2002

Labor scheduling with employee turnover and absenteeism

Fred F. Easton
*Syracuse University, School of Management, Robert H. Brethen Operations Management Institute*

John C. Goodale
*Ball State University, College of Business, Department of Management*

Follow this and additional works at: https://surface.syr.edu/mgt

Part of the Business Administration, Management, and Operations Commons

Recommended Citation
https://surface.syr.edu/mgt/5

This Article is brought to you for free and open access by the Whitman School of Management at SURFACE. It has been accepted for inclusion in Management by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.
Labor Scheduling with Employee Turnover and Absenteeism

Fred F. Easton  
Robert H. Brethen Operations Management Institute, School of Management  
Syracuse University  
Syracuse, NY 13244-2130  
(315) 443-3463; ffeaston@syr.edu

John C. Goodale  
Department of Management, College of Business  
Ball State University  
Muncie, IN 47306  
(765) 285-5323; jgoodale@bsu.edu

Working Paper SLSDL_MS_031502.doc

March 15, 2002

This project was partially supported by a research grant from the Center for Creation and Management of Digital Ventures at Syracuse University
Labor Scheduling with Employee Turnover and Absenteeism

Abstract

Most labor staffing and scheduling models presume that all employees scheduled for duty reliably report for work at the beginning of their shift. For industries with even moderate turnover or absenteeism, this assumption may be quite costly. We present a profit-oriented labor scheduling model that accounts for the day-to-day flux of employees and capacity induced by voluntary resignations, new hires, experience curves, and absenteeism. The proposed model also anticipates revenue losses due to reneging by customers whose patience decays exponentially with queue time. Our computational studies suggest that firms with comparatively high transaction volumes, long transaction times, and/or relatively tight profit margins may experience significant benefit from this approach. Compared with conventional labor scheduling models, the proposed method boosts average expected profits by more than 10 percent in certain operating environments.
Labor Scheduling with Employee Turnover andAbsenteeism

1. Introduction

A host of direct and indirect costs arise from the wake of each employee who voluntarily leaves an organization. Obvious expenses include the employer's recruiting, hiring, and training costs for a replacement employee. Until the vacancy is filled, employers may also face additional overtime costs, reduced productivity, increased customer queue times, lost sales and business opportunities, and the likelihood of additional turnover due to the extra work shouldered by coworkers of the departing employees (Herman, 1997; McConnel, 1999; Richardson, 1999).

Turnover is not just expensive; it is pervasive, arising in virtually all professions. Twenty percent of all new school-teachers and forty-four percent of all new lawyers quit within three years (Cooper, 2000; Flaherty, 1999). The average annual turnover rate (the number of annual resignations divided by the average workforce size) among call center employees is 31 percent (Karr, 1999), although larger operations (500 or more agents) average 61 percent/year. Zuber (2001), citing declining turnover rates for limited-service restaurant workers, reported that 1999 turnover averaged 123 percent in this industry.

Voluntary turnover problems also tend to be persistent and difficult to eliminate. Leonard's (1998) survey of human resource professionals found that 55 percent took measures to improve turnover, but only 10 percent reported noticeable improvements. Furthermore, many firms report that it takes 75 - 90 days to fill a vacant position (Fitz-enz, 1997; Matson, 1999), followed by weeks or months of training before a newly-hired employee becomes proficient.

While turnover results in permanent losses of human capital, unexpected absences due to illness or personal matters consume 2.5 percent of all scheduled work hours in the U.S. service
sector (Bureau of Labor Statistics, 2001). Some industries, however, suffer much higher absenteeism rates. Call center managers, for example, report losing 12 - 16 percent of all scheduled work hours to absenteeism (Call Center Ops, 2001).

The need to protect service delivery systems from turnover and absenteeism is well established. Turnover planning models have been developed for many industries and professions, including banking (Jones et al, 1973), engineering (Lapp & Thompson, 1974), law enforcement (Leeson, 1981), and the armed services (Charnes et al., 1972; Collins, et al, 1983; Eiger et al., 1988). Typically, these models anticipate the effects of turnover on existing staff and estimate the number of new employees that should be recruited into the organization each year at each grade to satisfy projected future staffing needs (Bartholomew, Forbes & McClean, 1991).

In this paper, however, we focus on the day-to-day operational impact of turnover and absenteeism, and developing techniques that mitigate their impact through short-term staffing and scheduling decisions. Our premise is that both service demand and employee availability are random variables. We model the day-to-day flux of employee resignations and new hires as a Markov process, then derive estimates for the probabilities of realizing different incumbency levels. From the underlying employee survivor function, we predict workforce experience levels and proficiencies. Modeling the service delivery system as a multi-server queue, we estimate revenue losses due to reneging when customer patience decays exponentially with waiting time. Finally, we devise a labor staffing and scheduling model that integrates workforce incumbency probabilities, experience levels, and random employee absenteeism with the object of maximizing expected profit under stochastic demand and impatient customers. The model determines the nominal workforce size, and how those employees should be deployed over time, to compensate for anticipated turnover and absenteeism.
To identify operating environments likely to benefit from this model, we compare its solutions with those from a conventional profit-oriented labor staffing and scheduling model that assumes stochastic service demand and deterministic labor supply. Our experiments reveal that even with modest turnover rates and brief job vacancies, conventional labor scheduling models consistently misstate ideal staffing levels and overstate expected profits and service levels. This limitation is more pronounced as a firm's average service rate or profit margin decreases, and as either its average turnover rate, mean job vacancy duration, absenteeism, or training time increases. In several operating environments, the proposed model improved expected profit by more than 10 percent over that earned with a conventional labor staffing and scheduling model.

