

2010

Intelligent Procedures for Intra-Day Updating of Call Center Agent Schedules

Vijay Mehrotra

University of San Francisco, vmehrotra@usfca.edu

O. Ozluk

R. Saltzman

Follow this and additional works at: <http://repository.usfca.edu/at>

 Part of the [Business and Corporate Communications Commons](#)

Recommended Citation

Vijay Mehrotra, Ozgür Ozlük and Robert Saltzman. Intelligent Procedures for Intra-Day Updating of Call Center Agent Schedules. *Production and Operations Management*. Volume 19, Issue 3, pages 353–367, May/June 2010. <http://dx.doi.org/10.1111/j.1937-5956.2009.01097.x>

This Article is brought to you for free and open access by the School of Management at USF Scholarship: a digital repository @ Gleeson Library | Geschke Center. It has been accepted for inclusion in Business Analytics and Information Systems by an authorized administrator of USF Scholarship: a digital repository @ Gleeson Library | Geschke Center. For more information, please contact repository@usfca.edu.

INTELLIGENT PROCEDURES
FOR INTRA-DAY UPDATING
OF CALL CENTER AGENT SCHEDULES

Vijay Mehrotra
Department of Finance and Quantitative Analytics
School of Business and Management
University of San Francisco
San Francisco CA 94117
vmehrotra@usfca.edu
Phone: (415) 817 4357
Fax: (415) 405 0364

Özgür Özlük
Department of Decision Sciences
College of Business
San Francisco State University
San Francisco CA 94132
ozgur@sfsu.edu
Phone: (415) 817 4353
Fax: (415) 405 0364

Robert Saltzman
Department of Decision Sciences
College of Business
San Francisco State University
San Francisco CA 94132
saltzman@sfsu.edu
Phone: (415) 338 1232
Fax: (415) 405 0364

ABSTRACT

For nearly all call centers, agent schedules are typically created several days or weeks prior to the time that agents report to work. After schedules are created, call center resource managers receive additional information that can affect forecasted workload and resource availability. In particular, there is significant evidence, both among practitioners and in the research literature, suggesting that actual call arrival volumes early in a scheduling period (typically an individual day or week) can provide valuable information about the call arrival pattern later in the same scheduling period.

In this paper, we develop a flexible and powerful heuristic framework for managers to make intra-day resource adjustment decisions that take into account updated call forecasts, updated agent requirements, existing agent schedules, agents' schedule flexibility, and associated incremental labor costs. We demonstrate the value of this methodology in managing the trade-off between labor costs and service levels to best meet variable rates of demand for service, using data from an actual call center.

Keywords: call center operations, doubly stochastic processes, workforce scheduling, integer programming, real time optimization

Date of Original Submission:	January 3, 2006
Date of Submission of First Revision:	January 21, 2007
Date of Submission of Second Revision:	May 7, 2008
Date of Submission of Third Revision:	December 29, 2008
Date of Acceptance:	April 15, 2009

1. INTRODUCTION

Effective and efficient operation of a telephone call center is largely dependent on strong workforce planning and management (Mehrotra 1997; Cleveland and Mayben 1997), as typically 60-80% of a call center's budget is spent on labor costs. As illustrated in Figure 1, this workforce planning and management involves three levels of decision-making: long term planning, short term scheduling, and real time schedule adjustments (Abernathy et al. 1973). Long term planning decisions address how many agents to hire and train at what times. These decisions are typically made 6-12 months ahead of time and take into account aggregate call forecasts, agent availability and productivity assumptions, and anticipated staff attrition rates. Short term scheduling decisions determine which agents are assigned to work on which shifts on which days at which times over the course of a scheduling period (typically one week). In most call center environments, scheduling is typically done 1-2 weeks ahead of time, with schedules communicated to individual agents so that they can plan accordingly. Scheduling decisions are based on estimated agent requirements (which in turn are based on detailed forecasts for call arrival patterns and service times as well as customer waiting time goals), shift definitions and restrictions, agent rosters and shift preferences, and absenteeism assumptions.

By contrast, real time schedule adjustments are made *after* agents have been hired and trained, agent schedules have been created, and additional details have become available about call volumes, absenteeism, and unanticipated off-phone activities such as training and meetings. Adjustments are then made on an intra-day basis to agents' schedules. There is typically a limited set of feasible adjustments that can be made due to human resource policies and practices (see Table 2 of Easton and Goodale 2005 for a good overview).

Insert Figure 1 Here

There is a substantial and growing body of academic research on call center workforce management (Gans et al. 2004, Aksin et al. 2007). The vast majority of this research focuses on call

forecasting, queueing, and short term scheduling, with a small amount of work on long term planning. Very little attention has been devoted to the issues and decisions associated with real time schedule adjustments, either for call centers or for other types of service systems.

However, there are several factors that make real time schedule adjustments vital for successful call center management. Most importantly, several researchers (Jongbloed and Koole 2001, Brown et al. 2002, Avradimis et al. 2004, Steckley et al. 2004, Weinberg et al 2006, and Shen and Huang 2006), have recently identified significant correlation between arrivals in different time intervals within the same day, and have suggested methods for updating call forecasts on an intra-day basis; a primary purpose for such updated call forecasts is to provide support for real time schedule adjustments. Secondly, given the lead time associated with schedule generation, many changes to employee availability can and do take place after the original schedules have been created. Thirdly, detecting how well the scheduled agent workforce actually matches the actual workload is often not possible for a given day until that day has begun, at which point responding to the incremental (positive or negative) demand may be crucial. Finally, managers regularly struggle with staffing tradeoffs, for while having too few agents on duty can lead to severe degradations in service quality, having too many agents results in low resource utilization and overspending of scarce financial resources.

The task of making intelligent real time schedule adjustments is a challenging undertaking. Managers have many possible ways of increasing staff levels (offering overtime to agents already scheduled to work, calling in additional agents from home or “borrowing” resources from another department, utilizing outsourcers, eliminating off-phone activities such as meetings or trainings) or decreasing total agent on-phone hours (offering Voluntary Time Off, moving agents to other activities or queues). All of this – when combined with multiple time periods, multiple agent types and costs, and broad classes of feasible shift change options – results in substantial combinatorial complexity.

Despite this complexity, recent empirical research (Mehrotra et al. 2006) estimates that over 70% of call centers routinely make real time schedule adjustments, with these decisions based largely

on experience and intuition. Typically, actual agent attendance and call volumes are observed and compared to forecasted values during the early part of the day, and then agent schedules are updated in an ad hoc manner based on these observations.

The objectives and main contributions of this paper are as follows: (1) to develop a new mathematical framework for real time schedule adjustments that reflects the operational characteristics of the call center environment; (2) to connect the growing literature on random arrival rates and inter-period correlations for inbound call centers to the problem of agent re-scheduling; (3) to identify and illustrate a tractable solution methodology for updating forecasts and determining cost effective schedule adjustments based on these updates; and (4) to illustrate the impact of making these real time schedule adjustments on operating costs and service quality.

In Section 2, we describe the real time management challenge in more detail and review the relevant research literature. In Section 3, we present a framework for managing Real Time Schedule Adjustments within call centers, including a workload forecast updating model and an integer programming formulation that reflects the costs and constraints associated with adjusting agent schedules on an intra-day basis. In Section 4, we demonstrate the value of using this framework to update schedules through an illustrative set of numerical examples. Finally, in Section 5 we conclude by discussing the importance of real time monitoring and management in call center operations as well as directions for future research.

2. REVIEW OF CALL CENTER WORKFORCE MANAGEMENT PROCESSES AND ASSOCIATED LITERATURE

Throughout this paper, we focus on call centers in which all calls are inbound calls from customers and all agents are capable of handling all calls. In such call centers, as in many labor-intensive industries (Hur et al. 2004), the essential scheduling challenge faced is cost-effectively matching the actual demand for service with the service delivery resources (which we refer to as “agents”) available to handle this workload.

The actual workload faced by an inbound call center is typically modeled as a stochastic process based on a random number of arriving calls and a randomly distributed service time for each call. Therefore, the forecasting of workload is an important part of the agent scheduling (and re-scheduling) process. There is extensive literature in this area that is well reviewed by Gans et al. (2004) and by Aksin et al. (2007). In particular, Thompson (1998) provides a general tutorial on demand forecasting for service systems while Andrews and Cunningham (1995) present a vivid case study of one call center's forecasting challenges.

