Articles and Chapters   School of Hotel Administration Collection

9-1992

Improving the Utilization of Front-Line Service Delivery System Personnel

Gary Thompson
Cornell University School of Hotel Administration, gmt1@cornell.edu

Improving the Utilization of Front-Line Service Delivery System Personnel

Abstract

There are two types of work typically performed in services which differ in the degree of control management has over when the work must be done. Serving customers, an activity that can occur only when customers are in the system is, by its nature, uncontrollable work. In contrast, the execution of controllable work does not require the presence of customers, and is work over which management has some degree of temporal control. This paper presents two integer programming models for optimally scheduling controllable work simultaneously with shifts. One model explicitly defines variables for the times at which controllable work may be started, while the other uses implicit modeling to reduce the number of variables. In an initial experiment of 864 test problems, the latter model yielded optimal solutions in approximately 81 percent of the time required by the former model. To evaluate the impact on customer service of having front-line employees perform controllable work, a second experiment was conducted simulating 5,832 service delivery systems. The results show that controllable work offers a useful means of improving labor utilization. Perhaps more important, it was found that having front-line employees perform controllable work did not degrade the desired level of customer service.

Keywords

linear programming, manpower planning, service operations management, simulation

Disciplines

Hospitality Administration and Management

Comments

Required Publisher Statement

*This paper was a finalist for the 1993 First Annual "Best Paper of the Year" given by the David Eccles School of Business.

Improving the Utilization of Front-Line Service Delivery System Personnel

Gary M. Thompson

David Eccles School of Business

University of Utah

1655 East Campus Center Drive

Salt Lake City, UT 84112

ABSTRACT

There are two types of work typically performed in services which differ in the degree of control management has over when the work must be done. Serving customers, an activity that can occur only when customers are in the system is, by its nature, uncontrollable work. In contrast, the execution of controllable work does not require the presence of customers, and is work over which management has some degree of temporal control. This paper presents two integer programming models for optimally scheduling controllable work simultaneously with shifts. One model explicitly defines variables for the times at which controllable work may be started, while the other uses implicit modeling to reduce the number of variables. In an initial experiment of 864 test problems, the latter model yielded optimal solutions in approximately 81 percent of the time required by the former model. To evaluate the impact on customer service of having front-line employees perform controllable work, a second experiment was conducted simulating 5,832 service delivery systems. The results show that controllable work offers a useful means of improving labor utilization. Perhaps more important, it was found that having front-line employees perform controllable work did not degrade the desired level of customer service.

Keywords: Linear Programming, Manpower Planning, Service Operations Management, and Simulation

Improving the Utilization of Front-Line Service Delivery System Personnel

INTRODUCTION

Front-line service delivery employees have as their primary duty the serving of customers. Since customer contact activities cannot be inventoried, these employees must be ready to serve the arriving customers. In some services, appointment systems offer a means of scheduling customers. In many more services, however, the arrival of customers is random, with a mean underlying customer arrival rate typically following some predictable pattern arising from the nature of the service. In the latter services, which are the focus of this paper, serving customers may be thought of as uncontrollable work (UW): customers cannot be served before they arrive, while service can be delayed only at the peril of customer dissatisfaction.

Even in services that respond to, rather than schedule, customer arrivals, some employees (not necessarily the front-line employees) normally do work over which management has some temporal control. We choose to call this controllable work (CW), since management has control, within limits, over when to do it. In grocery and convenience stores, serving customers is UW for the front-line employees (the cashiers), while CW may include such activities as cleaning the work area, maintaining equipment, and stocking shelves. For police patrol persons, responding to citizens' calls is UW while CW may include such duties as traffic control (speed traps) and monitoring neighborhoods.

Labor scheduling is the balancing act of attempting to have the correct number of front-line employees on hand to provide good customer service at all times. Effective labor scheduling is crucial to the efficient use of service delivery labor. Too few employees working results in poor, customer-annoying service, and may result in the loss of sales. Too many front-line employees working is undesirable since the customer contact activities cannot be inventoried. Labor scheduling has received a good deal of attention in the literature, which is attributable to the importance of services in the modem economy and the impact of labor scheduling on the overall efficiency of service organizations.

In the remainder of this paper relevant literature is reviewed, the optimal models for simultaneously scheduling shifts and CW are presented, the results of an initial experiment to evaluate the relative performance of the two optimal models are described and presented, and the

results of a second experiment to evaluate the impact on customer service of having front-line employees carry out CW are then presented. The paper concludes with a discussion of the generalizability and implications of the findings.

LITERATURE REVIEW

A trend toward increasing the flexibility employed by scheduling procedures characterizes much of the labor scheduling literature. Greater flexibility typically enables a better matching of the number of staff scheduled to the number of staff needed. Scheduling flexibility may take many forms: (1) large break windows (the times within shifts when breaks may be taken); (2) more shift-length alternatives; (3) greater start-time variation for shifts; (4) split-shifts (shifts with meal breaks exceeding 1.5 hours); (5) large start-time floats (the variation in starting times for an individual employee over a week); and (6) fewer days worked per week. Henderson and Berry [11] [12] found that fewer employees could satisfy customer demand with either an increase in the number of shift length alternatives or an increase in the range of shift starting times. Mabert and Watts [17] and Bailey and Field [3] had similar findings. Bailey [2] found that start-time floats reduced schedule costs, while Showalter and Mabert [23] found that allowing employees to work shorter shifts and fewer days per week improved tour schedules (schedules developed for a week at a time). Bechtold and Jacobs [6] presented an optimal shift scheduling model that implicitly incorporates breaks. They used this model [7] to evaluate the flexibility offered by more shift lengths and enlarged break windows, and found synergistic effects between the two.

The number of alternate shifts (or tours) provides a simple aggregate measure of the degree of scheduling flexibility existing in an environment. Table 1 shows that the number of distinct shifts (tours) has ranged from 26 to 6,588 (30 to 35,840) in the labor scheduling literature.

To date, the labor scheduling literature has focused implicitly on the scheduling of front-line employees doing UW. We extend the literature by broadening the shift scheduling problem to include the scheduling of CW for front-line employees.

TWO OPTIMAL INTEGER PROGRAMMING MODELS
FOR SIMULTANEOUSLY SCHEDULING SHIFTS
AND CONTROLLABLE WORK

The following subsections define relevant terms, present two optimal integer programming (IP) models for scheduling shifts and CW, and discuss the assignment of CW to shifts.

Terms

The following terms are used throughout the paper:

Planning period is the basic unit of time used in schedule development.

Type of controllable work is the set of different work functions or activities over which management has some temporal control.

Block of work is the standard duration of assignment of a single employee to a specific type of CW, measured in planning periods, and including any unproductive time required for switching between UW and CW.

Commencement window is the set of planning periods within which a block of CW must be started.

Starting time is a time (the beginning of a period) at which shifts, breaks, or blocks of CW may commence.

Model 1 (M1)

Define the following variables

$$c_{bj} = \begin{cases} 1, \text{if CW block } b \text{ commences in planning period } j \\ \qquad\qquad 0, \text{otherwise,} \end{cases}$$

x_n = the number of employees assigned to shift n.

With these variables, M1 is

$$\text{Minimize } Z = \sum_{n \in N} v_n \, x_n \qquad\qquad (1)$$

Subject to

$$\sum_{n \in N} a_{np} x_n - \sum_{b \in B} \sum_{j \in Abp} c_{bj} \geq r_p \qquad for\ p \in P \qquad (2)$$

$$\sum_{j=e_b}^{f_b} c_{bj} = 1 \qquad\qquad for\ b \in B \qquad (3)$$

$$x_n \geq 0 \ and\ integer \qquad\qquad for\ n \in N \qquad (4)$$

$$c_{bj} = 0,1 \qquad\qquad for\ b \in B\ and\ j = e_b, \dots, f_b, \quad (5)$$

Where data constants are

a_{np} = 1, if planning period p is a working period for shift n,

      0, otherwise,

d_b = the duration, in planning periods, of CW block b,

e_b = the earliest possible starting time for CW bock b,

r_p = the number of employees required in period p to perform the UW at the

      desired level of service, and

v_n = the relative cost of assigning an employee to work shift n,

And sets are

    B = the set of blocks of CW

    N = the set of staffable shifts,

    P = the set of planning periods in a work day,

    A_{bp} = the set of starting times for CW block b that result in a requirement for employees in

      period p

$$= \{j \mid \max(p - d_b + 1, e_b) \leq j \leq \min(p, f_b + d_b - 1)\}$$

Ml's objective (1) is to minimize the total cost of the scheduled shifts. Constraint set (2) satisfies employee requirements arising from both CW and UW. Constraint set (3) requires that each block of CW commences within its appropriate window. Constraint set (4) specifies the integrality of the shift variables, while constraint set (5) does the same for the CW variables. The Appendix presents a simple example of Ml.