The rest of the paper is organized in the following manner. In Section 2, we develop a framework for estimating workforce incumbency levels and experience levels. In Section 3, we characterize a multi-server queuing system with customers whose patience decays exponentially with queue time. In Section 4, we describe our labor staffing and scheduling model for stochastic service demand, impatient customers, and stochastic labor availability. In section 5, we describe our experiments to compare the proposed model with conventional techniques and discuss the results. Our summary of the project appears in Section 6.

2. Modeling turnover, absenteeism, and experience levels

Our goal in this research is to account for the stochastic nature of the labor supply in staffing and scheduling decisions for systems with random arrivals and impatient customers. We adopt the premise that employee resignations are in general random, independent events and that on average, fraction $\bar{T}$ of the total workforce will voluntarily leave the organization each year. We assume that once a vacancy occurs, the employer initiates efforts to recruit a replacement. After a random interval, a replacement worker arrives to staff the vacant position. We assume the new
worker's proficiency improves gradually with training and experience, approaching that of veteran staffers after a certain period of time. Finally, we assume that a random fraction of the employees scheduled for duty will fail to report due to unexpected illness. In the following subsections, we develop models to estimate: (1) the short-term impact of turnover on the firm's labor supply, (2) the effects of turnover on aggregate experience levels and productive capacity, and (3) the effects of absenteeism.

2.1 Turnover, recruitment, and workforce dynamics.

Let W be the nominal workforce size during the period of interest (W is an output of our proposed labor scheduling model, described in Section 4). In this section we estimate the probabilities that w = 0, …, W of those positions are occupied. If employees independently and randomly decide when to resign, the number of resignations/day should resemble a Poisson process. Let \( \gamma = T / 365 \) denote the expected fraction of the available workforce that resigns on a given day. For \( \gamma \) to remain stationary\(^1\), the number of voluntary departures will depend on the number of occupied positions. The Poisson probability that exactly k employees will resign during interval \( t \), given \( w \) positions are occupied, is:

\[
P(k \mid w) = (t \gamma w)^k e^{-t \gamma w} / k!.
\]  

When an employee resigns, a search for a qualified replacement begins. Let \( v \) be the number of days until the position is filled, and assume that \( v \) is exponentially distributed with mean \( \bar{v} \). At any instant \( 0 < w < W \) of the authorized positions might be occupied, depending on the flux of resignations and new recruits, with \( W - w \) active searches underway. On average \( (W - w) / \bar{v} \) vacant positions will be filled each day, so the Poisson probability that exactly k new recruits

\(^1\) Due to the increased workload and job stress for surviving workers, Tsui, Pearce, Porter, & Tripoli (1997) posit that the ratio of resignations to occupied positions tends to increase with the number of job vacancies.
report for duty during some interval $t$ is given by equation (1), substituting $(W - w)/\bar{v}$ for parameters $\gamma_w$. The probability of exactly one voluntary resignation during interval $h$ is:

$$P(1 \mid w) = (h \gamma_w)^1 e^{-h \gamma_w} / \gamma w \text{h as } h \to 0.$$  

(2)

The probability that exactly one new recruit arrives during interval $h$ also depends on the state variable $w$ and the number of authorized positions $W$. For state $w$, the probability that exactly one new recruit arrives during interval $h$ is $h(\gamma w - w)/\bar{v}$ as $h \to 0$.

The birth-death diagram in Figure 1 illustrates the flux of existing workers resigning and new workers recruited into the system. The nodes represent the number of occupied positions, ranging from 0 to $W$. The values within each node denote the probability of occupying that state. The directed arcs pointing to the right are the state-dependent probabilities of exactly one new recruit reporting for duty during interval $h$. The directed arcs pointing left are the state-dependent probabilities of exactly one employee resignation occurring during the same interval.

(please insert Figure 1 about here)

To maintain a stable workforce size, the flow leaving each state in Figure 1 must equal the flow into that state. The graph of this system is a tree, so it is reversible (Kelly, 1979) and its state probabilities can be estimated from the probability flux between any pair of adjacent states (Nelson, 1993). For each state, the following relationships must hold:

<table>
<thead>
<tr>
<th>State</th>
<th>Flow out = Flow in</th>
</tr>
</thead>
</table>
| 1     | $P_1 \gamma = P_0 W/\bar{v}$  
(3.1) |
| 2     | $P_2 2 \gamma = P_1 (W - 1)/\bar{v}$  
(3.2) |
| 3     | $P_3 3 \gamma = P_2 (W - 2)/\bar{v}$  
(3.3) |
| …    |                   |
| $W$   | $P_W W \gamma = P_{W-1} 1/\bar{v}$  
(3.W) |

Rearranging and substituting $P_0$ we have for state:
1. \[ P_1 = P_0 \frac{W}{\gamma v}, \] (4.1)

2. \[ P_2 = P_1 \frac{(W-1)}{(2\gamma v)} = P_0 \frac{W(W-1)}{2 \cdot 1\left(\gamma v\right)^2}, \] (4.2)

3. \[ P_3 = P_2 \frac{W-2}{3\gamma v} = P_0 \frac{W(W-1)(W-2)}{3 \cdot 2 \cdot 1\left(\gamma v\right)^3}, \] (4.3)

and in general,

\[ P_k = P_0 \frac{W!}{(W-k)!k!(\gamma v)^k} = P_0 \left(\frac{W}{k}ight) \left(\frac{1}{\gamma v}\right)^k. \] (4)

Since the state probabilities summed from 0 to W must equal 1, we require:

\[ P_0 \sum_{k=0}^{W} \left(\frac{W}{k}\right) \frac{1}{\left(\gamma v\right)^k} = 1. \] (5.1)

Exploiting the binomial expansion of \((1+x)^W\) (Beyer, 1981, p. 64), \(P_0\) simplifies to:

\[ P_0 = \left[ \sum_{k=0}^{W} \left(\frac{W}{k}\right) \frac{1}{\left(\gamma v\right)^k} \right]^{-1} = \left[ 1 + \frac{1}{\gamma v} \right]^{-1}. \] (5.2)