Within the call center literature, the standard forecasting approach is to treat call arrivals over the course of a day or week as a Non Homogeneous Poisson Process (NHPP) with piecewise constant arrival rates over specific time intervals of 15-, 30- or 60- minutes that are independent of each other. Indeed, there is significant theoretical and empirical evidence that supports the concept of modeling call arrivals as an NHPP. When one considers call arrivals as the superposition of arrivals from a large number of independent customers, the Palm-Khintchine Theorem shows that a Poisson Process provides a good approximation (Whitt 2002b). As part of an extensive empirical analysis of one call center's data, Brown et al. (2002) report that they could find no evidence to reject the hypothesis that call arrivals follow an NHPP.

The typical next step is to translate the forecasted arrival rates into a demand for agents, which depends not only on the workload forecast but also on a defined acceptable waiting time distribution. Grassman (1988) discusses many of the practical issues generally associated with this type of translation, while Green et al. (2001, 2003) describe the standard call center forecast translation process, which they refer to as the *stationary, independent, period by period* ("SIPP") method. The SIPP method treats each individual period as an independent stationary queueing system, and for each period sets the target number of servers to be the minimum number for which the acceptable waiting time distribution will be achieved in steady state, given the workload forecast. Green et al. (2001, 2003) also propose improvements to the way in which agent requirements are determined for call

centers with cyclic demand (2001) and call centers with limited daily operating hours (2003).

The translation of call arrival forecasts and target waiting time distributions into agent requirements depends implicitly on the assumption that the forecasted call arrival *rates* are known and deterministic. However, recent research has also shown that the assumption of a deterministic arrival rate within a given period is often invalid, which has implications for determining the number of agents to schedule in each period. Brown et al. (2002), studying data from a bank's call center, test and reject the hypothesis of deterministic (Poisson) arrival rates per time period. Similarly, Steckley et al. (2004) analyze data from several call centers' queues, statistically testing and rejecting the hypothesis that arrival rates are deterministic for the vast majority of queues and time periods studied. Whitt (1999) examines infinite-server systems as a mechanism for understanding the staffing levels required for systems in which the objective is to answer calls immediately, while Whitt (2002) also considers both random arrival rates and employee absenteeism (along with costs associated with servers, waiting time, and abandonment) in developing approximation techniques for determining the optimal number of servers for a given workload distribution. Steckley et al. (2008) provide approximation techniques for performance measures associated with waiting time distribution in the presence of a random arrival rate in order to facilitate the selection of the number of servers. Motivated by empirical observations of random arrival rates, Jongbloed and Koole (2001) consider the question of how to "schedule agents ... in a statistically correct way" when the arrival *rate* itself is a random variable. In particular, they propose a Poisson mixture model for arrivals within a specific time period, and then explore various methods for determining the number of agents to schedule. Robbins (2007) suggests a stochastic programming approach to scheduling agents while explicitly accounting for arrival rate uncertainty.

In addition to random arrival rates, there is also considerable evidence that the arrivals across different periods within the same day are correlated with one another; these correlations have in turn prompted researchers to create more sophisticated forecasting and staffing models. Motivated by empirical results that show strong correlations across periods within the same day, Avramidis et al.

(2004) develop and test several models in which the arrival *rate* for each interval of the day is a random variable that is correlated with the arrival rates of the other intervals. Brown et al. (2002) develops a non-linear least squares model in which a previous day's call volume is an independent variable in predicting the subsequent day's call volume, producing roughly a 50% reduction in the variability of the forecasted daily volumes.

Most recently, several researchers have confirmed the persistent presence of intra-day cross-period correlations in large call center datasets and developed sophisticated techniques for intra-day forecast updating. Weinberg et al. (2006) use data from a large North American bank to develop a two-way multiplicative Bayesian Gaussian model for forecasting call arrivals, with Monte Carlo Markov Chain (MCMC) methods used for parameter estimation; their empirical analysis shows strong intra-day correlations and substantial improvements in forecast accuracy based on MCMC parameter updating. Shen and Huang (2006) analyze data from a major financial services firm's inbound call center and demonstrate strong intra-day correlations; these results are used to motivate intraday forecast updating methods based on Singular Value Decomposition techniques and a Penalized Least Squares model.

Beyond call centers, this intra-period correlation has been investigated in many other settings as well. Bodily and Freeland (1988) examine several different forecast updating techniques for predicting overall product shipments based on initial observed orders, while Kekre et al. (1990) and Guerrero and Elizondo (1997) also examine the problem of updating a cumulative demand forecast based on a subset of actual demand. Hur (2002) proposes a variety of monitoring techniques to identify when new information might suggest the need to update previous forecasts.

These recent research results support the premise that new information about call arrivals during the first few time intervals of a given day may provide important insights into the distribution of calls over the remainder of the day, and in turn provide motivation for adjusting agent schedules in order to better meet an updated demand forecast.

Real time schedule adjustment processes start with initial forecasts and agent schedules, and

then seek to update them based on new information that has become available more recently. The three components of the re-scheduling process are analogous to the three components of the scheduling process, and thus the literature on real time schedule adjustments includes work on: (1) updating call forecasts; (2) revising resource requirements; and (3) updating agent schedules.

Interestingly, while there is great deal of literature on shift scheduling in general (Ernst et al. 2004) and in call centers in particular (Gans et al. 2004, Aksin et al. 2007), real time schedule adjustments have been studied far less extensively. Thompson (1996, 1999) has done some initial work on real-time schedule adjustments in service systems. Easton and Goodale (2005) propose a methodology for re-scheduling resources in service systems to account for absenteeism, focusing on systems in which there is quantifiable marginal revenue associated with handling customers that would otherwise abandon the system. Hur et al. (2004) provide an excellent overview of the literature associated with Real Time Schedule Adjustments while highlighting some of the challenges and also developing re-scheduling techniques specifically for the context of quick service restaurants. Citing several sources, including Cleveland and Mayben (1997) and Mabert (1991, 1995), Hur et al. (2004) also assert that “even the most accurate call center staff scheduling must be complemented by real-time schedule adjustment to achieve the target customer service level.” Despite this widespread belief, there is a surprising absence of research on real time agent schedule adjustments within the call center research literature.

3. INTRA-DAY SCHEDULE UPDATING METHODOLOGY

3.1. Overview

In this section, we describe our Intra-Day Schedule Updating methodology. The underlying business context is a (possibly virtual) call center that handles a single type of incoming phone call, with all agents being skilled to handle each call. At the beginning of the day, there is an initial call volume forecast for each period of the day, and an existing set of agents who are scheduled to handle this

workload. Agents are grouped into “types” by the specific details of their schedules, where an agent type is defined by (a) the period in which agents of this type begin their shift; (b) the specific periods during which they are available to handle calls; and (c) the period in which they complete their shift.

Once the day begins, managers observe the actual workload (and the actual attendance of the agents) at the end of each period, paying attention to the deviation from the forecasted workload (and the expected agent coverage levels). When the cumulative call volume deviates significantly from what was forecasted, the forecast for the remainder of the day is updated, an incremental demand for agents for each period for the remainder of the day is identified, and some or all agents’ schedules can be updated to reflect these changes in demand. If the value of these schedule changes exceeds the associated disruption costs, the schedule changes are communicated to agents and the updated schedules are followed for the remainder of the day.

Once the incremental demand for agents has been identified for subsequent periods of the day, the intra-day rescheduling model seeks to identify a cost effective solution that meets the updated agent requirements for each remaining period. To accommodate this rescheduling, agents of a particular type may be asked to transition to a different schedule, with this new schedule featuring at least one period in which this agent was previously working (or idle) but is now idle (working). It is assumed that the range of possible new schedules for each agent type, the costs associated with each of these transitions, and the disruption costs associated with changing agent schedules are all known prior to solution of the re-scheduling problem.