Excluding the CW term in constraint set (2), the model defined by (1), (2), and (4) is the set covering formulation of the shift scheduling problem frequently observed in the literature [1] [3] [4] [5] [6] [7] [8] [11] [12] [15] [20] [22] [23]. Examining constraint set (2), one may observe that the scheduling of CW is conceptually very similar to Bechtold and Jacobs' [6] [7] scheduling of implicit breaks: The number of employees doing CW (or on break) represents a usage (sink) of employees.

Model 2 (M2)

M2 makes use of the insight that blocks of each type of CW are undifferentiated, except for their commencement windows. Instead of defining variables based on the times at which each

block of CW may start (like the c_{bj} variables in Ml), M2 defines variables based on the times at which blocks of each CW type may start:

S_{tj} = the number of blocks of type-r CW starting in planning period j.

If there is any overlap in the commencement windows for blocks of a CW type, then M2 will require fewer integer variables than Ml. M2 is

$$\text{Min } Z = \sum_{n \in N} v_n x_n$$

Subject to

$$\sum_{n \in N} a_{np} x_n - \sum_{t \in T} \sum_{j \in R_{tp}} S_{tj} \geq r_p \qquad for\ p \in P \qquad (6)$$

$$\sum_{j \in Y} S_{tj} = \sum_{w \in W} m_{tw} \qquad for\ t \in T \qquad (7)$$

$$\sum_{\{j \in Y_t, j < p\}} S_{tj} \leq \sum_{\{w \in W_t, g_{tw} < p\}} m_{tw} \quad for\ t \in T\ and\ \{p\ |p > k_t \quad (8)$$

$$and\ p = g_{tw}\ for\ w \in W_t\},$$

$$\sum_{\{j \in Y_t, j < p\}} S_{tj} \leq \sum_{\{w \in W_t, h_{tw} > p\}} m_{tw} \quad for\ t \in T\ and\ \{p\ |p < q_t \quad (9)$$

$$and\ p = h_{tw}\ for\ w \in W_t\},$$

$$S_{tj} \geq 0\ and\ integer \qquad for\ t \in\ T\ and\ j \in Y_t, \qquad (10)$$

and

$$x_n \geq 0\ and\ integer \qquad for\ n \in\ N$$

where data constants are

g_{tw} = the earliest period in commencement window w of CW type t,

h_{tw} = the latest period in commencement window w of CW type t,

k_t = the earliest possible starting period for any block of CW of type t,

m_{tw} = the number of blocks of CW type t with commencement window w,

q_t = the latest possible starting period for any block of CW of type t, and

u_t = the duration, in planning periods, of any CW block of type t.

Sets are

R_{tp} = the set of start times for CW blocks of type t that result in requirement for

for employees in period p,

$$= \{j \mid j \in Y_t \ and \ p - u_t + 1 \leq \min \ (p, q_t + u_t - 1)\},$$

T = the set of CW types

Y_t = the set of possible start times for CW blocks of type t, and

W_t = the set of commencement windows for CW type t,

with N, P, a_{np}, r_p, and x_n as earlier defined.

Constraint set (6) satisfies the employee requirements arising from both CW and UW. Constraint set (7) fixes the number of blocks of each CW type. Constraint sets (8) and (9) ensure the blocks of each CW type neither start earlier or later than allowed, respectively. Constraint set (10) specifies the integrality of the CW variables. The Appendix presents a simple example of M2.

Post-Solution Assignment of Controllable Work Blocks to Shifts

Both Ml and M2 make use of implicit modeling in the sense that neither model explicitly matches CW to shifts. Such matching is necessary, however, before schedules developed by either model can be applied. Our procedure for doing this assigns CW blocks in the order of their scheduled starting times, with ties broken randomly. A block is assigned to the shift for which it comes closest to the beginning or end of a work stretch (the periods of work before or after a break), with ties broken randomly. Having CW at the beginning or end of a work stretch is beneficial since it reduces the number of work set-ups—the non-productive time required to switch between UW and CW—that employees must undertake within their work stretches. Although both models schedule enough labor to cover the CW, blocks may have to be split among employees (shifts) during the assignment process. This issue is discussed later.

EXPERIMENT 1—EVALUATING THE RELATIVE PERFORMANCE OF THE OPTIMAL MODELS

Experiment 1 (EXP1), which was designed with the goal of evaluating schedule generation times of Ml and M2, contained an extensive array of test problems. The following sections describe the structure and present the results of EXP1.

Structure of EXP1

Within this section the scheduling environment of EXP1 is described, experimental factors are identified, and details regarding schedule generation are provided.

Scheduling Environment. In developing the scheduling test environment, hour- long planning periods and a 20-hour operating day were arbitrarily selected. To highlight the value of CW in improving labor utilization, a very inflexible scheduling environment was desired and one was achieved by using only 9-hour shifts with an hour-long meal break taken in the fifth hour. Indeed, as these restrictions resulted in only 12 distinct shifts, the scheduling environment had exceedingly limited flexibility compared to the range of alternate shifts and tours considered in the labor scheduling literature (see Table 1).

Experimental Factors. Table 2 identifies the eight factors of EXP1. Three of the factors related to UW and four related to CW; the final factor was the optimal IP model. The diversity of these factors ensures that, in aggregate, EXP1 represents conditions existing in a wide range of service organizations. Representing conditions existing in a wide range of service organizations is desirable because it helps to ensure that a narrowly-effective model does not outperform a broadly-effective model.

One UW-related experimental factor identifies the shape of the employee requirements (RP), another the variability in the requirements (RV), and the third, the magnitude of the requirements (MR). RP had four levels (patterns), representing employee requirement patterns commonly observed in service organizations: unimodal (1 daily peak), bimodal (2 daily peaks), trimodal (3 daily peaks), and random (many daily peaks). RV, measured as the coefficient of variation of the employee requirements, had two levels: .2 and .6. Combining the eight resultant patterns with MR values of 5 and 20 employees needed every hour, on average, resulted in a total of 16 UW employee requirement curves, as illustrated in Figure 1.

Four experimental factors related to CW: the number of CW types (WT), the CW volume (LH), the CW block length (BL), and the CW timing flexibility (TF). WT was selected because M2 requires new variables and constraint sets for each additional type of CW. Thus, the hypothesis is that compared to M2, Ml would function better with greater numbers of CW types. LH enabled the investigation of the effect on labor utilization of increasing the amount of CW to be carried out by the front-line employees. Varying the CW timing flexibility is desirable for two reasons. First, increasing the flexibility in scheduling the CW should result in higher labor utilization. Second,

increasing the timing flexibility has more impact on the number of integer variables in Ml than in M2. Thus, it is expected that compared to Ml, M2 will generate optimal schedules faster when CW has greater timing flexibility. BL was selected for two reasons. First, shorter-duration blocks are expected to be more useful in improving labor utilization. Second, it enables one to measure how often the longer blocks are split between shifts during the assignment of CW to shifts.

The eight factors, and the levels of these factors, result in EXP1 having a total of 864 shift and CW scheduling problems. For every problem, both Ml and M2 generated optimal schedules, thus giving 1,728 observations in EXP1.

Schedule Generation Details. All investigations were conducted on an 80486-based personal computer, operating at 33 MHz. A two-phase procedure based on [20] was very useful in reducing the mean and maximum times of Ml's and M2's solution to optimality. First, the procedure solved the relaxed linear programming version of a model using SAS-OR [21]. From this solution, the total number of shifts scheduled was determined. Let this quantity equal TS_{lp}. A constraint was then added to the model and solved a second time in its IP form using the branch and-bound procedure of SAS-OR [21]. The additional constraint served to set the sum of the scheduled shifts to equal or exceed the smallest integer at least as large as TS_{lp}.

Results

Table 3 presents selected results from an ANOVA analysis using schedule generation time as a dependent variable. For the 864 problems in EXP1, Ml and M2 generated schedules in an average of 31.49 and 25.42 seconds, respectively, a difference significant at the .0001 level. It appears that the problem of splitting blocks across shifts is slight since, for each model, only six of a possible 5,040 CW blocks had to be split across shifts.

Figure 2 illustrates significant MD-based first-order interaction effects for schedule generation times. Ml only had a lower mean schedule generation time than M2 when the CW timing flexibility was low. M2's schedule generation time advantage over Ml was greater with (a) shorter CW blocks (see Figure 2a); (b) greater volumes of CW (see Figure 2b); greater flexibility in scheduling CW (see Figure 2c); and (d) fewer types of CW (see Figure 2d). These results are generally consistent with expectations.