For a nominal workforce size \(W\), the average number of employees available for duty is:

\[ \overline{W} = \sum_{k=0}^{W} kP_k. \] (6)

In Table 1, we illustrate the effects of turnover at different average turnover rates \(\overline{T}\) and average vacancy durations \(\overline{v}\). For example, with a nominal workforce \(W = 10\), the chance of being fully staffed (i.e., \(P(k=10)\)) is 85 percent when annual turnover is 20% and positions remain vacant for an average of 30 days. The likelihood of at least one vacant position increases with turnover and the average duration of a vacancy. For \(\overline{T} = 60\%\) and \(\overline{v} = 60\) days, for example, the model predicts the organization will operate short-handed more than 60 percent of the time.

(please insert Table 1 about here)
2.2 Turnover and Aggregate Workforce Experience Levels

To maintain a stable workforce, new employees must be recruited to replace those who voluntarily leave the organization. Until they gain experience, those recruits may not be as productive as veteran workers. To illustrate, suppose newly-hired employees achieve full productivity after \( \tau \) days of experience, and that during training, they process transactions at average rate \( \mu_1 \). After \( \tau \) periods of experience, their average service rate increases to \( \mu_2 \). Since average employee tenure is \( 1/T \), and employees randomly and independently decide when they will resign, the probability that an employee will remain with the organization for at least \( \tau \) days may follow the exponential distribution. Under this distribution, the likelihood of remaining with the firm for at least \( \tau \) varies inversely with both \( \frac{1}{T} \) and \( \tau \).

While Bartholomew et al (1991) suggest that each employee cohort may have its own characteristic turnover rate, we shall assume for now this rate is identical for all employees. Let \( (1-\phi) \) be the fraction of employees who resign during the interval \([0, \tau]\). Expressed as a survivor function, the expected fraction of new hires who reach full productivity (or the fraction of veteran employees likely to still be working \( \tau \) periods from now), is:

\[
\phi = p(x > \tau) = e^{-\tau T}.
\]

For example, with \( \frac{1}{T} = 0.6/\text{yr} \) and \( \tau = 2 \) months, \( \phi \) is about 90 percent. During any one training cycle, we should expect \((1-\phi)W\) resignations. If the organization plans to maintain a workforce of size \( W \), on average \((1-\phi)W\) new employees will always be in training. The average processing time per transaction \( \mu^{-1} \) depends on the expected mix of experienced and inexperienced employees in the workforce, so:

\[
\frac{1}{\mu} = (1-\phi)\mu_1 + \phi \mu_2.
\]
2.3 Absenteeism

Even those employees who fully intend to continue their relationship with the organization occasionally get sick or must attend to personal business during the times they are scheduled to work. In 2000, U.S. industry lost 2.0% of all scheduled work hours due to unscheduled absences (Bureau of Labor Statistics, 2001), and 2.5% in the service sector. For industries like call centers, however, anecdotal reports of 12 - 16 percent absenteeism are not unusual (Call Center Ops, 2001). CCH (2000) reports that 41 percent of the HR professionals responding to their annual survey now view unscheduled absenteeism as a serious problem in their organizations. Furthermore, few of the managers surveyed expect absenteeism to improve in the near future.

Mitchell (2001) observed that the direct costs of absenteeism are a significant component of total payroll expense. For example, employee benefit plans often indemnify workers from incidental absence and short-term illness. However, unscheduled absenteeism also adversely affects productivity, especially when compounded by employee turnover. Let $A$ be the ratio of scheduled labor hours lost to unplanned absenteeism divided by total scheduled hours. Although $H$ workers may have been originally scheduled for duty at a particular time, turnover and absenteeism will reduce the expected number of workers who actually report for duty to $H(1 - A)\bar{W}/W$.

3. Expected loss due to reneging in $M/M/S$

Most workforce management systems use $M/M/S$ queueing models (Erlang C) to estimate customer delays for different arrival rate/staffing level combinations (Durr, 1994; Cleveland, 1999). $M/M/S$ assumes queued customers patiently advance in line once every $1/\mu$ time units, on average. In practice, their patience often decays exponentially with average waiting time.
(Sasser et al, 1991; Whitt, 1999b) and some customers may renege (abandon the queue) before service begins.

While reneging customers speed up the line for others still in the queue, their behavior reduces the effective arrival rate (\( \lambda_{\text{eff}} \)) and sales/period. We can estimate \( \lambda_{\text{eff}} \) and lost revenue from the definition for server utilization. Let \( F_k \) = probability that the M/M/S: reneging system is occupied by \( k \) customers. With \( s \) servers each operating at mean service rate \( \mu \), server utilization \( U_{\text{eff}} \) is:

\[
U_{\text{eff}} = \frac{\lambda_{\text{eff}}}{s \mu} = \sum_{i=0}^{s} \left( \frac{i}{s} \right) F_i + \sum_{j=s+1}^{\infty} \frac{s}{s} F_j = \sum_{i=0}^{s} \left( \frac{i}{s} \right) F_i + \left( 1 - \sum_{j=0}^{s} F_j \right),
\]

so the effective arrival rate \( \lambda_{\text{eff}} \) must equal \( U_{\text{eff}} \mu \).

Reneging behavior impacts total revenue, and is therefore an important consideration for staffing and scheduling decisions. Unfortunately, the recent literature contains surprisingly few references on this topic. Boots and Tijms (1999) estimate loss probabilities for systems where customers renege after waiting a fixed time period. For systems that serve a population of humans, however, it is more likely that patience varies from customer to customer. Gross and Harris (1998, p. 94) outlined a single-server model where the propensity to renege increases exponentially with queue time. Below, we extend that notion to the multi-server case.