3.2. Initial Schedule Parameters and Notation

To represent the call center’s daily operations and initial agent schedules, for a given day we define the following notation:

Initial Parameters

$T =$ number of periods within a day

$N =$ number of agent types

$m_i =$ number of type i agents originally scheduled to work on this day

$b_{it} =$ 1, if type i agents are scheduled to take inbound calls in period t ,
0, otherwise, for $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$

$d_t =$ number of agents scheduled to take inbound calls in period t , for $t = 1, 2, \dots, T$

For any given vector $\mathbf{m} = \{m_i\}$ and matrix $\mathbf{B} = \{b_{it}\}$, we can compute $\mathbf{d} = \{d_t\}$ using $\mathbf{d} = \mathbf{Bm}$.

The case where $b_{it} = 0$ for all $t = 1, 2, \dots, T$ corresponds to all agents of type i being currently unscheduled for the particular day in question. In practice an agent of this type may be someone who is available to be called in from home with some lead-time, an employee in another department who is capable of handling these calls, or an agent that is available from a third party or “outsourced” call center (see Easton and Goodale (2005) for a good discussion of these types of contingent resource options). For this type of agent, the parameter m_i corresponds to the maximum number of agents of this type who are available for duty as part of an updated agent schedule.

3.3 Monitoring Call Volumes and Identifying Incremental Demand for Agents

Prior to each time period u , our methodology monitors actual call arrivals and compares them to forecasted call volumes. When the cumulative call volume deviates significantly from the expected cumulative call volume, we identify u as a possible schedule updating period and estimate an incremental demand for agents. This incremental agent demand then serves as an input into the agent rescheduling model described in Section 3.4. Our description below uses the following notation and associated definitions:

Agent Demand Updating: Definitions and Notation

$N_t =$ random variable representing the number of calls arriving in period t

$\lambda_t =$ expected value of N_t

$C_u =$ random variable representing the cumulative number of calls arriving in periods $1, 2, \dots, u-1$

$x_t =$ actual number of calls observed in period t , for $t = 1, 2, \dots, u-1$

$S_u =$ actual cumulative number of calls arriving in periods $1, 2, \dots, u-1$

$\hat{d}_t =$ number of agents required to take inbound calls in period t after the updated forecast, for $t = u, u+1, \dots, T$

$\delta_t =$ the amount of staff change justified by the updated forecast and observed agent attendance levels in period t , for $t = u, u+1, \dots, T$

3.3.1. Modeling Call Arrivals

We model N_t , as a stochastic process with random arrival rates in which the number of calls arriving in each period may be correlated with the number of calls arriving in previous periods. This general model allows for a wide range of call forecasting techniques, including the distributional forecasting models presented in Whitt (1999), Brown et al. (2002), Avramidis et al. (2004), Weinberg et al. (2006), and Shen and Huang (2006).

3.3.2. Comparing Actual Call Arrivals to Forecasted Call Arrivals

Our first criterion for determining whether or not to adjust agent schedules at the beginning of a given period u is to compare actual calls to the distributional call forecasts. Immediately prior to period u , we monitor the actual cumulative call arrivals S_u . Given the distributional forecasts for the call volumes N_1, N_2, \dots, N_{u-1} , we also have a distributional forecast for the cumulative call volume C_u for periods $1, 2, \dots, u-1$. Based on the distributional forecast for C_u , we can compute $\pi_u \equiv P(C_u > S_u)$. We interpret a small value (respectively, a large value) of π_u to suggest that the actual cumulative call volume is significantly greater than expected (significantly lower than expected). Our methodology utilizes threshold values p_1 and p_2 , where $0 \leq p_1 \leq p_2 \leq 1$, to determine whether the call center is potentially

overstaffed (if $\pi_u \leq p_1$) or potentially understaffed (if $\pi_u \geq p_2$).

On the other hand, we infer that S_u is relatively close to $E[C_u]$ if $p_1 \leq \pi_u \leq p_2$, and consequently we do not consider updating agent schedules prior to period u . In such cases, the remaining steps described in the sections below are not executed for period u .

3.3.3. Updating Forecasts Based on Actual Call Volumes S_u

If the call center appears to be overstaffed or understaffed based on the criteria described in the previous section, the next step in our methodology is to update the forecasts. This is defined quite generally, and can be either an update to the forecasted mean arrival rates λ_t for $t=u, u+1, \dots, T$ (which we denote $\hat{\lambda}_t^u$) or an update to the distributional forecast for N_t for periods $t=u, u+1, \dots, T$ (which we denote \hat{N}_t^u). In either case, the purpose of the updated forecast is to determine an updated target for agent requirements, which we discuss in the next section.

3.3.4. Updating Target Agent Requirements and Incremental Per Period Agent Demand

Given an updated forecast for the call arrival rates $\hat{\lambda}_t^u$, one can use the standard SIPP translation or one of several variants in the literature (as discussed in Green et al. 2001) to determine the value of \hat{d}_t^u , the minimum number of agents needed to achieve the desired waiting time objective in periods $t = u, u+1, \dots, T$. Alternately, the additional information contained in the updated distributional forecasts \hat{N}_t^u for periods $t = u, u+1, \dots, T$ can be used to determine \hat{d}_t^u using approximation techniques such as those presented by Steckley et al. (2008). Throughout the remainder of this section, we suppress the superscript on \hat{d}_t^u for clarity.

Next, for all periods $t \geq u$, we compute $\delta_t = (\hat{d}_t - d_t)$ where the first term on the right hand side corresponds to the target number of agents for period t based on the updated forecast and the second term corresponds to the number of agents originally scheduled for period t .

For clarity of exposition, we have omitted any adjustment for absenteeism in our description of the methodology here. However, absenteeism can be incorporated simply by adjusting the number of scheduled agents d_t by the number of scheduled agents who as of the end of period $u - 1$ are projected to be absent during each period prior to computing the δ_t values for all periods $t \geq u$. Similarly, although our presentation above focuses on changes in call volume forecasts, we note that changes to service time parameter forecasts and/or the waiting time distribution (“service level”) objectives for the remaining periods $u, u + 1, \dots, T$ can also be considered in a straightforward manner.

3.4. Formulation of the Re-Scheduling Model

Once the change in agent demand δ_t has been determined, we then solve an integer programming model to determine a cost effective agent re-scheduling plan. This model is formulated in detail in this section, first in general terms and then for two important special cases.

Schedule Updating Parameters

$K_i =$ number of schedule transitions available to a type i agent

$$r_{ikt} = \begin{cases} 0 & \text{with transition } k, \text{ a type } i \text{ agent remains status quo in } t \\ 1 & \text{with transition } k, \text{ a type } i \text{ agent with } b_{it} = 0 \text{ takes inbound calls in } t \\ -1 & \text{with transition } k, \text{ a type } i \text{ agent with } b_{it} = 1 \text{ takes time off in } t \end{cases}$$

Cost and Saving Parameters

$p_{ik} =$ the penalty of making the transition k for a type i agent

$c_{ikt}^+ =$ the cost of having a type i agent make a transition in period t where $r_{ikt} = 1$

$c_{ikt}^- =$ the savings from having a type i agent make a transition in period t where $r_{ikt} = -1$

Decision Variables

y_{ik} = the number of type i agents who transition to the new schedule k

The general schedule updating model determines the number of agents of each type that should make each type of feasible transitions in order to accommodate the forecast update while minimizing the total cost of the transitions. Mathematically, the objective is to:

$$\text{Minimize } \sum_{i=1}^N \sum_{k=1}^{K_i} p_{ik} y_{ik} + \sum_{i=1}^N \sum_{k=1}^{K_i} \sum_{t:r_{ikt}=1} c_{ikt}^+ y_{ik} - \sum_{i=1}^N \sum_{k=1}^{K_i} \sum_{t:r_{ikt}=-1} c_{ikt}^- y_{ik}$$

Subject to

$$\sum_{i=1}^N \sum_{k=1}^{K_i} r_{ikt} y_{ik} \geq \delta_t \quad \forall t = u, u+1, \dots, T \quad (4)$$

$$\sum_{k=1}^{K_i} y_{ik} \leq m_i \quad \forall i \quad (5)$$

$$y_{ik} \in \mathbb{Z}^+ \cup \{0\} \quad \forall i, k \quad (6)$$

3.4.1. Special Case I: The Understaffed Call Center

In the intra-day context, we describe a call center as being “understaffed” if $\mathbf{d} \leq \hat{\mathbf{d}}$ and $d_t < \hat{d}_t$ for at least one period after the forecast update in period u . When a call center is understaffed, management is confronted with a difficult choice: maintaining current staffing levels is likely to result in long waiting times and high abandonment rates, while increasing staffing levels means additional labor expenditures. In such cases, our modeling framework seeks to enable managers to improve service quality while at the same time cost-effectively managing incremental labor costs.