EXPERIMENT 2—EVALUATING THE IMPACT

Experiment 2 (EXP2) was designed with the goal of evaluating the impact on customer service and labor utilization of having the front-line employees carry out both UW and CW. To this end, EXP2 is broader in scope than EXP1. First, EXP2 includes the setting of the UW employee requirements. Second, EXP2 simulates service delivery systems to measure both customer service and actual labor usage. Because of its superiority over Ml, only M2 is used in generating optimal schedules in EXP2. In the subsections that follow, the structure and results of EXP2 are presented.

Structure of EXP2

The scheduling environment in EXP2 is identical with that used in EXP1. The following subsections identify the experimental factors; describe the process of simulating service delivery systems; specify assumptions of the simulation; and identify performance measures.

Experimental Factors. The factors of EXP2, which Table 4 identifies, can be categorized as relating to UW and to CW. The three CW-related factors—BL, TF, and LH—are the same as in EXP1, and EXP2 includes them for the reasons presented earlier. In EXP2, all blocks of CW are of the same type (this is equivalent to WT=1 in EXP1).

Five factors related to UW: the true, or underlying, customer arrival-rate pattern (AP), the variation in the underlying customer arrival-rate pattern (AV), the mean duration of customer service (SD), the desired level of customer service (SL), and the accuracy of customer arrival forecasts (FA). AP had three levels: a unimodal pattern, a bimodal pattern, and a trimodal pattern. Sinusoidal curves were used to allow the true customer arrival rate to change continuously over the operating day. Because customer arrivals to the system were simulated using randomly and exponentially distributed interarrival times, the actual arrival rate of customers to the system in any simulated day would typically be very different from the true rate. AV, measured by the coefficient of variation in the underlying customer arrival-rate pattern, had levels of .25 and .50. Combining the levels of AP and AV yielded 6 customer arrival-rate curves, each having average arrivals of 60 customers per hour, as illustrated in Figure 3.

SD had three levels—1, 4, and 16 minutes. Simulated customer service times were exponentially distributed to reflect the high degree of variability in service times commonly occurring in service organizations. In our investigation, we had the goal of serving 90 percent of

customers with maximum waiting times, for the three levels of SL, of 3, 1, and 1/3 minutes.

FA was selected as a factor to see if CW would be useful only in theory. If having employees perform CW did not harm customer service when arrival forecasts were highly accurate, but hurt customer service when arrival forecasts were less accurate, one could make such a judgment. For the three levels of FA, we used historical information from 1, 4, and 16 days of simulated past arrivals to the service system in developing expected customer arrivals for each planning period. Obviously, using more information should result in better forecasts.

Simulation Process. For each of two replications of the 6 customer arrival-rate curves, 56 days worth of information on the service delivery system were generated and stored in data files. Sixteen days were historical, while the remaining days enabled the simulation of the service system's future operation. An explanation of each follows. An average number of past customer arrivals to the service system was calculated for each planning period in the operating day using information appropriate to the level of FA. The mean number of customer arrivals in a period yielded an average arrival rate that, in turn, was an input to an M/M/c queuing model. The number of employees needed for the UW (the right-hand side of equation (6)) was the smallest staff size serving 90 percent of customers within the time limit specified by the level of SL.

The 40 days of data used in simulating the future operation of the service system were, for each customer, stored as an arrival time and a service-duration determining random number. To control variance, the 12 customer arrival data sets were applied with all combinations of FA, SD, SL, BL, TF, and LH. Elaborating on this, the 27 combinations of FA, SD, and SL, for each of 12 customer arrival data sets, yielded a total of 324 UW-employee-requirement scenarios. These 324 scenarios, combined with the 18 variations of BL, TF, and LH, yielded 5,832 service environments. M2 was used to generate and the methodology of simulation to evaluate schedules for each service environment. As with EXP1, all investigations were conducted on an 80486-based personal computer. The simulation model was coded in FORTRAN while the optimal M2 schedules were obtained using the process outlined earlier.

Simulation-Related Assumptions. In conducting the simulation experiment, it was assumed: (1) no changes in the schedule during its implementation; (2) employees worked as scheduled (no absenteeism, for example); (3) no customers left the queue; (4) a constant mean service time across the simulation period (the employees do not speed up if the facility is busy, for example); and (5) stationary underlying customer arrival-rate curves. The relaxation of these

assumptions is addressed later.

Performance Measures. Three measures of service system performance were selected: (1) PM1—the percentage of customers served within the desired maximum waiting time; (2) PM2—the average actual idle time of the front-line employees, in hours; and (3) PM3—the utilization of the front-line employees. PM2 equals the total scheduled front-line labor hours, less the total hours the front-line employees spend doing CW and UW (in the latter case, a quantity lower than the sum of right-hand sides from constraint set (6) because of the idle time implicit in the specified service level). A key assumption in measuring the value of CW is that any CW moved from the back-shop to the front-line employees yields a one-for-one savings in the required back-shop labor.

PM1 gauges the impact on customer service of CW, while PM2 and PM3 measure the effects of CW on labor usage. Both PM2 and PM3 are useful indicators, since higher utilization of the front-line employees (PM3) could arise with increasing volumes of CW simply from the same unutilized time spread across greater total labor hours (in which case there would be no change in PM2).

Results

Table 5 presents significant CW-based terms from an ANOVA model of the PM1 results from EXP2. As CW volume (LH) increased from 10 to 20 to 40 labor hours, the actual level of service, measured as the percentage of customers served within the specified limit on waiting time, fell from 96.53 percent to 96.03 percent to 95.44 percent. Figure 4 illustrates LH-based first-order interaction effects for PM1. The most important observation from this figure is that the actual level of customer service always exceeded the desired level (recall that 90 percent of customers should be served within the specified limit on waiting time).

Table 6 shows that lower idle times occurred with (a) higher volumes of CW, (b) greater flexibility in CW timing, and (c) shorter blocks of CW. This table also shows that scheduling x additional hours of CW reduced the front-line employees' idle time less than x hours, and that this conversion factor exhibited diminishing returns (increasing CW from 10 to 20 hours reduced idle time by 7.13 hours, but further increasing CW from 20 to 40 hours only reduced idle time by 11.00 additional hours). Given the assumption that any CW moved from the back-shop to the front-line employees yields a one-for-one savings in back-shop labor hours, the service system, as a whole,

reduced its paid labor hours by 7.13 (11.00) hours as the volume of CW done by front-line employees increased from 10 to 20 (20 to 40) hours. PM3 results are consistent with those for PM2. For example, increasing the CW from 10 to 20 to 40 hours raised actual average utilization of the front-line employees from 51.14 percent to 54.02 percent to 58.97 percent.

Finally, Table 6 shows a diminishing impact of increasing the CW timing flexibility. Lengthening the average commencement window from 2 to 4 hours reduced idle time by 5.8hours, but further lengthening the average commencement window to 8 hours only lowered idle time by an additional .44 hours.

Table 7 presents significant CW-based terms from an ANOVA model for PM2. Figure 5 illustrates CW-based interactions for PM2. As expected, shorter duration blocks were more effective in reducing idle time. By far the greatest difference in idle time between blocks of 1 and 2 hours occurred with low CW timing flexibility, as illustrated in Figure 5a. Figure 5b shows that the greatest benefits of increasing CW timing flexibility occurred with higher volumes of CW. Finally, Figure 5c illustrates that higher volumes of CW were more useful in reducing idle time when the service duration, and hence the mean UW employee requirements, were higher.

DISCUSSION

This section offers a discussion focusing on the relative performance of Ml and M2, the likely impact of relaxing the simulation-related assumptions, managerial issues associated with the use of CW, and possible research extensions.

Relative Model Performance

The results from EXP1 show the broad superiority of M2 over Ml. M2 generated schedules in approximately 81 percent of the time required by Ml, a difference significant at the .0001 level. Ml only generated schedules more quickly than M2 with low flexibility in the timing of CW, a not-unexpected result. Recall that M2, by defining variables for the times at which blocks of a particular type of CW may start, will have fewer variables than Ml when block commencement windows overlap. Clearly, the degree of commencement window overlap will be lower with lower levels of CW timing flexibility.

Although M2's superiority is consistent with the results of other researchers using implicit modeling to reduce the required number of variables [6], M2's relative advantage

compared to Ml is lower than for other implicit models [6], the reason for this is that Ml did not require a huge number of integer variables in the problems comprising EXP1. It is to be expected, however, that M2's relative superiority would increase the duration of planning periods decreased to 30 or 15 minutes.