If customers are willing to wait an average of \( \alpha^{-1} \) before reneging (exponentially distributed), the probability that queue time exceeds \( T \) is \( P(q>T) \approx e^{(-\alpha T)} \). From the relationship between the Poisson and the exponential, \( \alpha T e^{(-\alpha T)} \) is the probability that exactly one customer reneges during interval \( T \). As \( T \) approaches zero, this probability approaches \( \alpha h \). The birth-death diagram in Figure 2 illustrates the dynamics of an M/M/S queue with reneging. Each circle represents a possible system state, corresponding to the number of customers in the system. The arcs
pointing right indicate the arrival of a new customer, with the probability of exactly one Poisson arrival during a brief instant of time shown above the arc. The arcs pointing left represent the flow of customers departing a given state, which could be due to a Poisson service completion (at rates up to \( s \mu \)) or a reneging customer. Departures from state \( k>s \) occur if any of the customers at queue positions 1 to \( k-s \) renege during interval \( h \), so the departure probabilities for state \( k \) must include the chance of a single service completion plus the summed reneging probabilities from any queue position up to and including position \( k-s \).

(please insert Figure 2 about here)

The birth-death diagram forms a tree, so the system is reversible. If the system is stationary, conservation of flow requires:

\[
\text{State} \quad \text{Flow out} \quad = \quad \text{Flow in} \tag{10}
\]

<table>
<thead>
<tr>
<th>State</th>
<th>Flow out</th>
<th>= Flow in</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( F_1 \mu )</td>
<td>= ( F_0 \lambda )</td>
</tr>
<tr>
<td>2</td>
<td>( F_2 2\mu )</td>
<td>= ( F_1 \lambda )</td>
</tr>
<tr>
<td>3</td>
<td>( F_3 3\mu )</td>
<td>= ( F_2 \lambda )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>s</td>
<td>( F_s s\mu )</td>
<td>= ( F_{s-1} \lambda )</td>
</tr>
<tr>
<td>( s+1 )</td>
<td>( F_{s+1} (s\mu + \alpha) )</td>
<td>= ( F_s \lambda )</td>
</tr>
<tr>
<td>( s+2 )</td>
<td>( F_{s+2} (s\mu + 2\alpha) )</td>
<td>= ( F_{s+1} \lambda )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

The birth-death diagram for the system has two distinct sections: the portion from state 0 to state \( s \), where the probabilities of a single service completion steadily increase with occupancy; and the portion from state \( s \) to \( \infty \), where all servers are busy and departures may be due to either a service completion or a reneging customer. For a stationary system, the probability of occupying state \( k \) in an M/M/s system with reneging is \( F(k) \), where:
The state probabilities must sum to 1.0. Thus:

$$F_k = \left\{ \begin{array}{ll} F_{k-1} \left( \frac{\lambda}{k \mu} \right) = F_0 \left( \frac{\lambda}{\mu} \right)^k \frac{1}{k!}, & \text{for } 0 < k \leq s \\ F_{k-1} \left( \frac{\lambda}{s \mu + (k-s) \alpha} \right) = F_0 \left( \frac{\lambda}{\mu} \right)^{s} \prod_{i=1}^{i=s} \left( \frac{\lambda}{s \mu + i \alpha} \right) & \text{for } s < k \leq \infty \end{array} \right.$$  \hspace{1cm} (11)

The infinite sum in (13) converges to zero rapidly. While a closed form solution for (13) may exist, we find that an efficient approximation technique such as Newton's method and the relationship in (12) can be used to estimate $F_0$ with adequate precision.

$$\sum_{k=0}^{\infty} F_k = \sum_{i=0}^{s} F_i + \sum_{j=s+1}^{\infty} F_j = F_0 \left[ \sum_{i=0}^{s} \left( \frac{\lambda}{\mu} \right)^i \frac{1}{i!} + \left( \frac{\lambda}{\mu} \right)^{s} \sum_{j=s+1}^{\infty} \prod_{i=1}^{i=s} \frac{\lambda}{s \mu + i \alpha} \right] = 1.0$$  \hspace{1cm} (12)

and therefore:

$$F_0 = \left\{ \frac{\sum_{i=0}^{s} \left( \frac{\lambda}{\mu} \right)^i + \left( \frac{\lambda}{\mu} \right)^{s} \sum_{j=s+1}^{\infty} \prod_{i=1}^{i=s} \frac{\lambda}{s \mu + i \alpha} }{\sum_{i=0}^{s} \prod_{k=1}^{i=s} \frac{\lambda}{s \mu + k \alpha}} \right\}^{-1}$$  \hspace{1cm} (13)

The infinite sum in (13) converges to zero rapidly. While a closed form solution for (13) may exist, we find that an efficient approximation technique such as Newton's method and the relationship in (12) can be used to estimate $F_0$ with adequate precision.

4. Employee Staffing and Scheduling Decisions

Many service organizations with turnover and absenteeism problems face consumer demands that vary from hour-to-hour and day-to-day. Their production capacity is difficult to store, and they risk lost sales when their services cannot be produced upon demand. Furthermore, their service capacity decisions, in the form of employee work schedules, are usually made long before that demand is realized.

Employee scheduling problems have been modeled as deterministic generalized set covering problems (Bailey, 1985; Dantzig, 1954) and deterministic goal programs (Andrews & Parsons,
1989; Baker, 1976; Keith, 1979; Mabert & Watts, 1982). In both models, target staffing requirements for each period are exogenous parameters. In their labor scheduling models for stochastic demand (SLS/D), Thompson (1995b) and Easton & Rossin (1996) integrated labor requirements planning with staffing and scheduling decisions, explicitly accounting for the stochastic nature of service demand.