One way to increase the staffing levels is by extending one or more agents’ shifts by one or more periods; this extension of an agent’s shift is known as “overtime” (OT) and typically has a pay premium associated with it. In this case, we assume that an agent who is done with his regular shift immediately starts working overtime if he is offered OT. Another way to increase staffing levels is to add agents who are not already scheduled to handle these calls, either by calling in agents from home,

from another department or group, or from an outsourcer. The challenge is to determine how many of each type of resource to add to the schedule over which periods of the day.

We now present an integer programming model to determine a cost effective mix of OT offers and additional agents (how many agents and for what periods) to remedy the understaffing that is projected for periods $t = u, u + 1, \dots, T$. For clarity of exposition (and without loss of generality), we will assume that $u = 1$ for the remainder of the model section.

For this special case, we will need the following additional notation:

Parameters for the Understaffed Model

$o_i =$ overtime cost per period of a type i agent

$p_{it} =$ cost of making a schedule change to a type i agent in period t

$e_i =$ daily cost of bringing in a call-in agent of type i

$L_i =$ The last period in which a type i agent is on staff w.r.t. his original schedule

Decision Variables for the Understaffed Model

$Y_{it} =$ the number of type i agents that work overtime in period t .

$Q_{jt} =$ the number of type j call-in agents that start working in period t .

In our formulation, we assume that there are two types of call-in agents: Part-Time agents available for 4 hours only and Full-Time agents available for 8 hours.

The following integer programming model is defined using the decision variables Y_{it} . However to relate these variables to the general formulation, we have

$$Y_{it} = \sum_{k=1}^t y_{ik} \quad (7)$$

where transition k is defined for a type i agent as the agent continuing work after his regular shift ends

and working overtime for k periods before leaving for the workday. Therefore, if a type i agent is on staff until period L_i on a given day, he is allowed to have one of $T - L_i$ transitions to a new schedule, where transition k is defined as the agent working overtime from period $L_i + 1$ to $L_i + k$.

Given the above parameters and decision variables, the OT integer program is defined as follows:

$$\text{Minimize } \sum_{i=1}^N \sum_{t=2}^T p_{it} (Y_{it} - Y_{i,t-1}) + \sum_{i=1}^N \sum_{t=1}^T o_i Y_{it} + \sum_{j=1}^2 \sum_{t=1}^T e_j Q_{jt}$$

Subject to:

$$Y_{it} \leq Y_{i,t-1} \quad \forall i, t > L_i + 1 \quad (8)$$

$$Y_{it} = 0 \quad \forall i, t < L_i + 1 \quad (9)$$

$$Y_{it} \leq m_i \quad \forall i, t = L_i + 1 \quad (10)$$

$$\sum_{i=1}^N Y_{it} + \sum_{s=\max\{0,t-15\}}^t Q_{1s} + \sum_{s=\max\{0,t-31\}}^t Q_{2s} \geq \delta_t \quad \forall t \quad (11)$$

$$Y_{it} \in \mathbb{Z}^+ \cup \{0\} \quad \forall i, t \quad (12)$$

In the above formulation, the objective function has three components. The first component calculates the total schedule change penalty for the number of type i agents who just had a schedule change at period t using $(Y_{it} - Y_{i,t-1})$. The second component computes the additional cost caused by agents who are asked to stay for overtime. The last term calculates the cost of the call-in agents.

The constraints sets of the formulation can be described as follows: (8) after their regular shift is over, the number of type i agents working overtime must be a monotonically non-increasing function in t ; (9) by definition, no agent can work overtime during his/her originally scheduled periods; (10) for any agent type, one may request overtime work from as many employees of that type as the number that are on duty that day; and (11) the total number of agents that are offered OT and that are called in

should be sufficient to account for the demand shift.

3.4.2. Special Case II: The Overstaffed Call Center

In the intra-day context, we describe a call center as being “overstaffed” if $\mathbf{d} \geq \hat{\mathbf{d}}$ and $d_t > \hat{d}_t$ for at least one period after the forecast update in period u . When a call center is overstaffed, management’s goal is to modify the original schedules in order to reduce staffing costs and increase agent utilization levels while simultaneously maintaining high service levels and low abandonment rates.

In an overstaffed situation, management often has the opportunity to save money by giving agents a chance to leave work early; this is known as Voluntary Time Off (VTO). In this case, we assume that once an agent has taken VTO, he will not return to work the rest of the workday. The challenge is to offer VTO to the right mix of agents, given the updated demand levels \hat{d}_t , the mix of agents already scheduled, and the dynamics of each agent’s work schedule as represented by the elements of the shift matrix \mathbf{B} . Hence, the goal is to determine the number of agents of each type i to receive VTO at the beginning of each period t , $t - u, u + 1, \dots, T$. Below, we present an integer programming model to determine a cost effective of VTO offers.

For this special case, we will need the following additional notation:

Parameters

$s_i =$ cost savings per period by giving VTO to a type i agent

$p_{it} =$ cost of making a schedule change to a type i agent in period t

$F_i =$ the first period in which a type i agent is on staff in the original schedule

Decision Variables

$W_{it} =$ the cumulative number of type i agents that are given voluntary time off up to and including period t .

The following integer programming model is defined using the decision variables W_{it} . However to relate these variables to the decision variables in our general formulation, we note that

$$W_{it} = \sum_{k=1}^t y_{ik}, \quad (13)$$

where transition k is defined as the agent taking VTO starting period $k - 1$ and not returning to work for the rest of the workday. Hence, if a type i agent is on staff for j periods on a given day, he is allowed to have one of j transitions to a new schedule. The VTO integer program can now be defined as follows:

$$\text{Minimize } \sum_{i=1}^N \sum_{t=1}^T p_{it} (W_{it} - W_{i,t-1}) - \sum_{i=1}^N \sum_{t=1}^{L_i} s_i W_{it}$$

Subject to

$$W_{it} \geq W_{i,t-1} \quad \forall i, t \geq F_i \quad (14)$$

$$W_{it} = 0 \quad \forall i, t < F_i \quad (15)$$

$$W_{it} \leq m_i \quad \forall i, t = L_i \quad (16)$$

$$W_{it} = W_{i,t-1} \quad \forall i, t > L_i \quad (17)$$

$$\sum_{i=1}^N B_{it} W_{it} \leq \max(-\delta_t, 0) \quad \forall t \quad (18)$$

$$W_{it} \in \mathbb{Z}^+ \cup \{0\} \quad \forall i, t \quad (19)$$

In the above formulation, the objective function has two components. The first component calculates the total schedule change penalty for the number of type i agents who just had a schedule change at period t using $(W_{it} - W_{i,t-1})$, while the second component computes savings caused by agents who are asked to take VTO.

The constraint sets of the formulation can be described as follows: (14) the total number of type i agents offered VTO is non-decreasing as a function of time, starting in the first period in which a

type i agent is on staff; (15) no agent can be offered VTO for any period before he starts his shift; (16) for any agent type, one may offer VTO to as many employees as the number that are on duty that day of that type; (17) no agent can be offered VTO for any period that is after the completion of his or her shift; and (18) for each time period, the number of agents offered VTO cannot be greater than the change in the demand for agents.

3.4.3. Characterization of the Rescheduling Models

The OT model described above contains $T \times N + T \times 2$ general integer decision variables and $T \times N + T$ constraints, while the VTO model has $T \times N$ general integer decision variables and $T \times N + T$ constraints. For the numerical examples presented in the following section, the optimization models were solved with Premium Solver (Fylstra et al. 1998) on a Pentium-class laptop computer in 2-3 seconds each .