Impact of the Simulation Assumptions in EXP2

Relaxing the first assumption (implementing the schedule as planned) would require the use of real-time control activities. Real-time control seeks to lower labor costs by such actions as sending employees home early (without pay) should excess staff be present at the current time, or, if customer demand is straining the system at the current time, to lower customers' waiting times (and in doing so raise the level of service) by an action such as extending the length of a shift. Real-time control activities are beneficial due to the natural variation in customer arrival and service times. Obviously, then, CW and real-time control are complementary, since CW uses the idle labor that occurs consistently during the same periods, and that arises because the number of employees scheduled cannot be matched precisely to the number of employees needed to carry out the UW.

Relaxing the other simulation-related assumptions would have one or both of two effects: real-time control would become desirable and the simulation would be necessarily more complex. Assumptions that, when relaxed, have the former effect include the employees working as scheduled, stationary underlying customer arrival- rate curves, and no customer balking. Presuming that the desired service level is great enough to make balking an uncommon occurrence, then balking will occur only when the facility is unusually busy. These busy times will occur, regardless of the use of CW, because of the natural variation in customer demand.

Assumptions that, when relaxed, have the latter effect include the employees working as scheduled, a constant mean service time, and stationary underlying customer arrival-rate curves. Should the assumption of the employees working as scheduled be relaxed, the desired level of service used in setting the UW employee requirements may be inflated so that the resultant actual level of service approximates the true desired level. Relaxing the assumption of the constancy of the mean service time would simply require that process of setting the UW employee requirements account for the variable, but presumably predicable, service time. Finally, relaxing the assumption of stationary underlying customer arrival-rate curves would necessitate effective forecasting of the

daily planning periods' customer arrival rates and, perhaps, increase the value of real-time control or tinkering with the desired level of service.

To summarize, CW only uses labor that is, on average, surplus to that needed to provide the desired level of customer service. Even relaxing all assumptions, idle time will still consistently occur at certain times. Thus, it is hard to envision how relaxing any of the simulation-related assumptions will jeopardize CW's improved labor utilization.

Managerial Issues Associated with the Use of Controllable Work

Flexibility Issues. In addressing the managerial implications of high and low contact services, Chase and Aquilano argued that

> Whenever possible a distinction should be made between the high-contact and low-contact elements of a service system. This can be done by a separation of functions, all high-contact activities should be performed by one group of people, all low-contact by another. This minimizes the influence of the customer on the production process and provides opportunities to achieve efficiency where it is actually possible. [10, p. 101]

We fully recognize the advantages of maintaining a technological core, but note that prohibiting high-contact employees (the front-line employees doing UW) from undertaking low-contact activities (CW) restricts the options available to managers. Clearly, there are advantages to a full-time staff—their knowledge of the service and their commitment to the organization are typically higher than for part-time employees. The results from the EXP2 show that managers may improve efficiency substantially by having front-line service delivery personnel carry out CW, even when limited shift scheduling flexibility- given by the type of shifts that full-time employees most commonly work-exists. Most importantly, the improvement in labor usage did not arise at the expense of a lower-than-desired level of customer service. The results of EXP2 showed that as the volume of CW increased, the amount of idle time decreased, but that the decrease in idle time was generally less than the increase in the CW labor hours. The implication of this finding is that to obtain the highest possible labor utilization, managers should combine the flexibility options at their disposal (e.g., using alternate break placements besides CW).

Controllable Work Block Length. Set-up time is a relevant managerial issue when considering the use of CW, primarily for its influence on the choice of an appropriate block length for the CW. Set-up time, which represents unproductive time, may include: (1) a wrap-up period

for the employee to leave the original task in a desired state; (2) transit time for tasks carried out at different locations; and (3) a warm-up period for the employee to get up to speed on the new task. In determining an appropriate block length for a particular type of CW, the characteristics of, including set-up times for, the work should be considered. For example: because of the lost time, a person must do at least *x* minutes of CW to make it beneficial to switch from UW to CW and later back to UW. One can evaluate the impact of set-ups by comparing idle time across different volumes of CW, and adjusting for set-up time. Consider, for example, the results reported in Table 6. With 10 hours of CW, average idle time was 143.64 hours. As CW increased to 20 hours, average idle time decreased to 136.51 hours. From the reduction in idle time a maximum set-up time of 7.13 hours is obtained, which translates to 71.3 percent of the CW hours scheduled. Thus, set-up time could represent up to 71.3 percent of the block duration and CW would still be beneficial in lowering idle time.

As managers of manufacturing operations can, service managers can work at lowering set-up times. In effect, the procedure presented here for assigning CW blocks to shifts serves to decrease set-up time, since it reduces the needed number of set-ups. Another way of reducing set-up time is using CW that can be done near the UW and by having CW and UW designed to be easy to stop and start. Having CW carried out near the UW does not necessarily mean within the view of customers, since this may result in higher-than-expected customer dissatisfaction.

Reducing set-up times may allow the use of smaller blocks of CW. The expected benefits of smaller blocks are twofold, as we have seen in EXP2: smaller blocks offer greater labor utilization and short blocks are less likely to be split between shifts during the assignment of CW to shifts. Rather than scheduling short-duration blocks, they may be used in real-time to improve labor usage at those times when customer demand is, due to natural variation, lighter than anticipated.

A Methodology for Improving the Utilization of Front-line Employees. It is useful to consider the steps a manager might take in applying the results of this research. An 11-step procedure is presented that may be helpful in focusing one's thoughts on the tasks involved in using CW to improve the utilization of front-line employees:

1. Develop a base on which to measure the value of CW: measure the level of customer service currently provided by the service delivery system and measure both scheduled and actual idle time of the front-line employees. Collect information on

customer arrival and service time distributions that later will be used in a simulation of the service system.

2. Identify CW tasks that are candidates for assignment to the front-line employees. These tasks should be of somewhat long duration (at least as long as the duration of planning periods used in the scheduling process), and there should be some latitude in when they can be undertaken. Also, management should be able to reduce back-shop labor appropriately if front-line employees performed a candidate task.

3. Identify the periods within which each candidate task must be done.

4. Identify the skills necessary for the front-line employees to execute each candidate task.

5. Develop a labor schedule with front-line employees doing both UW and CW.

6. Simulate the operation of the service delivery system with the new labor schedule and evaluate the performance of the simulated system.

7. Iteratively repeat steps 4 and 5, trying different combinations of CW tasks, to identify the best mix of CW and UW. If the front-line employees need new skills to do the CW tasks, determining the best combination of tasks may require weighing the improvement in labor utilization against training costs.

8. Train the front-line employees to execute the CW tasks identified as best.

9. Implement the new scheduling process, so that front-line employees do both UW and CW. Reduce the back-shop labor, by the amount of CW transferred to the front-line employees, through job reassignment, attrition, or layoffs.

10. Monitor the performance of the service delivery system.

11. Periodically evaluate the current performance of the service delivery system against its historical performance. Repeat steps 2 through 9 when it appears that doing so would be beneficial.

Step 7 is key to the effective use of CW. Given the nature of UW employee requirement patterns, all CW types are not likely to be of equal value in reducing employee idle time. Consequently, it is very important to identify the mix of CW resulting in the greatest reduction in employee idle time. Moreover, the potential necessity of training employees before they can carry out CW makes it very desirable to evaluate potential system improvement via the methodology of simulation. Decision scientists can make valuable contributions to service managers in all steps

of the process identified above, but particularly at steps 1, 5, 6, 7, and 9.

Research Extensions

Both Ml and M2 are readily extendable to the development of tour (weekly) work schedules. However, the difficulty of solving flexible tour (or, for that matter, shift) scheduling problems to optimality (see, from Table 1, that the largest optimally- solved labor scheduling problem had 970 shifts [6]) warrants the development of effective heuristics. In initial testing, a two-step heuristic functioned poorly relative to the optimal models. This heuristic's first step optimally scheduled shifts to cover only the UW employee requirements. The second step used both additional shifts and any surplus of staff from step I in optimally covering the CW requirements. The implication of the poor labor utilization provided by this heuristic compared to that provided by Ml and M2 is that heuristics should not employ a sequential approach to scheduling shifts and CW, but instead attempt to use the CW at all times during the scheduling of shifts. One possible approach for a started-from- scratch heuristic would be to assign CW to specific periods initially and then add shifts or tours to the schedule. Schedule improvement activities, such as moving or adjusting shifts and moving CW, could then be undertaken. Alternately, starting solutions may be provided to a construction/ improvement heuristic by solving the linear programming relaxation of Ml or M2.

Given the range of flexibilities incorporated in labor scheduling procedures and the steady advance of computer capabilities, perhaps the time has come to develop an integrative model incorporating the complete range of flexibility options that might be available to managers of service delivery systems. An integrative model would be useful not only in practice, but also for identifying synergistic effects between scheduling flexibility options. As noted earlier, these results have offered indirect evidence of the value of combining CW with other flexibility options.