All of these models assume deterministic employee availability. Since turnover and absenteeism reduce the likelihood that the system will be fully staffed, their solutions may overstate labor expenses. Because some employees scheduled for duty at a particular time may have resigned or be absent, customer queues at those times may be longer, and more customers may renege, than these models anticipate. Thus, existing labor staffing and scheduling models may also overestimate revenues.

These limitations motivate our proposed model for labor staffing and scheduling decisions under stochastic demand and labor supply (SLS/DL). The model assumes that each completed transaction earns an incremental contribution (in dollars) before labor expense, and that customer arrivals/hour and hourly service rates per employee are random variables. In addition, the model assumes customers have limited patience and will abandon the queue after waiting an exponentially distributed amount of time. What makes this model distinctive, however, is that it adjusts the workforce size and the number of workers scheduled for duty at a particular time to account for turnover, the mix of trainees and veteran employees, and absenteeism. To fully describe our labor staffing and scheduling model for variable demand and turnover, we rely on the following notation:
**Model Parameters**

**Workforce Characteristics**
- $\overline{T}$ = average annual turnover rate (number of voluntary resignations/yr ÷ average workforce size), Poisson-distributed
- $\nu$ = average vacancy duration (days), distributed exponential;
- $A$ = average absenteeism rate; total hours lost due to unexpected absences ÷ total scheduled and paid hours;
- $\tau$ = length of the training period for new employees;
- $\mu_1$ = mean service rate for employees in training (Poisson-distributed);
- $\mu_2$ = mean service rate for experienced employees (Poisson-distributed);
- $\mu$ = aggregate service rate (see equations (7) and (8)).

**Customer and Market Characteristics**
- $\lambda_i$ = average customer arrival rate forecasted for period i (Poisson-distributed);
- $\alpha^{-1}$ = average customer patience, exponentially distributed.
- $R_i$ = average contribution per completed customer transaction, before labor costs, for period i.

**Schedule Characteristics**
- $I$ = set of time periods in planning horizon, indexed $i=1,\ldots,I$;
- $K$ = set of allowable work schedules for employees, indexed $k=1,\ldots,K$;
- $a_{ik} = 1$ if $i$ is a working period in schedule $k$, 0 otherwise, for $i=1,\ldots,I$;
- $k=1,\ldots,K$;
- $C_k$ = Wage cost to assign one employee to schedule $k$.

**Decision Variables and Consequence Variables**
- $X_k$ = number of employees assigned to schedule $k$, $\forall k \in K$;
- $P_k$ = probability of exactly $k$ incumbent employees (equations (4.6) and (5.6));
- $W$ = nominal workforce size
- $\bar{w}_i$ = expected number of employees who actually report for duty in period $i$;
- $F_k$ = probability of exactly $k$ customers in M/M/S: reneging system (equations (11) and (13));
- $\lambda_{effi}$ = effective customer arrival rate for period $i$, adjusted for impatient customers who abandon the queue before initiating service (equation (9)).
The object of our stochastic labor scheduling model with labor turnover and absenteeism (SLS/DL) is to maximize expected profits. Using the above notation, we determine the number of employees who should be assigned to each feasible work schedule that will:

\[
\text{SLS/DL: Maximize } Z = \sum_{i \in I} \lambda_{\text{eff}} R_i - \left( \frac{W}{\overline{W}} \right) \sum_{k \in K} C_k X_k
\]

subject to:

\[
W = \sum_{k \in K} X_k
\]

\[
\overline{W} = \sum_{j=1}^{W} j P_j
\]

\[
w_i = \sum_{k \in K} a_{ik} x_k , \forall \ i \in I; \quad (16)
\]

\[
\overline{w}_i = w_i \left(1 - A \left( \frac{\overline{W}}{W} \right) \right) , \forall \ i \in I; \quad (17)
\]

\[
\lambda_{\text{eff}} = \mu \left[ \sum_{k=0}^{\overline{w}_i} k F_k + \overline{w}_i \left(1 - \sum_{k=0}^{\overline{w}_i} F_k \right) \right] , \forall \ i \in I; \quad (18)
\]

\[x_k \geq 0 , \text{ and integer, } \quad k \in K. \quad (19)
\]

The objective function (13) computes the expected contribution for the solution by determining expected revenues and expenses. The first term calculates the product of the effective number of customer arrivals and the marginal contribution per arrival. The second term computes the expected labor expenses for the solution, after adjusting for turnover (the ratio $\overline{W}/W$ is the fraction of occupied positions). Note that labor expense includes the expected wages for employees with unplanned absences.

Equation (14) is a consequence variable that determines the nominal workforce size. Equation (15) adjusts this total for expected turnover. Equation (16) calculates the nominal
number of employees scheduled for duty during period i. Lacking more complete information about individual workers, we assume that employee resignations and absences are dispersed uniformly throughout the workforce. Equation (17) adjusts the number of employees scheduled for duty during each period to reflect uniform turnover and absenteeism.

Equation (18) determines the effective arrival rate, adjusting the nominal arrival rate for customers who abandon the queue. This relationship is adapted from equation (9). The state probabilities $F_k$ used in equation (18) are based on those described in equations (10) and (13).

SLS DL is similar to SLS D when $p(W) = 1.0$ and $A = 0.0$ (i.e., when turnover and absenteeism are zero). Both models include non-linear and discontinuous terms, and it is often difficult to obtain provably optimal solutions. Fortunately, meta-heuristics like simulated annealing (Brusco & Jacobs, 1998) and hybrid genetic algorithms (Easton & Mansour, 1999) work well with this class of problems. However, unless labor scheduling decisions that account for employee turnover and absenteeism provide significant economic benefits for decision makers, there is little reason to deploy such models. In the next section, we hypothesize the types of operating environments where the advantages of SLS DL are likely to be noteworthy. We then describe our experiments to ascertain the validity of those hypotheses.