3.5 Determining the Value of Schedule Changes and Making the Rescheduling Decision

After solving the appropriate agent rescheduling integer model for a given rescheduling period u , the next step in the process is to estimate the net value associated with the proposed schedule adjustment. Once this value has been estimated, a final decision is made about whether or not to implement the updated schedules prior to period u .

3.5.1 Quantifying the Value of a Schedule Update

The net value of an updated agent schedule is comprised of three components: (1) the impact of the updated schedule on direct labor costs; (2) the impact on customer service quality, as measured by the Service Level changes resulting from the updated schedules; and (3) the “disruption costs” that result from changing individual agents’ schedules. We discuss modeling and calculation of each of these components below.

The increase (or decrease) in direct labor costs associated with a schedule update is defined as

the additional cost (savings) incurred as a result of agents working overtime (taking voluntary time-off) and can be obtained directly from the objective function value of the integer programming model. For a schedule update taking place prior to period u , we denote this value as Δ_L^u and note that a positive (negative) value for Δ_L^u corresponds to an increase (decrease) in direct labor costs relative to the labor costs associated with the base schedule.

To quantify the value of updated agent schedules on customer service quality in future periods, we first model a cost c_S associated with each call that waits in queue longer than the target service level period. For a proposed updated schedule (and associated schedule updating period u), we then estimate the number of calls that are expected to wait longer than the desired service level interval, which we denote as SL_0 , based on the updated forecast and the original agent schedule. We then update the agent schedules for periods $u, u+1, \dots, T$ as described above and use the results along with the updated forecasts to estimate the number of calls that are expected to wait longer than the desired service level interval, which we denote SL_u , based on these updates. The value of the impact of a schedule update on service levels prior to period u , which we denote Δ_S^u , is then estimated by $(SL_0 - SL_u) c_S$, where a positive (negative) value of Δ_S^u corresponds to a decrease (increase) in the costs associated with not meeting Service Levels.

Finally, to model disruption costs associated with updated agent schedules, we include a cost c_d for each individual agent whose schedule is changed in the updating process. In addition, we define γ_u to be the number of agents whose schedules would be updated if results of the rescheduling model were implemented prior to period u , where:

$$\gamma_u \equiv \sum_{i=1}^n \max_{t=u, u+1, \dots, T} Y_{it} \quad (\text{OT Case})$$

$$\gamma_u \equiv \sum_{i=1}^n \max_{t=u, u+1, \dots, T} W_{it} \quad (\text{VTO Case}),$$

where Y_{it} (W_{it}) is as defined in Section 3.4.1 (3.4.2) above. The total disruption cost associated with

updating agent schedules prior to period u is denoted Δ_D^u and is calculated as $\gamma_u c_d$.

3.5.2 Determining Whether or Not to Implement the Proposed Schedule Update

The total costs associated with a schedule update prior to period u is denoted Δ^u , with

$\Delta^u \equiv \Delta_L^u + \Delta_S^u + \Delta_D^u$. Our methodology implements the updated schedule prior to period u only if

$\Delta^u < 0$, which means that the net value (cost) of the updated schedule is positive (negative).

If $\Delta^u \geq 0$, schedules are not updated prior to period u , the actual call volume x_u for period u is observed, and the process described in Sections 3.3 and 3.4 above is repeated.

4. EMPIRICAL TESTING OF INTRA-DAY RESCHEDULING METHODOLOGY

4.1 Numerical Experiments: Description and Input Parameters

To test the rescheduling framework described in the previous section, we used data from an actual call center's operations to conduct a series of numerical experiments. For each day of operations, our experimental framework is illustrated in Figure 2. Our approach is to start with day-of-operations forecasts ("original forecasts") and the associated agent schedules ("original schedules"). From here, we observe the actual call arrivals for each day and use this information to create "updated forecasts." Next, we examine the updated forecasts in conjunction with the original schedules to determine if and when to update agent schedules. Finally, using the actual call arrivals for the entire day, we compare the performance of the call center when the original schedules are used for the entire day with the performance of the system when our rescheduling framework is used to update the schedules. This process is described in more detail below.

Insert Figure 2 Here

4.1.1 Overview of Call Center and Determination of Original Schedules

The call center that was used for our numerical experiments was originally presented by Saltzman (2005). This call center provides support for an on-line e-commerce portal and is open from 6:00 am to 9:00 pm, with a schedule granularity of 15-minute increments. The center employs 32 types of agents, 10 with part time and 22 with full time schedules.

In order to calculate the initial target staffing levels d_t , we used day-of-operations forecasts for call arrival rates λ_t and average service rates μ_t . From here, the minimum number of agents required per period is calculated using the traditional Stationary Independent Poisson Process (SIPP) transformation with a service level goal of answering 99% of calls within 300 seconds. Finally, the original agent schedules were created based on these day-of-operations forecasts and service level targets by using the schedule optimization algorithm from Saltzman (2005).

It is important to note that our day-of-operations forecasts λ_t can differ significantly from the two-week moving average forecasts actually used by this call center to determine agent requirements and create agent schedules. Specifically, we use the ratio of actual calls to forecasted calls during the first n days of the week to update the original forecast to get a day-of-operations forecast for the $n+1^{st}$ day. In addition, our original schedules are based on these day-of-operations forecasts, whereas the actual schedules for this call center were based only on the original forecasts.

4.1.2 Determining Whether or Not to Consider Updating Agent Schedules

The general process for determining when and how to update agent schedules is described in detail in Sections 3.3 – 3.5 above. Below, we provide additional details about the specific criteria used in our numerical experiments.

For each day in our experiments, the first allowable period for re-scheduling is $u = 9$, meaning that all of the estimates used for forecast updates are based on at least eight observations (that is, at least two hours of data). This is consistent with both current industry practice and with the empirical

results of Shen and Huang (2006) about the performance of this forecast updating method, which suggests that accuracy increases greatly after the first few periods. Our last allowable period for re-scheduling is $u=17$, meaning that the decision on whether or not to update call forecasts and agent schedules is based on at most sixteen observations (that is, four hours of data).

For a given period u , we determine whether or not actual cumulative call arrivals are significantly higher (or lower) than forecasted cumulative call arrivals based on the value of π_u as described in Section 3.3.2. In particular, in our experiments, we use the threshold value $p_1 = 0.2$ ($p_2 = 0.8$) to determine whether call volumes are significantly higher (lower) than expected, which in turn suggests that the system is potentially understaffed (overstaffed). Values of p_1 and p_2 closer to (or farther from) 0.5 correspond to a lower threshold (higher threshold) for concluding that the actual call arrivals have deviated significantly from the expected values.

4.1.3 Updating Forecasts and Agent Requirements

If we have determined that the call volumes are significantly higher or lower than expected, our next step is to make use of the call volumes observed in $1, 2, \dots, u-1$ to update the demand for agents in period t for $t = u, u+1, \dots, 60$. This is done in two steps.

The first step is to create an updated forecast for the arrival rate in period t as of the beginning of period u , which we denote $\hat{\lambda}_t^u$, where $\hat{\lambda}_t^u \equiv \lambda_t \frac{S_u}{E[C_u]}$. The second step is to use the updated forecast in the standard SIPP procedure (as discussed in Green et al. 2001, 2003) to determine the updated agent requirements $\hat{d}_u, \hat{d}_{u+1}, \dots, \hat{d}_{60}$. The values of the incremental agent demand levels $\delta_u, \delta_{u+1}, \dots, \delta_{60}$ are then computed as the difference between $d_u, d_{u+1}, \dots, d_{60}$ and $\hat{d}_u, \hat{d}_{u+1}, \dots, \hat{d}_{60}$.

4.1.4 Creating Updated Agent Schedules and Evaluating the Associated Benefits

After determining the updated agent requirements and incremental agent demand levels, for understaffed cases (overstaffed cases) the integer program to update schedules from Section 3.4.1

(Section 3.4.2) is run to determine the updated schedules. Once the updated schedules have been identified, we determine the net value of updating the schedules as of time u by following the procedure described in Section 3.5, using the following cost parameters.