Table 1: Information on relevant labor scheduling literature.

Reference	Scheduling Problem Shift (S), Tour (T)	Planning Period Duration (Minutes)	Number of Alternate Shifts/Tours	Solution Procedure[a]
Baker, Crabill, and Magazine [1]	S	NA[b]	NA[b]	CFS
Bailey [2]	T	60	35840	LP
Bailey and Field [3]	S	30	72	LP
Bartholdi [4]	S	NA[b]	NA[b]	LP
Bechtold, Brusco, and Showalter [5]	T	60	168	LP,H[c]
Bechtold and Jacobs [6]	S	30	970	IP
Bechtold and Jacobs [7]	S	30	444	IP
Bechtold and Showalter [8]	T	60	56	H
Buffa, Cosgrove, and Luce [9]	S	30	500	H
Henderson and Berry [11]	S	15	100	LP
Henderson and Berry [12]	S	15	100	LP[d]
Holloran and Byrn [13]	T	30	5000	LP
Keith [14]	S	15	NA[e]	LP
Li, Robinson, and Mabert [15]	T	120	132	LP
Loucks and Jacobs [16]	T	60	81[f]	H
Mabert and Watts [17]	T	60	30	LP
McGinnis, Culver, and Deane [18]	T	60	504	H
Moondra [19]	S	60	26	LP
Morris and Showalter [20]	T	60	168	LP
Segal [22]	S	30	300	LP
Showalter and Mabert [23]	T	60	320	LP
Thompson [24]	S	15	6588	LP

[a]CFS=closed form solution, LP=linear programming based heuristic, H=heuristic (but not LP-based), IP=integer programming.

[b]These models were for cyclical scheduling problems, and applied for any planning period duration providing no breaks were scheduled for any shift.

[c]Bechtold et al. [5] evaluated published tour scheduling heuristics, some of which were LP-based.

[d]Henderson and Berry [12] developed an LP-based optimal branch and bound procedure.

[e]Not provided.

[f]Number of shifts. The number of acceptable tours was not identified.

Table 2: The factors and levels of the factors used in Experiment 1.

Factor	Factor Abbreviation (Number of Levels)	Unit of Measure	Levels
Uncontrollable work employee requirements pattern	RP (4)	A pattern	Unimodal, Bimodal, Trimodal, and Random (see Figure 1)
Variation in the uncontrollable work employee requirements	RV (2)	A coefficient of variation	.2 and .6
Mean requirement for employees arising from the uncontrollable work	MR (2)	Number of employees needed each period, on average	5 and 20
Number of types of controllable work	WT (3)	A number	1, 2, and 3
Volume of controllable work	LH (3)	The number of employee-hours required to do the controllable work	10, 20, and 40
The degree of flexibility in the timing of the controllable work	TF (3)	The average number of periods in the commencement window for each block of controllable work	Low (2), Medium (4), and High (8)
Controllable work block length	BL (2)	Hours	1 and 2
Model used in scheduling shifts and controllable work	MD (2)	An optimal model	M1 and M2

Table 3: Selected terms from an ANOVA model of the results for Experiment 1 using schedule generation time as the dependent variable.

Source	DF	Sum of Squares	Mean Square
Model	431	235615.49	546.672
Error	1296	17788.22	13.726
Corrected Total	1727	253403.71	
F Value = 39.83,		Pr > F = .0000,	R^2 = .929803

Source	DF	ANOVA SS	Mean Square	F Value	Pr > F
MD	1	15851.21	15851.21	1154.87	.0001
TF*MD	2	19876.82	9938.41	724.09	.0001
LH*MD	2	6655.70	3327.85	242.46	.0001
BL*MD	1	2811.96	2811.96	204.87	.0001
WT*MD	2	1163.80	581.90	42.40	.0001

Notes: Only MD and MD-based first-order interaction terms significant at a level greater than .01. Schedule generation time is the total time required to generate optimal schedules on a 80486-based personal computer. MD=shift and controllable work scheduling model; TF=controllable work timing flexibility; LH=labor hours required to do the controllable work; BL=controllable work block length; WT=number of controllable work types. (RP—the uncontrollable work employee requirements pattern—was not included as a factor in the ANOVA model. The complete ANOVA results are available upon request.)

Table 4: The factors and levels of the factors used in Experiment 2.

Factor	Factor Abbreviation (Number of levels)	Unit of Measure	Levels
Underlying (true) customer arrival-rate pattern	AP (3)	A sinusoidal pattern	Unimodal, Bimodal, and Trimodal (see Figure 3)
Variation in the true customer arrival-rate	AV (2)	A coefficient of variation	.25 and .50
Accuracy of the forecasts of customer arrivals	FA (3)	Number of historical days used in calculating average customer arrivals for each daily period	1, 4, and 16
Mean duration of customer service	SD (3)	Minutes	1, 4, and 16
Desired level of customer service	SL (3)	The maximum waiting time within which 90 percent of the customers are to be served	3, 1, and 1/3 minutes

Notes: Experiment 2 also uses three controllable work-based factors used in Experiment 1: LH—the labor-hours of controllable work; TF—the controllable work timing flexibility; and BL—the controllable work block length.

Table 5: Selected terms from an ANOVA model of the results for Experiment 2 using the actual customer service level (PM1) as the dependent variable.

Source	DF	Sum of Squares	Mean Square
Model	485	18476.12	38.095
Error	5346	7511.81	1.405
Corrected Total	5831	25988.00	
F Value = 27.11,		Pr > F = .0000,	R^2 = .710951

Source	DF	ANOVA SS	Mean Square	F Value	Pr > F
LH	2	1155.58	577.79	411.20	.0001
SD*LH	4	98.00	24.50	17.44	.0001
AP*SD*LH	8	90.52	11.31	8.05	.0001
AP*LH	4	33.35	8.34	5.93	.0001
AP*AV*LH	4	22.85	5.71	4.07	.0027

Notes: Only LH-based terms significant at a level greater than .01. LH=the hours of controllable work to be scheduled; AP=the customer arrival-rate pattern; AV=the variation in the customer arrival-rate pattern; and SD=the mean service duration. (The following factors were not included in the ANOVA model: TF, the controllable work timing flexibility, and BL, the controllable work block length. The complete ANOVA results are available upon request.)

Table 6: Performance by the levels of selected factors from Experiment 2.

Performance Measure	Factor	Factor Levels		
		Low	Medium	High
PM1	LH	96.53	96.03	95.44
PM2	LH	143.64	136.51	125.51
PM2	TF	139.28	133.41	132.97
PM2	BL	133.75	NA	136.69
PM3	LH	51.14	54.02	58.97
PM3	TF	54.03	55.10	55.18
PM3	BL	55.03	NA	54.49

Notes: PM1=percentage of customers served within the specified waiting time limit; PM2=actual idle time of the front-line employees, in hours; PM3=actual utilization of the front-line employees, in percent; LH=hours of controllable work to be scheduled; TF=controllable work timing flexibility; BL=controllable work block length; NA=not applicable (BL has only two levels).

Table 7: Selected terms from an ANOVA model of the results for Experiment 2 using actual idle hours of the front-line employees (PM2) as the dependent variable.*

Source	DF	Sum of Squares	Mean Square
Model	323	54948989.91	170120.71
Error	5508	11135456.44	2021.69
Corrected Total	5831	66084446.34	
F Value = 84.15,		Pr > F = .0000,	R^2 = .831497

Source	DF	ANOVA SS	Mean Square	F Value	Pr > F
LH	2	324613.60	162306.80	80.28	.0001
TF	2	48209.12	24104.56	11.92	.0001
SD*LH	4	64267.34	16066.83	7.95	.0001
BL	1	12669.96	12669.96	6.27	.0123
BL*TF	2	15020.71	7510.36	3.71	.0244
TF*LH	4	21692.88	5423.22	2.68	.0299

*Only controllable work-based terms significant at a level greater than .05.

Notes: LH=hours of controllable work to be scheduled; TF=controllable work timing flexibility; SD=mean service duration; and BL=controllable work block length. (FA, the accuracy of customer arrival forecasts, and SL, the desired level of customer service, were not included as factors in the ANOVA model. The complete ANOVA results are available upon request.)

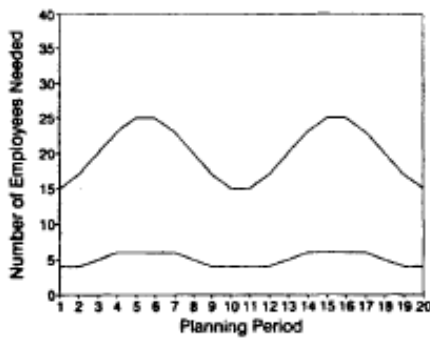**Figure 1:** Uncontrollable-work employee requirement curves used in Experiment 1.