5. SLS DL Model Assessment

Turnover and absenteeism reduce effective service capacity. As the intensities of these factors increase, labor staffing and scheduling models that anticipate their impact should outperform approaches that ignore them. However, other operating factors may either heighten or lessen their economic impact. For example, if scheduling policies and consumer demand patterns interact to provide significant surplus labor, employee turnover or absenteeism may actually reduce operating expenses without adversely affecting revenues.
To help identify the market and service delivery system characteristics where SLS/DL is likely to provide a significant performance advantage over SLS/D, we simulated 1024 different operating environments. The characteristics of these labor staffing and scheduling environments are summarized in Table 2. For each environment, we modeled the labor scheduling problem as both an SLS/D problem and as an SLS/DL problem, then obtained near-optimal solutions for each. Finally, we computed the performance improvement attributable to the SLS/DL model.

(please insert Table 2 about here)

SLS/DL addresses four aspects of labor supply uncertainty that have been ignored in earlier labor scheduling models: turnover rate, average vacancy time, the time for new employees to achieve proficiency, and absenteeism. To assess the effectiveness of SLS/DL, we simulated labor environments with turnover rates $\bar{T} = 0.0, 0.2, \text{and } 0.6$; job vacancy times of $\bar{v} = 30$ and $\bar{v} = 60$ days; and absenteeism rates of $A = 0.00$ and 0.05/scheduled hour. The turnover rates are similar to call center averages reported by Karr (1999). The average job vacancy times are similar to those reported by Fitz-enz (1997). The simulated absenteeism rates straddle the US service sector average (BLS 2001), but are probably much lower that those in some industries.

Task complexity affects the level of training required for new hires, and in high-turnover labor environments may have a significant impact on productivity. Therefore, our simulations included two levels for task complexity and new-hire training: 0 days with initial productivity of 100% of veteran employees, and 20 days, with initial productivity averaging 60%.

Systems with a small number of high-speed servers tend to have shorter mean queue times than systems with identical capacity that use more, but slower, servers (Hillier & Lieberman, 1995). With more workers subject to resignation and absenteeism, systems with slower servers may lose more revenue than fast-server systems due to increases in average customer queue
times and reneging rates. Thus these slower-server systems are more likely to benefit from the SLS/DL model than systems with high service rates. Guided by earlier studies (Andrews & Parsons, 1993 and Thompson (1997)), we simulated systems with average service rates of $\mu = 4/hr$ and $8/hr$ per employee.

Potential gains attainable with SLS/DL may be influenced by customer arrival characteristics. Like Thompson (1995a, 1997), we simulated daily demand patterns with different numbers of demand peaks per day (unimodal and trimodal), different amplitudes (COV = 0.25 and 0.50), and arrival rates (400/hr and 800/hr). Customer patience may also influence the improvements attainable with the SLS/DL model. Less patient customers are more apt to renge when queue times increase due to turnover and absenteeism. Therefore, we simulated systems with average customer patience of 120 seconds and 240 seconds. Finally, as profit margins narrow, workforce utilization levels are often driven higher. Systems with high worker utilization may be more vulnerable to loss from absenteeism and turnover, because they have smaller capacity buffers. In our study, we simulated average revenue/transaction values of $7.00 and $14.15 (adapted from Andrew and Parsons, 1993).

**Other parameters:**

*Full- and part-time employees.* We allowed solutions with a mix of full- and part-time employees, provided at least 50% of all employees were full-timers (Easton & Rossin, 1991)

*Hourly wage rate.* We assumed hourly wage rates for full- and part-time employees were $15.00 and $12.00, respectively.

*Expected orders/arrival:* We assumed 80% of all transactions produced income. The balance were transactions such as status inquiries or information requests.
We conducted a full-factorial experiment by first solving SLS/D problems for the 128 different combinations of service delivery features and market characteristics described in Table 2. These solutions, obtained with Easton & Mansour's (1999) distributed genetic algorithm, provided benchmarks for assessing the improvements attainable with SLS/DL. The solution procedure converged to answers that were on average within 0.08 percent of a computed bound.

SLS/D ignores turnover and absenteeism, so a solution for a problem with 60 percent annual turnover is identical to one with zero turnover. To complete our benchmarks, we "priced" each SLS/D solution with the SLS/DL objective function assuming the different turnover rates, job vacancy times, absenteeism levels, and training times described in Table 2. Finally, we solved each problem as an SLS/DL problem and compared its expected profit with the profit for the corresponding benchmark SLS/D solution.

We then regressed the experimental factors against SLS/DL performance to identify the environmental characteristics where SLS/DL could be expected to provide significant benefit over existing labor scheduling methods. The results, reported in Table 3, indicate that all 10 attributes play a significant role in explaining the improvements obtained with SLS/DL. As expected, the regression coefficients confirm that increasing absenteeism and employee turnover seriously degrade SLS/D schedule performance, creating opportunities for SLS/DL to find alternative solutions with significantly greater profits.

(Please insert Table 3 about here)

Under high turnover and absenteeism, SLS/DL returned an average of 4.9% more profit than the SLS/D solutions (see Table 4). However, SLS/DL achieved better than average results in operating environments with comparatively high customer arrival rates, low profit margins, and more patient customers. We also found larger than average profit increases when new workers
require more initial training and are less productive than veteran employees. In Table 4, we "drill down" through the various market attributes, demand attributes, and service delivery attributes to zero in on the operating characteristics where SLS/DL provided the greatest benefit.