For our numerical experiments, we used a standard labor cost of \$20/hour for all agents of all types as in Saltzman (2005) as well as values of \$15/hour for the savings rate s_i associated with an hour of Voluntary Time Off for all agents and a value of \$27/hour for the overtime cost o_i for all agents who are already scheduled. In addition, we used a value of \$36/hour for the hourly cost of for any previously-unscheduled agents who are called in to meet the increased demand levels. In practice, call-in agents are asked to show up whenever needed and on very short notice, which is the justification for the higher overtime costs. We assessed a cost of \$25/call for each call not answered within the Service Level target period. Finally, we included a charge of \$5 for each individual agent whose schedule was changed, and this disruption charge was explicitly included in our calculation of the net value of any given schedule update, as described in Section 3.5.

4.2 Experimental Results

4.2.1 Initial Results

Our numerical experiments are based on six and a half weeks of day-of-operations forecasts and call arrival data, for a total of thirty two distinct days of call center operations. Our presentation of numerical results is focused on the financial impact of the schedule updates. At the end of each day, we can estimate the financial impact of the schedule update decision by calculating the performance metrics using *actual call volumes*, first with the original agent schedules and then with the updated agent schedules.

For three of these thirty two days, the model either did not identify any significant differences between the original call arrival forecast and the actual call arrivals or estimated that the value of updating agent schedules was less than the cost of disruption. Of the remaining twenty nine days,

sixteen were identified as “understaffed” and thirteen were identified as “overstaffed.” For our initial tests, we used the updated value for the mean arrival rate $\hat{\lambda}_i^u$ to determine the updated agent requirements \hat{d}_i^u , as described in Section 3.3.4 above.

When the early periods’ call arrivals suggest that the call center is understaffed, our model seeks to increase staffing levels to meet service level goals and reduce lost calls for the remainder of the day. For days that were identified as understaffed, Table 1 illustrates the impact of the schedule updating model on these metrics. In particular, we note that schedule updating has a strong impact on the achieved Service Levels (with the mean increasing from 83.50% under the original schedules to 89.07% in the presence of updated schedules). As such, the cost of calls that fail to meet the Service Level objective decreases substantially as a result of our intra-day rescheduling, in all cases exceeding the cost of additional staff and the cost of disruption.

Insert Table 1 Here

When the early periods’ call arrivals suggest that the call center is overstaffed, our model seeks to reduce staffing levels, reducing labor costs and thereby increasing agent utilization levels, while also continuing to achieve the service level objectives. For days that were identified as overstaffed, Table 1 illustrates the impact of the schedule updating model. In the overstaffed case, we see that the cost of lost calls grows as a result of the decision to update agent schedules and that on average this cost exceeds the amount that is saved as a result of decreased labor costs. In addition, we notice that even the original schedules fail to meet the service level objectives over the course of the day; this reflects the fact that, for this call center, some days’ actual call arrivals are greater than expected in later periods even though the early periods are lower than expected. We address this phenomenon with additional experiments, which are described in the next section.

Insert Table 2 Here

4.2.2 Incorporating Distributional Information In Determining Updated Agent Requirements

In our next round of numerical experiments, we explored the impact of using additional distributional information to determine the updated agent requirements \hat{d}_t^u , as described in Section 3.3.4 above. In particular, in understaffed (overstaffed) situations, we define values $k_1 > 0$ ($k_2 > 0$) and then use $\hat{\lambda}_t^u + k_1\sigma_t$ ($\hat{\lambda}_t^u + k_2\sigma_t$) as the updated mean arrival rates that are used to determine \hat{d}_t^u , where σ_t is the standard deviation of the call volume distribution for period t for $t = u, u+1, \dots, 60$.

The parameters k_1 and k_2 can be interpreted as “insurance” against the remaining arrival rate uncertainty, with larger k_i values corresponding to more conservative estimates of the updated arrival rate. In the understaffed case, a positive value of k_1 corresponds to a more aggressive attempt to meet service levels in periods $u, u+1, \dots, 60$ by adding more staff hours during the re-scheduling process than in our initial experiments, which correspond to the case where $k_1 = 0$. Conversely, in the overstaffed case, a positive value of k_2 corresponds to a more conservative approach to releasing resources that might be needed in the event of higher-than-expected call volumes in future periods than in our initial experiments, which correspond to the case where $k_2 = 0$.

During the sixteen days that were identified as understaffed, we ran experiments with values of k_1 ranging from 0.25 to 2.5 with increments of 0.25. The results are shown in Table 3. While increased values of k_1 always correspond to higher target values of \hat{d}_t^u and therefore lower costs associated with calls that fail to meet service level, it is interesting to note that the net benefit peaks when $k_1 = 1.25$. For $k_1 > 1.25$, the mean incremental benefits associated with improving service levels fell short of the cost of increased staff hours and disruption.

Insert Table 3 Here

On the other hand, the thirteen days identified as overstaffed show somewhat different

dynamics. In particular, the higher the value of k_2 , the smaller the change to staffing levels and the lower the loss associated with calls in future periods that do not achieve the service level goal.

Insert Table 4 Here

5. CONCLUSIONS AND NEXT STEPS FOR RESEARCH

In this paper, we have described an important business problem associated with Real Time Schedule Adjustments for call center operations. The methodology that we have developed here enables managers to update daily workload forecasts and demand for agent resources by leveraging information obtained from observations of call traffic early in the day, and to then use these updated demand levels to intelligently re-schedule agents across a range of feasible adjustments. Our experimental results, based on data from an actual sales-and-service call center, show that there is significant business value associated with such intra-day adjustments when the call center appears to be understaffed. When this particular call center appears to be overstaffed, on the other hand, there appears to be significant risk associated with releasing agents from their schedules, though our rescheduling framework provides an approach to mitigating this risk.

In developing our methodology, we have connected major ideas from the forecasting and forecast updating literature and from the optimal shift scheduling literature to produce an integrated solution for schedule updating. In addition, we have demonstrated the use of this methodology to update day-of-operations forecasts and agent schedules for an actual call center.

We believe that the area of intra-day/real time schedule adjustments has received insufficient research attention, both in the call center literature and in the personnel scheduling literature overall. As such, we conclude by suggesting several other related research questions.

One obvious critical input to this process is the joint distribution of (and correlation between) calls during different intervals within the same day. While a few papers have appeared in the literature

recently (Brown et al. 2002, Avramidis et al. 2004, Weinberg et al. 2006, Shen and Huang 2006), there is still a need for additional work in this area, as improved intra-day forecast updates will lead to better intra-day schedule updates.

A related issue is the process of determining of initial (and updated) staffing levels and the impact of these decisions on agent schedules. Historically, agent requirement calculations have been made under the assumption of a deterministic arrival rate (the so-called SIPP method) and in turn agent schedules have been based on these agent requirement levels. Recent work by Steckley et al. (2008) proposes an approximation method for determining staffing requirements to achieve a certain service level objective while explicitly accounting for arrival rate variability. As such, another avenue for investigating Real Time Schedule Updates for call centers is to examine the effect of using such approximations for determining the agent requirement levels used to determine the initial agent schedules and/or updated agent schedules.

For understaffed situations, our model assumes that there are additional resources available on an on-call basis that can be added to the schedule to help meet the higher-than-expected demand levels. In practice, such agents are often likely to be contracted by a third-party, a relationship commonly referred to as “outsourcing.” Because of the rapid growth in the call center outsourcing industry, contracting and utilizing contingent resources has recently been explored by several researchers, including Milner and Lennon-Olsen (2006), Bhandari et al. (2006), and Gans and Zhou (2007). Exploring contract structures and contingent capacity planning models in the context of schedule updating, variable arrival rates, and cross-period correlation is of both theoretical and practical interest.

Finally, there is clearly a need to extend this type of schedule updating framework to include more sophisticated and increasingly common call arrival and call management issues such as customer messages leading to callbacks (Armony and Maglaras 2004) and skill-based routing (Pot and Koole 2005, Wallace and Whitt 2005, and Harrison and Zeevi 2005).

ACKNOWLEDGEMENTS

We would like to thank the editor and the anonymous referees for the invaluable feedback they provided during the revision of this manuscript.