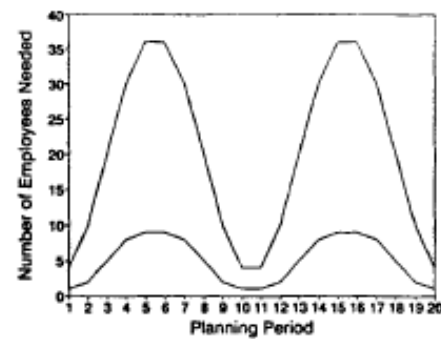**a.** The two low-variation unimodal employee requirements patterns.
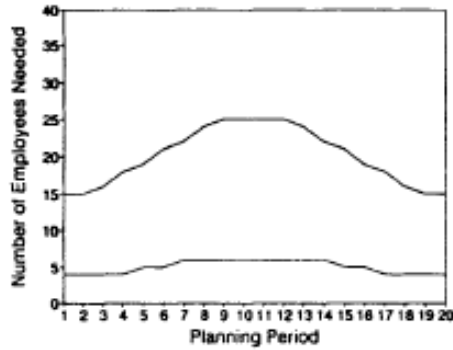
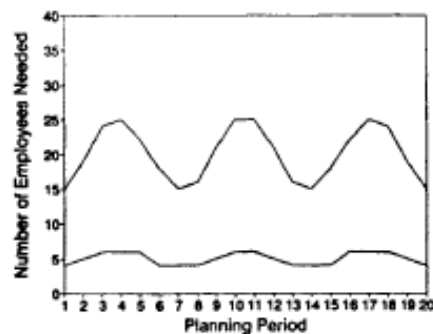**b.** The two high-variation unimodal employee requirements patterns.
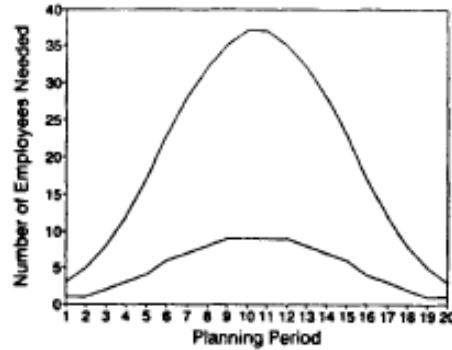
c. The two low-variation bimodal employee requirements patterns.
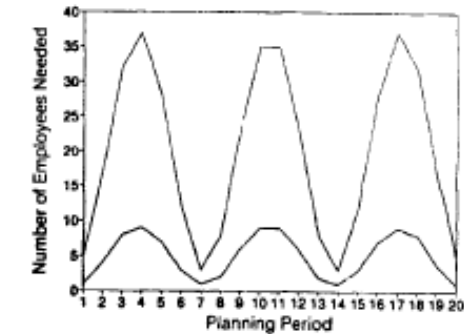

d. The two high-variation bimodal employee requirements patterns.

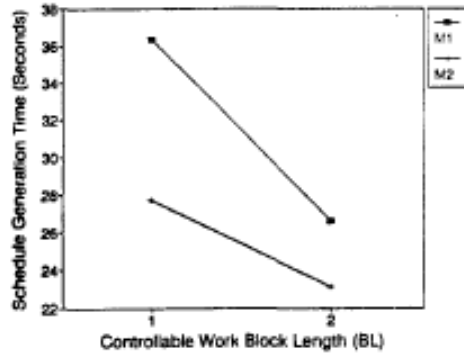**e.** The two low-variation trimodal employee requirements patterns.

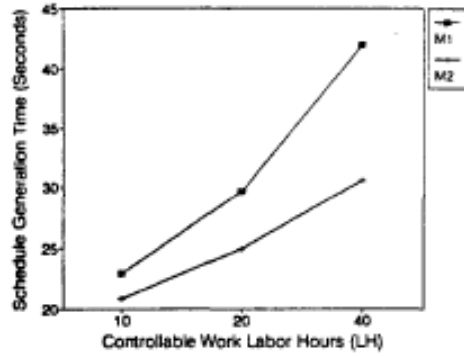**f.** The two high-variation trimodal employee requirements patterns.

Figure 2: MD-based first-order interactions for schedule generation times in Experiment 1 (MD=shift and controllable work scheduling model).
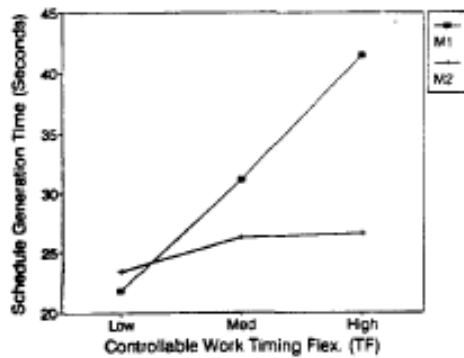
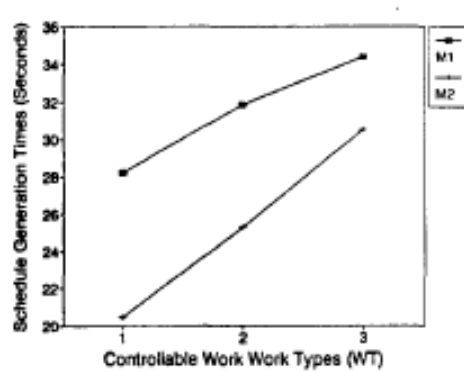**a.** Schedule generation time, by MD, as a function of BL.

b. Schedule generation time, by MD, as a function of LH.

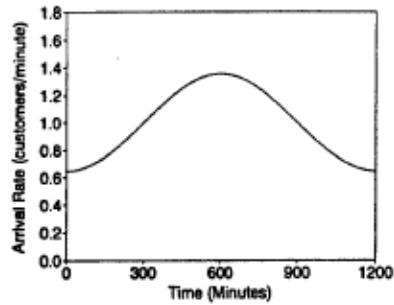**c.** Schedule generation time, by MD, as a function of TF.

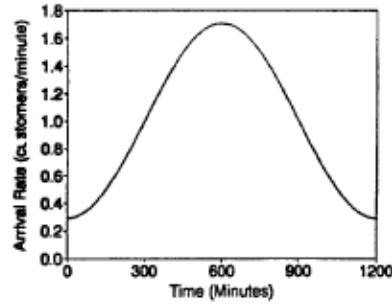**d.** Schedule generation time, by MD, as a function of WT.

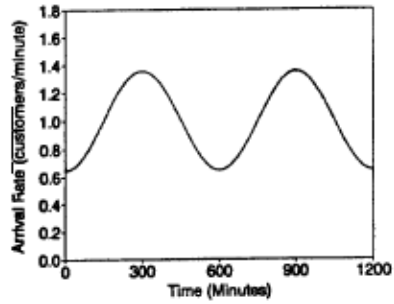**Figure 3:** Customer arrival-rate curves for Experiment 2.

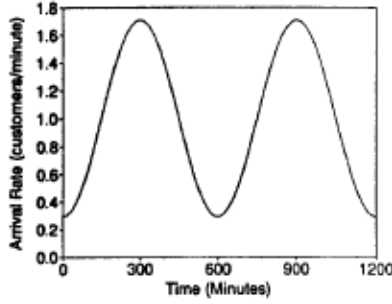a. The low-variation unimodal customer arrival-rate pattern.

b. The high-variation unimodal customer arrival-rate pattern.
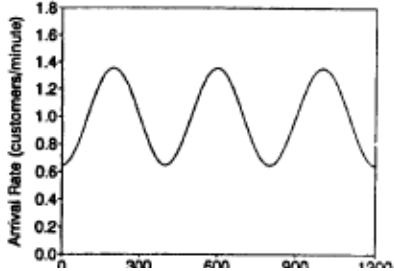
c. The low-variation bimodal customer arrival-rate pattern.

d. The high-variation bimodal customer arrival-rate pattern.

e. The low-variation trimodal customer arrival-rate pattern.