(please insert Table 4 about here)

In Table 4.A, we isolate the market attributes where SLS/DL produces above-average improvements. As the regression coefficients in Table 3 indicate, the SLS/DL's performance improves with customer arrival rates, decreases with revenue per order, and improves with customer patience. With standard labor scheduling models, worker utilization tends to increase with transaction volume. The loss of a busy worker to absenteeism or turnover has a greater impact on revenue than the loss of a worker with lower utilization. A similar effect occurs with tighter profit margins (or less revenue/order). To maximize profitability, SLS/D tries to squeeze surplus capacity from the system. It is thus more vulnerable to revenue losses when absenteeism and turnover increase queue times. With greater average customer patience, SLS/D allows longer queues. When absenteeism and turnover deplete capacity, queued customers tend to wait longer than expected and are more likely to renege, thereby reducing revenue and profits.

Table 4.B focuses on the pattern of service demand, decomposing the column of Table 4.A with the greatest average improvement. With multiple daily demand peaks, work rules about minimum shift length make it difficult to avoid significant overstaffing. These systems thus tend to be less vulnerable to absenteeism and turnover, leaving few opportunities for SLS/DL to improve on the SLS/D solutions. As the cell averages in Table 4.B and the regression coefficients of Table 3 show, SLS/DL's advantages decrease with the number of daily demand peaks. On the other hand, SLS/DL apparent performance advantages increase slightly with
demand "peakiness." Doubling the amplitude of demand leads to small, but significant performance improvements.

In Table 4.C we decompose the column from Table 4.B with the greatest gains to examine how service delivery system characteristics influence SLS/DL's performance. Generally, we find SLS/DL's advantages are more pronounced in slower systems, which tend to have more employees subject to turnover and absenteeism. Thus SLS/D solutions for slower systems tend to lose a larger number of servers than fast systems with comparable capacity, leading to a greater percentage increase in average queue times (Hillier & Lieberman, 1995). Increased task complexity (training time), coupled with higher absenteeism and turnover, reduces the effective service rate. The SLS/D solutions tend to have longer than expected queues when absenteeism and turnover strike, reducing revenues as customers become impatient and renege.

6. Discussion

This paper characterized the impact of employee turnover and absenteeism on workforce availability and integrated those results into a stochastic labor staffing and scheduling model we labeled SLS/DL. We completed a series of experiments to help isolate the operating environments most likely to benefit from this approach. Our preliminary tests indicate that firms with low turnover, low absenteeism, short average vacancy times, comparatively high unit revenues, and multiple daily demand peaks benefit least from the SLS/DL model. In contrast, firms with comparatively long transaction times, high customer arrival rates, relatively tight profit margins, more patient customers, and greater task complexity (training time) should experience significant benefits from this approach.

These results suggest a number of directions for future research. For example, we assumed interchangeable employees with common skill sets, scheduling preferences, availability and
reliability. In future research, we hope to apply this methodology to systems with extant, heterogeneous employees. We also assumed single stage processes in fairly large-scale systems. For smaller, lower volume operations or systems with a sequence of operations, it may be more difficult to realize performance advantages of the magnitude reported here. Finally, SLS/DL considers only one strategy for coping with turnover -- overstaffing. Using other coping strategies, such as temporary workers, overtime, reassigning cross-trained employees, etc., may provide additional benefits.

7. References


Berman, D. 2000. "We've got to get our hands on some workers." Business Week (January 31), F.4


Management Science. 37(11), 1441-1451.


Herman, R.F. 1999. Hold on to the people you need. HRFocus. 76(6), pg. S11.


Tsui; Anne S. J. Pearce; L. Porter; and A. Tripoli; 1997. Alternative approaches to the employee-organization relationship: Does investment in employees pay off? *Academy of Management Journal*; 40(5) 1089-1121


Figure 1: Birth-death diagram for worker flux under state-dependent resignations and recruitment
Figure 2.
Birth-death diagram for M/M/s with reneging

\[ \lambda h \rightarrow \mu h \rightarrow \sigma h \rightarrow \sigma h + \alpha h \rightarrow \sigma h + 2\alpha h \rightarrow \sigma h + 3\alpha h \]
Table 1: State probabilities for number of employees available for duty when nominal workforce size \( W = 10 \)

<table>
<thead>
<tr>
<th>Number of Available Employees</th>
<th>( \bar{\nu} = 30 \text{ days} )</th>
<th>( \bar{\nu} = 60 \text{ days} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{T} = 0.2/\text{yr} )</td>
<td>( P_k )</td>
<td>( P_k )</td>
</tr>
<tr>
<td>0</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>1</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>2</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>3</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>4</td>
<td>0.000000</td>
<td>0.000001</td>
</tr>
<tr>
<td>5</td>
<td>0.000001</td>
<td>0.000023</td>
</tr>
<tr>
<td>6</td>
<td>0.000026</td>
<td>0.000350</td>
</tr>
<tr>
<td>7</td>
<td>0.000628</td>
<td>0.004263</td>
</tr>
<tr>
<td>8</td>
<td>0.011463</td>
<td>0.038901</td>
</tr>
<tr>
<td>9</td>
<td>0.139466</td>
<td>0.236650</td>
</tr>
<tr>
<td>10</td>
<td>0.848417</td>
<td>0.719811</td>
</tr>
</tbody>
</table>