REFERENCES

1. Abernathy , W. J., N. Baloff, J.C. Hershey, and S. Wandel. 1973. A three-stage manpower planning and scheduling model - a service-sector example, *Operations Research* **22** (3) 693–711.
2. Aksin, Z., M. Armony, and V. Mehrotra. 2007. The Modern Call-Center: A Multi-Disciplinary Perspective on Operations Management Research, *Production and Operations Management* **16** (6) 665-688.
3. Andrews, B.H. and S.M. Cunningham. 1995. L.L. Bean improves call-center forecasting, *Interfaces* **25** (6) 1–13.
4. Armony, M. and C. Maglaras. 2004. On customer contact centers with a call-back option: Customer decisions, routing rules, and system design, *Operations Research* **52** (2) 271–292.
5. Avramidis A.N., A. Deslauriers, and P. L’Ecuyer. 2004. Modeling daily arrivals to a telephone call center, *Management Science* **50** (7) 896–908.
6. Bhandari., M. Harchol-Balter, and A. Scheller-Wolf. 2008. An Exact and Efficient Algorithm for the Constrained Dynamic Operator Staffing Problem for Call Centers, *Management Science* **54** (2) 339-353.
7. Bodily , S.E. and J.R. Freeland. 1988. A simulation of techniques for forecasting shipments using firm orders-to-date, *Journal of the Operational Research Society* **39** (9) 833–846.
8. Brown, L. et al. 2005. Statistical analysis of a telephone call center: A queueing-science perspective, *Journal of the American Statistical Association* **100** (469) 36-50.
9. Cleveland, B. and J. Mayben. 1997. *Call center management on fast forward*, Call Center Press, Annapolis, MD.

10. Easton, F.F and J. Goodale. 2005. Schedule recovery: Unplanned absences in service operations, *Decision Sciences* **36** (3) 459–488.
11. Ernst, A.T., H. Jiang, M. Krishnamoorthy, B. Owens, and D. Sier. 2004. An annotated bibliography of personnel scheduling and rostering, *Annals of Operations Research* **127** (1-4) 21 – 144.
12. Fylstra, D., L. Lasdon, J. Watson, and A. Waren. 1998. Design and use of the Microsoft Excel Solver, *Interfaces*, **28** (5) 29–55.
13. Gans, N., G. Koole, and A. Mandelbaum. 2003. Telephone call centers: tutorial, review and research prospects, *Manufacturing and Service Operations Management* **5** (2), 79–141.
14. Gans , N. and Y.-P. Zhou. 2007. Call-routing schemes for call-center outsourcing. *Manufacturing & Service Operations Management* **9** (1) 33-50.
15. Grassman,W.K. 1988. Finding the right number of servers in real-world queueing systems, *Interfaces* **18** (2) 94–104.
16. Green, L. V., P. J. Kolesar, and J. Soares. 2001. Improving the SIPP approach for staffing service systems that have cyclic demands, *Operations Research* **49** (4) 549–564.
17. Green, L. V., P. J. Kolesar, and J. Soares. 2003. An improved heuristic for staffing telephone call centers with limited operating hours, *Production and Operations Management* **12** (1), no. 1, 1–16.
18. Guerrero, V. M. and J. A. Elizondo. 1997. Forecasting a cumulative variable based on its partially accumulated data, *Management Science* **43** (6) 879–889.
19. Harrison, J. M. and A. Zeevi. 2005. A method for staffing large call centers using stochastic fluid models, *Manufacturing & Service Operations Management* **7** (1) 20-36.
20. Hur, D. 2002. A comparative evaluation of forecast monitoring systems in service organizations, 33rd Annual Meeting of the Decision Sciences Institute, San Diego, CA, USA.
21. Hur, D., V.A. Mabert, and K.M. Bretthauer. 2004. Real-time work schedule adjustment decisions: An investigation and evaluation, *Production and Operations Management* **13** (4) 322–339.

22. Jongbloed, G. and G. Koole. 2001. Managing uncertainty in call centers using Poisson mixtures, *Applied Stochastic Models in Business and Industry* **17** (2001), 307–318.
23. Kekre, S., T.E. Morton, and T.L. Smunt. 1990. Forecasting using partially known demands, *International Journal of Forecasting* **6** (1) 115–125.
24. Mabert, V.A. 1991. *Ohio valley national bank*, Cases in Manufacturing and Services Management, Prentice Hall, Englewood Cliffs, New Jersey.
25. Mabert, V.A. 1995. *Psi energy's call center*, unpublished case study, Indiana University, IN.
26. Mehrotra, V. 1997. Ringing up big business, *OR/MS Today* **24** (4) 18–24.
27. Mehrotra, V., C. Wright, and S. Patel. 2009. An investigation into the business processes, job descriptions, and human resource factors associated with successful workforce management practice, Working Paper, San Francisco State University, San Francisco, CA.
28. J. Milner and T. Lennon-Olson. 2008. Service Level Agreements in Call Centers: Perils and Prescriptions. *Management Science* **54** (2) 238-252.
29. Pot, A. and G. Koole. 2005. Approximate dynamic programming for multi-skill call centers, *Proceedings of the 2005 Winter Simulation Conference*, Piscataway, NJ.
30. Robbins, T. R. 2007. Addressing Arrival Rate Uncertainty in Call Center Workforce Management, *Proceedings of the 2007 IEEE/INFORMS International Conference on Service Operations and Logistics, and Informatics*, Philadelphia, PA.
31. Saltzman, R., 2005. A hybrid approach to minimizing the cost of staffing a call center, *International Journal of Operations and Quantitative Management* **11** (1) 1–14.
32. Shen, H. and J. Z. Huang. 2008. Interday forecasting and Intraday updating of call center arrivals, *Manufacturing and Service Operations Management* **10** (3) 391-410.
33. Steckley, S., S. Henderson, and V. Mehrotra. 2004. Service system planning in the presence of a random arrival rate, working paper, Department of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY.

34. Steckley, S., S. Henderson, and V. Mehrotra. 2009. Forecast Errors in Service Systems. *Probability in the Engineering and Information Sciences* 23 (2) 305-332.
35. Thompson, G. M. 1996. Controlling actions in daily workforce schedules, *Cornell Hotel and Restaurant Administration Quarterly* 37 (2) 82–96.
36. Thompson, G. M. 1998. Labor scheduling, part 1: forecasting demand, *Cornell Hotel and Restaurant Administration Quarterly* 39 (5) 22–31.
37. Thompson, G. M. 1999. Labor scheduling, part 4: Controlling workforce schedules in real-time, *Cornell Hotel and Restaurant Administration Quarterly* 40 (3) 85–96.
38. Wallace, R. and W. Whitt. 2005. A Staffing Algorithm for Call Centers with Skill-Based Routing. *Manufacturing and Service Operations Management* 7 (4) 276-294.
39. Weinberg, J., J., L. D. Brown, and J. R. Stroud. 2007. Bayesian Forecasting of an Inhomogeneous Poisson Process with Applications to Call Center Data. *Journal of the American Statistical Association* 102 (480) 1185-1198.
40. Whitt, W. 1999. Dynamic staffing in a telephone call center aiming to immediately answer all calls, *Operations Research Letters* 24 (5) 205–212.
41. Whitt, W. 2002. Stochastic models for the design and management of customer contact centers: some research directions, Working Paper, Columbia University, New York, NY.
42. Whitt, W. 2002. *Stochastic-process limits*, Springer Series in Operations Research, Springer, New York.
43. Whitt, W. 2004. Staffing a call center with uncertain arrival rate and absenteeism, working paper, Department of Industrial Engineering and Operations Research, Columbia University, New York, NY.