f. The high-variation trimodal customer arrival-rate pattern.
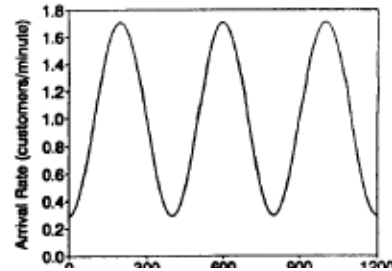
g. The two low-variation random employee requirements patterns.
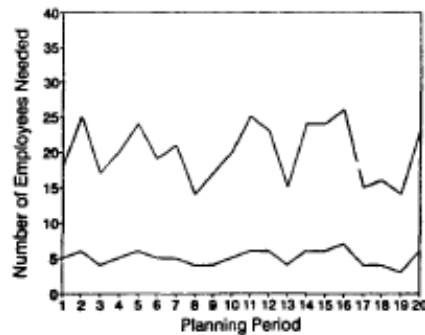
h. The two high-variation random employee requirements patterns.
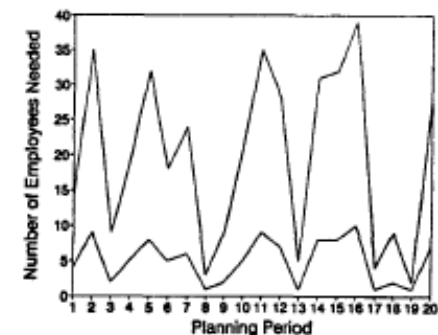
Figure 4: LH-based first- and second-order interactions for the actual level of customer service in Experiment 2.

a. Actual customer service, by controllable work labor hours, as a function of the customer arrival-rate pattern.
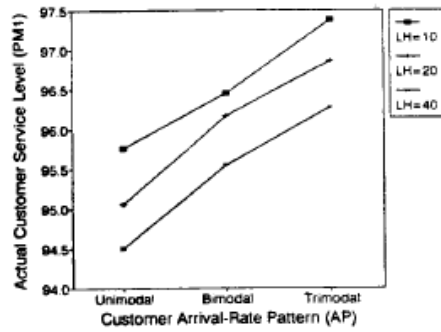
b. Actual customer service, by controllable work labor hours, as a function of mean service duration.
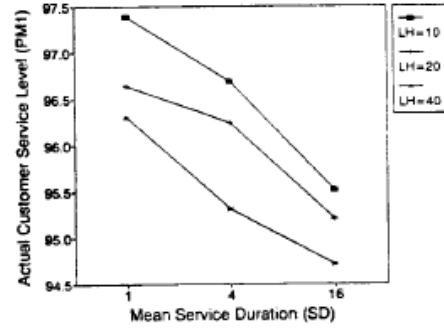
Figure 5: Controllable work-based first-order interactions for actual idle time of the front-line employees in Experiment 2 (BL=controllable work block length, in hours; LH=labor hours required to do the controllable work).

a. Actual idle hours, by BL, as a function of TF.
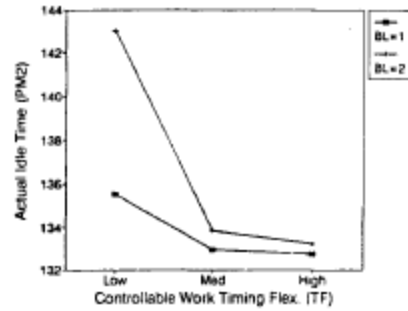
b. Actual idle hours, by LH, as a function of TF.
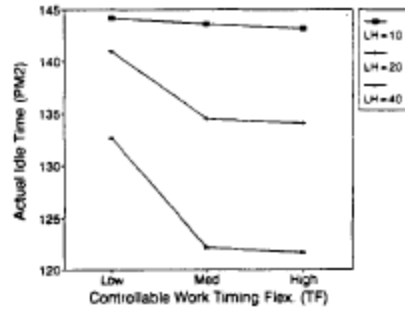
c. Actual idle hours, by LH, as a function of SD.
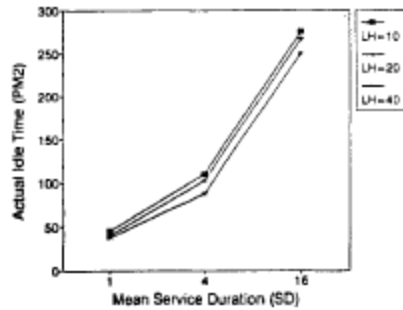
REFERENCES

[1] Baker, K. R., Crabill, T. B., & Magazine, M. J. An optimal procedure for allocating manpower with cyclic requirements. *AIIE Transactions,* 1973, 5(2), 119-126.

[2] Bailey, J. Integrated days off and shift personnel scheduling. *Computers and Industrial Engineering,* 1985, 9(4), 395-404.

[3] Bailey, J., & Field, J. Personnel scheduling with flexshift models. *Journal of Operations Management,* 1985, 5(3), 327-338.

[4] Bartholdi, J. J. A guaranteed-accuracy round-off algorithm for cyclic scheduling and set covering *Operations Research,* 1981, 29(3), 501-510.

[5] Bechtold, S. E., Brusco, M. J., & Showalter, M. J. A comparative evaluation of labor tour scheduling methods. *Decision Sciences,* 1991, 22(4), 683-699.

[6] Bechtold, S. E., & Jacobs, L. W. Implicit optimal modeling of flexible break assignments in labor staffing decisions for service operations. *Management Science,* 1990, 56(11), 1339-1351.

[7] Bechtold, S. E., & Jacobs, L. W. Improvement of labor utilization in shift scheduling for services with implicit optimal modeling. *International Journal of Operations and Production Management,* 1991, *11(2),* 54-69.

[8] Bechtold, S. E., & Showalter, M. J. A methodology for labor scheduling in a service operating system. *Decision Sciences,* 1987, 78(1), 89-107.

[9] Buffa, E. S., Cosgrove, M. J., & Luce, B. J. An integrated work shift scheduling system. *Decision Sciences,* 1976, 7, 620-630.

[10] Chase, R. B., & Aquilano, N. J. *Production and operations management: A life cycle approach* (5th ed.). Homewood, IL: Irwin, 1989.

[11] Henderson, W. B., & Berry, W. L. Heuristic methods for telephone operator shift scheduling: An experimental analysis. *Management Science,* 1976, 22(12), 1372-1380.

[12] Henderson, W. B., & Berry, W. L. Determining optimal shift schedules for telephone traffic exchange operators. *Decision Sciences*, 1977, 8(2), 239-255.

[13] Holloran, T. J., & Bym, J. E. United Airlines station manpower planning system. *Interfaces,* 1986, 76(1), 39-50.

[14] Keith, E. G. Operator scheduling. *AllE Transactions,* 1979, 7/(1), 37-41.

[15] Li, C., Robinson, E. P., Jr, & Mabert, V. A. An evaluation of tour scheduling heuristics with differences in employee productivity and cost. *Decision Sciences,* 1991, 22(4), 700-718.

[16] Loucks, J. S., & Jacobs, F. R. Tour scheduling and task assignment of a heterogeneous work force: A heuristic approach. *Decision Sciences,* 1991, 22(4), 719-738.

[17] Mabert, V. A., & Watts, C. A. A simulation analysis of tour-shift construction procedures. *Management Science,* 1982, 28(5), 520-532.

[18] McGinnis, L. F., Culver, W. D., & Deane, R. H. One- and two-phase heuristics for workforce cftwtiiling. *Computers and Industrial Engineering,* 1978, 2(1), 7-15.

[19] Moondra, S. L. An L. P. model for work force scheduling for banks. *Journal of Bank Research,* 1976, 7(4), 299-301.

[20] Morris, J. G., & Showalter, M. J. Simple approaches to shift, days-off and tour scheduling problems. *Management Science,* 1983, 29(8), 942-50.

[21] *SAS users guide: Statistics, Version 5.* Cary, NC: SAS Institute, Inc., 1985.

[22] Segal, M. The operator-scheduling problem: A network-flow approach. *Operations Research,* 1974, 22(4), 808-823.

[23] Showalter, M. J., & Mabert, V. A. An evaluation of a full-/part-time tour scheduling methodology. *International Journal of Operations and Production Management,* 1989, 8(7), 54-71.

[24] Thompson, G. M. Shift scheduling when employees have limited availability: An L. P. approach. *Journal of Operations Management,* 1990, 9(3), 352-370.

APPENDIX

A 14-hour day, comprised of hour-long planning periods shall be used for the sample scheduling environment. Shifts are nine hours long with an hour-long break taken in the fifth hour and all are of equal cost. UW employee requirements for periods 1 through 14 are 1, 2, 2, 3, 3, 3, 3, 4, 3, 3, 2, 2, 1, and 1 employee(s), respectively. Table A1 gives information on the two available types of CW.