\( \bar{\nu} = \text{Avg. Workforce Size} = \)

\[ \begin{array}{cccccc}
\end{array} \]
<table>
<thead>
<tr>
<th>Experimental Factor</th>
<th>Number of Levels</th>
<th>Factor Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Employee Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turnover rate</td>
<td>3</td>
<td>0, 20%, and 60% per year</td>
</tr>
<tr>
<td>Average job vacancy time</td>
<td>2</td>
<td>30 and 60 days</td>
</tr>
<tr>
<td>Absenteeism rate</td>
<td>2</td>
<td>0 and 5.0%</td>
</tr>
<tr>
<td><strong>Service Delivery Features</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average service rate for fully-trained employees</td>
<td>2</td>
<td>$\mu_2=8$ and 4 transactions per hour</td>
</tr>
<tr>
<td>Initial training period/trainee productivity</td>
<td>2</td>
<td>0 days @ 100% $\mu_2$ and 20 days @ 60% $\mu_2$</td>
</tr>
<tr>
<td><strong>Market Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average arrival rate</td>
<td>2</td>
<td>$\lambda=400$ and 800 arrivals/hour,</td>
</tr>
<tr>
<td>Daily demand pattern</td>
<td>2</td>
<td>Unimodal and trimodal sinusoidal patterns</td>
</tr>
<tr>
<td>Amplitude of demand</td>
<td>2</td>
<td>Coefficient of variation (COV) of 0.25 and 0.50; sine function amplitudes of 0.353 and 0.706</td>
</tr>
<tr>
<td>Unit contribution before labor</td>
<td>2</td>
<td>$7.00$ and $14.15$ per order (80% of all arrivals are orders, 20% are non-revenue producing)</td>
</tr>
<tr>
<td>Average customer patience</td>
<td>2</td>
<td>120 seconds and 240 seconds, exponentially distributed</td>
</tr>
</tbody>
</table>
### Table 3: Regression Coefficients for Environmental Factors Influencing SLS/DL Performance

<table>
<thead>
<tr>
<th>Attribute Levels</th>
<th>Coefficients</th>
<th>t Statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average arrival rate/hr</td>
<td>400/hr</td>
<td>800/hr</td>
<td>1.215E-05</td>
</tr>
<tr>
<td>Demand Amplitude</td>
<td>0.353</td>
<td>0.706</td>
<td>-0.0070</td>
</tr>
<tr>
<td>Number of Demand Peaks/day</td>
<td>1</td>
<td>3</td>
<td>-0.0019</td>
</tr>
<tr>
<td>Average Revenue</td>
<td>$7.00/order</td>
<td>$14.15/order</td>
<td>-0.0006</td>
</tr>
<tr>
<td>Average Patience</td>
<td>120 sec</td>
<td>240 sec</td>
<td>1.192E-05</td>
</tr>
<tr>
<td>Average Service Rate</td>
<td>4/hr per employee</td>
<td>8/hr per employee</td>
<td>-0.0021</td>
</tr>
<tr>
<td>Mean Training Time/Proficiency</td>
<td>0 days @ 100%</td>
<td>20 days @ 60%</td>
<td>-0.0121</td>
</tr>
<tr>
<td>Average Turnover Rate</td>
<td>0.2/year</td>
<td>0.6/year</td>
<td>0.0482</td>
</tr>
<tr>
<td>Average Vacancy Time</td>
<td>30 days</td>
<td>60 days</td>
<td>0.0004</td>
</tr>
<tr>
<td>Mean Absenteeism</td>
<td>0% scheduled hrs/day</td>
<td>5% scheduled hrs/day</td>
<td>0.2987</td>
</tr>
</tbody>
</table>
### Table 4: Drill-down Relationships: Environmental Factors and SLS/DL Performance

#### A. Market Attributes

<table>
<thead>
<tr>
<th>Arrival rate</th>
<th>400/hr (80% orders)</th>
<th>800/hr (80% orders)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue/Order</td>
<td>$7</td>
<td>$14.15</td>
</tr>
<tr>
<td>Avg. Patience</td>
<td>120s</td>
<td>240s</td>
</tr>
<tr>
<td>Tbar</td>
<td>Vbar</td>
<td>Absent</td>
</tr>
<tr>
<td>0.2/yr</td>
<td>30 days</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>60 days</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>1.47</td>
</tr>
<tr>
<td>0.6/yr</td>
<td>30 days</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>2.29</td>
</tr>
<tr>
<td></td>
<td>60 days</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>4.53</td>
</tr>
</tbody>
</table>

### B. Demand Attributes

Arrivals=800/hr, Revenue=$7/order, Patience=240s

<table>
<thead>
<tr>
<th>Demand Peaks/day</th>
<th>Demand Amplitude</th>
<th>low</th>
<th>high</th>
<th>low</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tbar</td>
<td>Vbar</td>
<td>Absent</td>
<td><strong>SLS/DL Profit Increase (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2/yr</td>
<td>30 days</td>
<td>0.0%</td>
<td>0.41</td>
<td>0.31</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td>1.88</td>
<td>2.10</td>
<td>1.39</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>60 days</td>
<td>0.0%</td>
<td>0.71</td>
<td>0.77</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td>2.67</td>
<td>3.03</td>
<td>2.05</td>
<td>1.36</td>
</tr>
<tr>
<td>0.6/yr</td>
<td>30 days</td>
<td>0.0%</td>
<td>1.65</td>
<td>1.73</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td>4.04</td>
<td>4.51</td>
<td>3.18</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>60 days</td>
<td>0.0%</td>
<td>4.20</td>
<td>4.53</td>
<td>3.38</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td>7.48</td>
<td>8.21</td>
<td>6.09</td>
<td>4.28</td>
</tr>
</tbody>
</table>

### C. Service Delivery Attributes

Arrivals=800/hr, Revenue=$7/order, Patience=240s

<table>
<thead>
<tr>
<th>Service Rate</th>
<th>Peaks/day =1, Amplitude = high</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/hr</td>
<td>8/hr</td>
</tr>
<tr>
<td>Training time/Initial Proficiency</td>
<td>0 day/100%</td>
</tr>
<tr>
<td>Tbar</td>
<td>Vbar</td>
</tr>
<tr>
<td>0.2/yr</td>
<td>30 days</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
</tr>
<tr>
<td></td>
<td>60 days</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
</tr>
<tr>
<td>0.6/yr</td>
<td>30 days</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
</tr>
<tr>
<td></td>
<td>60 days</td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
</tr>
</tbody>
</table>