APPENDIX A: DETAILS OF PARAMETERS FOR THE NUMERICAL RESULTS

Since \mathbf{B} (a 60 by 32 matrix) is too large to display properly, we have provided the start period, end period and break times for each agent type. The first five and the last five agent types are part-time schedules. Part-time agents take one 15-minute break, while full-time workers take one 15-minute, one 30-minute and a final 15-minute break.

$$T = 60 \text{ (each period is 15-minutes)}$$

$$N = 32 \text{ (10 part-time and 22 full-time agent types)}$$

$$m_i = 1, 4, 5, 0, 0, 1, 1, 4, 1, 0, 2, 5, 1, 0, 0, 0, 5, 0, 0, 4, 0, 2, 1, 2, 1, 6, 0, 0, 0, 2, 1, 4$$

$$d_i = 3, 3, 8, 8, 14, 14, 20, 19, 24, 24, 26, 25, 27, 30, 25, 25, 29, 27, 28, 28, 24, 26, 27, 27, 35, \\ 29, 32, 32, 30, 30, 32, 34, 30, 34, 28, 32, 27, 31, 23, 27, 21, 21, 16, 16, 23, 23, 22, 20, 23, \\ 23, 16, 17, 17, 17, 17, 17, 13, 13, 10, 10$$

$$F_i = 1, 5, 9, 13, 17, 1, 1, 3, 3, 5, 5, 7, 7, 9, 9, 11, 11, 17, 17, 19, 19, 23, 23, 25, 25, 27, 27, 28, \\ 32, 36, 40, 44$$

$$L_i = 17, 21, 25, 29, 33, 34, 34, 36, 36, 38, 38, 40, 40, 42, 42, 44, 44, 50, 50, 52, 52, 56, 56, 58, \\ 58, 60, 60, 44, 48, 52, 56, 60$$

$$\text{Break 1} = 8, 12, 16, 20, 26, 9, 12, 11, 10, 15, 13, 15, 13, 18, 17, 20, 21, 26, 25, 28, 27, 32, 31, 34, \\ 33, 35, 34, 36, 40, 44, 48, 52$$

$$\text{Break 2} = -, -, -, -, -, 17, 20, 19, 18, 23, 21, 23, 21, 26, 25, 28, 29, 34, 33, 36, 35, 40, 39, 42, 41, 43, \\ 42, -, -, -, -, -$$

$$\text{Break 3} = -, -, -, -, -, 29, 30, 27, 26, 33, 31, 33, 31, 36, 35, 38, 39, 42, 41, 44, 43, 48, 47, 52, 51, 51, \\ 52, -, -, -, -, -$$

The forecasted service rates per period (μ_t) are:

$$\mu_t = 48, 48, 28, 28, 14, 14, 11, 11, 11, 11, 11, 11, 10, 10, 11, 11, 11, 11, 11, 11, 10, 10, 9.9$$

,9.9, 10, 10, 10, 10, 11, 11, 10, 10, 11, 11, 10, 10 ,9.8 ,9.8, 11, 11, 11, 11, 10, 10, 11, 11,
10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 11, 11 ,9.1 ,9.1

TABLES

Table 1: Summary Results for the Days Identified As Understaffed

Performance Measure	“End Of Day” Comparison Based on Actual Call Arrivals
Mean Service Level	
With the Updated Schedule	83.50%
With the Original Schedule	89.07%
Δ	5.57%
Mean Cost of Not Meeting SL Targets	
With the Updated Schedule	\$ 12,798
With the Original Schedule	\$ 19,112
Δ	-\$ 6,314
Payroll Impact of Update	\$ 3,902
Disruption Impact of Update	\$ 61

Table 2: Summary Results for the Days Identified As Overstaffed

Performance Measure	“End Of Day” Comparison Based on Actual Call Arrivals
Service Level	
With the Updated Schedule	81.33%
With the Original Schedule	97.37%
Δ	-16.04%
Cost of Not Meeting SL Targets	
With the Updated Schedule	\$ 18,505
With the Original Schedule	\$ 2,693
Δ	\$ 15,812
Payroll Impact of Update	-\$ 2,015
Disruption Impact of Update	\$ 335

Table 3: Summary Results for the Days Identified As Understaffed with different k_1 values

k_1	SL w/o update	SL w/ update	Cost of Calls Lost w/o Update (\$)	Cost of Calls Lost w/ Update (\$)	Cost of Update (\$)	Cost of Disruption (\$)	Net Benefit (\$)
0.25	83.50%	91.34%	19,112	10,189	4,678	72	4,172
0.50	83.50%	92.32%	19,112	9,066	5,374	77	4,595
0.75	83.50%	93.36%	19,112	7,865	6,131	89	5,027
1.00	83.50%	94.43%	19,112	6,643	6,926	97	5,445
1.25	83.50%	95.53%	19,112	5,313	7,752	104	5,943
1.50	83.50%	96.30%	19,112	4,417	8,600	108	5,987
1.75	83.50%	96.68%	19,112	3,980	9,466	112	5,554
2.00	83.50%	97.07%	19,112	3,512	10,320	118	5,161
2.25	83.50%	97.61%	19,112	2,862	11,199	125	4,926
2.50	83.50%	97.72%	19,112	2,710	12,080	125	4,197

Table 4: Summary Results for the Days Identified As Overstaffed with different k_2 values

k_2	SL w/o update	SL w/ update	Cost of Calls Lost w/o Update (\$)	Cost of Calls Lost w/ Update (\$)	Cost of Update (\$)	Cost of Disruption (\$)	Net Benefit (\$)
0.25	97.37%	84.76%	2,693.02	15,127.92	(1,790.14)	327.86	(10,972.61)
0.50	97.37%	87.43%	2,693.02	12,573.83	(1,562.14)	321.79	(8,640.45)
0.75	97.37%	89.25%	2,693.02	10,806.28	(1,333.71)	311.79	(7,091.33)
1.00	97.37%	90.94%	2,693.02	9,128.89	(1,141.29)	294.29	(5,588.87)
1.25	97.37%	92.96%	2,693.02	7,105.36	(937.71)	279.29	(3,753.92)
1.50	97.37%	94.41%	2,693.02	5,668.62	(753.86)	265.71	(2,487.46)
1.75	97.37%	95.48%	2,693.02	4,580.73	(603.00)	249.29	(1,534.00)
2.00	97.37%	96.23%	2,693.02	3,841.74	(459.43)	231.07	(920.37)
2.25	97.37%	96.77%	2,693.02	3,295.02	(342.86)	212.14	(471.29)
2.50	97.37%	97.06%	2,693.02	3,000.02	(232.71)	181.79	(256.08)

FIGURES

Figure 1: Three Level Hierarchical Model of Call Center Resource Planning

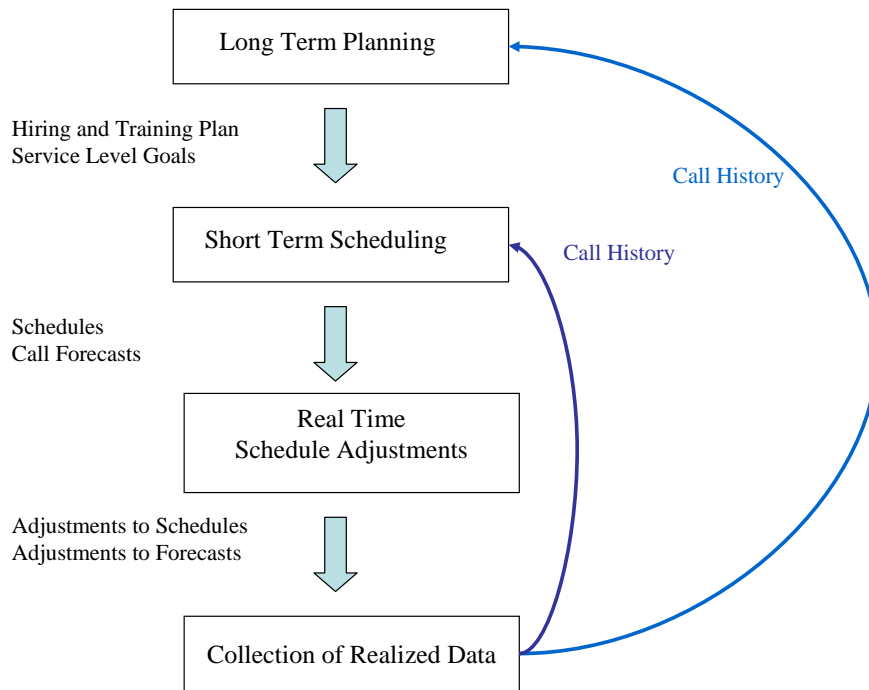


Figure 2: Rescheduling Framework and Experimental Workflow