Ml for this problem is:

$$\text{Min } x_1 + x_2 + x_3 + x_4 + x_5 + x_6$$

subject to

$$
\begin{array}{llllll}
x_1 & & & & \geq 1 & \quad 1 \\
x_1 + x_2 & -c_{1,02} & & & \geq 2 & \quad 2 \\
x_1 + x_2 + x_3 & -c_{1,03}\ \ -c_{1,02} & & & \geq 2 & \quad 3 \\
x_1 + x_2 + x_3 + x_4 & -c_{1,04}\ \ -c_{1,03}\ \ -c_{2,04} & & & \geq 3 & \quad 4 \\
x_2 + x_3 + x_4 + x_5 & -c_{1,04}\ \ -c_{2,05}\ \ -c_{2,04}\ \ -c_{4,05} & & & \geq 3 & \quad 5 \\
x_1 + x_3 + x_4 + x_5 + x_6 & -c_{2,06}\ \ -c_{2,05}\ \ -c_{4,06} & & & \geq 3 & \quad 6 \\
x_1 + x_2 + x_4 + x_5 + x_6 & -c_{2,07}\ \ -c_{2,06}\ \ -c_{4,07}\ \ -c_{5,07} & & & \geq 3 & \quad 7 \\
x_1 + x_2 + x_3 + x_5 + x_6 & -c_{2,08}\ \ -c_{2,07}\ \ -c_{4,08}\ \ -c_{5,08} & & & \geq 4 & \quad 8 \\
x_1 + x_2 + x_3 + x_4 + x_6 & -c_{2,09}\ \ -c_{2,08}\ \ -c_{4,09}\ \ -c_{5,09} & & & \geq 3 & \quad 9 \\
x_2 + x_3 + x_4 + x_5 & -c_{3,10}\ \ -c_{2,10}\ \ -c_{2,09}\ \ -c_{5,10} & & & \geq 3 & \quad 10 \\
x_3 + x_4 + x_5 + x_6 & -c_{3,11}\ \ -c_{3,10}\ \ -c_{2,11}\ \ -c_{2,10}\ \ -c_{5,11} & & & \geq 2 & \quad 11 \\
x_4 + x_5 + x_6 & -c_{3,12}\ \ -c_{3,11}\ \ -c_{2,11}\ \ -c_{5,12} & & & \geq 2 & \quad 12 \\
x_5 + x_6 & -c_{3,13}\ \ -c_{3,12} & & & \geq 1 & \quad 13 \\
x_6 & -c_{3,13} & & & \geq 1 & \quad 14 \\
\end{array}
$$

(with Hr column shown at right)

$$
\begin{array}{lll}
c_{1,02} + c_{1,03} + c_{1,04} & = 2 & \text{(Blocks 1,6)} \\
c_{2,04} + c_{2,05} + c_{2,06} + c_{2,07} + c_{2,08} + c_{2,09} + c_{2,10} + c_{2,11} & = 1 & \text{(Block 2)} \\
c_{3,10} + c_{3,11} + c_{3,12} + c_{3,13} & = 1 & \text{(Block 3)} \\
c_{4,05} + c_{4,06} + c_{4,07} + c_{4,08} + c_{4,09} & = 1 & \text{(Block 4)} \\
c_{5,07} + c_{5,08} + c_{5,09} + c_{5,10} + c_{5,11} + c_{5,12} & = 1 & \text{(Block 5)} \\
\end{array}
$$

all variables ≥ 0 and integer

One optimal solution to this problem is $x_2 = x_4 = 2$, $x_1 = x_6 = 1$, and $c_{1,02} = c_{1,04} = c_{2,06} = c_{3,11} = c_{4,07} = c_{5,07} = 1$.

Table A1: Information on the controllable work to be scheduled in the sample problem.

Block Number	Controllable Work Type	Block Length	Earliest Desirable Starting Period	Latest Desirable Starting Period
1	1	2	2	4
2	1	2	4	11
3	1	2	10	13
4	2	1	5	9
5	2	1	7	12
6	1	2	2	4

With equation references given in parentheses, M2 for this problem is:

$$\text{Min } x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \qquad (1)$$

subject to

x_1					$\geq 1 \quad (6)$
$x_1 + x_2$		$-s_{1,02}$			≥ 2
$x_1 + x_2 + x_3$		$-s_{1,03}$	$-s_{1,02}$		≥ 2
$x_1 + x_2 + x_3 + x_4$		$-s_{1,04}$	$-s_{1,03}$		≥ 3
$x_2 + x_3 + x_4 + x_5$		$-s_{1,05}$	$-s_{1,04}$	$-s_{2,05}$	≥ 3
$x_1 \quad + x_3 + x_4 + x_5 + x_6$		$-s_{1,06}$	$-s_{1,05}$	$-s_{2,06}$	≥ 3
$x_1 + x_2 \quad + x_4 + x_5 + x_6$		$-s_{1,07}$	$-s_{1,06}$	$-s_{2,07}$	≥ 3
$x_1 + x_2 + x_3 \quad + x_5 + x_6$		$-s_{1,08}$	$-s_{1,07}$	$-s_{2,08}$	≥ 4
$x_1 + x_2 + x_3 + x_4 \quad + x_6$		$-s_{1,09}$	$-s_{1,08}$	$-s_{2,09}$	≥ 3
$x_2 + x_3 + x_4 + x_5$		$-s_{1,10}$	$-s_{1,09}$	$-s_{2,10}$	≥ 3
$x_3 + x_4 + x_5 + x_6$		$-s_{1,11}$	$-s_{1,10}$	$-s_{2,11}$	≥ 2
$x_4 + x_5 + x_6$		$-s_{1,12}$	$-s_{1,11}$	$-s_{2,12}$	≥ 2
$x_5 + x_6$		$-s_{1,13}$	$-s_{1,12}$		≥ 1
x_6			$-s_{1,13}$		≥ 1

$$
\begin{aligned}
s_{1,02} + s_{1,03} + s_{1,04} + s_{1,05} + s_{1,06} & \\
+ s_{1,07} + s_{1,08} + s_{1,09} + s_{1,10} + s_{1,11} + s_{1,12} + s_{1,13} &= 4 \qquad (7)\\
s_{2,05} + s_{2,06} + s_{2,07} + s_{2,08} + s_{2,09} + s_{2,10} + s_{2,11} + s_{2,12} &= 2
\end{aligned}
$$

$$
\begin{aligned}
s_{1,02} + s_{1,03} & \geq 2 \qquad (8)\\
s_{1,02} + s_{1,03} + s_{1,04} + s_{1,05} + s_{1,06} + s_{1,07} + s_{1,08} + s_{1,09} & \geq 3\\
s_{2,05} + s_{2,06} & \geq 1
\end{aligned}
$$

$$
\begin{aligned}
s_{1,12} + s_{1,13} & \geq 1 \qquad (9)\\
s_{1,05} + s_{1,06} + s_{1,07} + s_{1,08} + s_{1,09} + s_{1,10} + s_{1,11} + s_{1,12} + s_{1,13} & \geq 2\\
s_{2,10} + s_{2,11} + s_{2,12} & \geq 1
\end{aligned}
$$

$$\text{all variables 0 and integer} \qquad (4),(10)$$

One optimal solution to this problem is $x_2 = 2, x_1 = x_4 = x_5 = 1$, and $x_6 = s_{1,02} = s_{1,04} = s_{1,09} = s_{1,11} = s_{2,08} = s_{2,09} = 1$.

Table A2 compares the solutions to the above formulations to the solution of the same problem without CW. Although another shift is necessary when the front-line employees also do CW, scheduled idle time for the front-line employees fell from seven to five hours. This results in the scheduled utilization of the front-line employees increasing from 82.5 percent to 89.6 percent. Assuming that the organization saves ten paid, back-shop hours with the transfer of ten hours of CW to the front-line employees, the net savings to the organization, of having the front-line employees perform ten hours of CW, is two paid hours (ten hour reduction in back-shop hours, less the eight hour increase in front-line labor hours).

Table A2: The effects of scheduling controllable work in the sample problem.

	No Controllable Work To Be Scheduled	Ten Hours of Controllable Work To Be Scheduled*
Number of shifts scheduled	5	6
Paid front-line labor hours	40	48
Uncontrollable work labor hours	33	33
Controllable work labor hours	0	10
Scheduled idle hours**	7	5
Scheduled utilization**	82.50	89.58
Reduction in labor hours for the service system***	0	2

*Based on the timing restrictions identified in Table A1.

**Actual idle times will be higher and actual labor utilization will be lower than presented because of the employee idle time inherent in delivering the desired level of customer service.

***Represents the benefit to the service system of having the front-line employees perform ten hours of CW. Assumes that the ten hours of CW transferred from the back-shop to the front-line employees save ten paid hours of back-shop labor.

Figure A1: Solutions to M1 and M2 for the sample problem.

Notes: uuu=an employee available to perform uncontrollable work; ccc=an employee performing controllable work; bbb=an employee on a meal break; TAE=total net number of employees available to perform the uncontrollable work; and SIT=scheduled employee idle time, in employee-periods.

Figure A1 illustrates possible work schedules for the sample problem for each model. With either solution there is a good degree of latitude in assigning the CW blocks to shifts, but the illustrated matchings came from applying the assignment procedure presented earlier. Note that although the solutions are different, both are optimal.