Macro- and Micro-Modeling of Field Service Dynamics

Jack Homer
Homer Consulting
Voorhees, New Jersey 08043
JackBHomer@compuserve.com

Abstract

Manufacturers of high-tech capital equipment typically offer their customers after-sales support in the form of repairs and preventive maintenance, as well as bug fixes and upgrades. These services require a field engineering workforce which tends to grow with the installed base of customer-owned equipment, but whose growth is also affected by other variables, such as equipment reliability and workforce experience. The company's ability to match field service capacity to demand affects customer satisfaction, and thus, future demand for its products and services.

A system dynamics model to investigate field service issues was developed for a major producer of diagnostic equipment for semiconductor manufacturing. This strategic model is broad in scope and treats variables in an aggregate and deterministic way that is typical for such models. This approach is adequate in most respects, but begs the key question of how service readiness (the adequacy of the workforce to handle service job volume and variety) is affected by the cross-training of workers on multiple product types. As a result, it proved useful to supplement the strategic "macro" model with a "micro" model that focuses on the daily queueing and assignment of service jobs. The micro model specifies the product-by-product skill levels of individual field engineers in a local service area and includes an optimization algorithm for job assignment. The micro model provides accurate information on expected service performance as a function of workload and extent of cross-training. This information is used for calibrating the strategic model and may also be used for making tactical manpower decisions at the local level.

Background

A system dynamics modeling project was undertaken by a major producer of diagnostic machines used in semiconductor wafer fabrication. The company's products range from simpler machines requiring little maintenance to complex systems requiring more frequent maintenance. All equipment comes with a one-year warranty on parts and service, after which the customer has the option of continuing with a service contract. Customers often select service contracts for the more complex and essential pieces of equipment, for which downtime must be minimized and do-it-yourself repairs are difficult at best.

The company's new product sales have grown rapidly over the last several years, though they do follow the ups and downs of the semiconductor industry's persistent two-year business cycle. This overall growth in sales has led to robust growth in the installed base of equipment and similar expansion of the workforce of field service engineers. Although field service does not generate much for the company in the way of profits, it is nearly as important as product performance and fair price are for the company's continued success in the marketplace. As the field service workforce has grown, its planning, organization, and management have taken on increasing complexity and significance for the company. (Richmond 1994 describes the inevitable evolution of a high-tech manufacturer from a primary sales focus to an increasing field service focus.) For example, while the workforce has always been segmented by local territory or hub, it has only recently been segmented by customer account as well. This action was taken so that major customers would generally have their service done by engineers with whom they are familiar and who have worked on their equipment before.

With the field service organization at a critical juncture, senior management turned to system dynamics as a tool for the analysis of strategies that seek to trim workforce costs without jeopardizing customer satisfaction. Some managers felt that the field workforce had grown too large relative to the installed base, as supported by anecdotal evidence of excess idle time. Others claimed that the idle time was primarily a reflection of workload imbalance due to inadequate product cross-training, and pointed to the fact that some field engineers worked long hours with no idle time while other engineers often had hours or days between service jobs. These managers felt that additional cross-training should be implemented before considering a hiring freeze or similar drastic action.

The modeling work was done with a team that included both senior staff and field managers. As often happens, our discussions tended to move back and forth from aggregate concepts, such as customer satisfaction or workforce utilization, to more detailed ones, such as distinctions among various types of products, customers, and service contracts. The team agreed that, for purposes of strategic analysis, the appropriate model would be a "macro" one with a time horizon of five to ten years that aggregated across most of the detailed distinctions. But the team also wanted to look closely at how a service job backlog can develop, along with idle time, when workforce skills do not match up well with service requirements. We required a separate "micro" model to investigate this question, a stochastic job-queueing model that focused at the local hub level and represented product types and field engineers individually, and that had the ability to assign engineers to product-specific jobs in a realistic and effective way.

The remainder of this paper presents an overview of the macro and micro models. It will be shown how the micro model has played a key role in supporting the macro model, particularly in helping to specify the nonlinear function through which cross-training impacts the workforce's ability to handle service volume and variety.

A Strategic Model of Field Service

The idea of modeling service delivery is hardly new to system dynamics, though the number of published works is still modest. Much of the early work was in the area of health care and education (Levin et al. 1976). Later works of note include models of People Express Airlines (Graham et al. 1992), Hanover Insurance claims processing (Senge and Sterman 1992), NatWest Bank lending (Oliva 1996), and DuPont chemical plant equipment maintenance (Carroll, Sterman and Marcus 1998). These models have led to useful recommendations; for example, that a service company should hire steadily rather than in spurts to avoid problems of inexperience, should hire enough workers to avoid overwork and a drift to low standards, and (in the case of equipment service) should give preventive maintenance high priority to avoid a spiral of equipment failures.

The field service strategic model, about 500 equations in all, contains sectors dealing with the installed base, service job volume, workforce and its utilization, service readiness and performance, customer satisfaction, and financial results. The installed base is disaggregated only by service billing category: machines under warranty, machines under service contract, and machines billed by time and material. These categories are important not only for financial reasons, but also because they differ in terms of typical service requirements. For example, machines billed by time and material typically generate far fewer repair calls than do machines under contract or under warranty.

Improvements in technology, associated with better product reliability or serviceability, enter the installed base through newly installed machines and upgrades of older machines. New machines are also associated with a number of product bugs that require time to discover, solve, and eliminate through engineering change order (ECO) jobs. Machine reliability, as measured by mean time between failure (MTBF), is determined jointly by the level of technology, the number of bugs outstanding, and the frequency of preventive maintenance (PM) service. Failures generate repair jobs, which, being unscheduled (unlike ECO, PM, and upgrade jobs) yet requiring immediate attention, are at the heart of the workforce planning puzzle that confronts a company servicing a variety of products.

Customer satisfaction with service is a function of two aspects of performance: machine downtime and service frequency (greater being worse in both cases.) These metrics are directly related: Downtime = Service frequency x Downtime per service. In the case of repair jobs, downtime per service represents the sum of call response time and repair time. Call response time is measured from when the customer first calls the central response center until an assigned field engineer arrives at the customer site. Response time may be extended beyond the minimum preparation and travel time when needed repair parts are not locally available and have to be shipped overnight, or when no engineer with appropriate product training is available during the current work shift. Repair time is determined by product serviceability, the skill level of the assigned engineer, level of fatigue, and the ability of Technical Support to assist field engineers (especially less experienced ones) over the phone.

Since the strategy model aggregates across all product types, a literal matching of product-specific skills to service jobs cannot be done as it is in the job-queueing model. Instead, the strategy model describes the adequacy of workforce skills using the concept of service readiness. A low level of readiness implies an imbalance that leads to additional overtime and extends call response times. If all engineers were trained on all product types, then service readiness could simply be defined as the ratio of repair hours available to repair hours required. (Repair hours available per engineer equals 40x52 baseline hours per year less time away for holidays, personal days off, training, and installation and upgrade job assignments.) But if, practically speaking, each engineer can become trained and skilled on only a handful of products at most, service readiness will be reduced to some extent below what the aggregate ratio of available to required hours alone would suggest.

It is important to know not only the size of the field workforce and its level of cross-training, but also the mean skill level that is brought to bear on repair jobs. The applied skill level affects required repair time, and thus, machine downtime. To compute the mean skill level applied, one must first understand how skill is gained over time on the product for which an engineer is initially trained (primary skill), as well as on any products for which the engineers has been cross-trained (secondary skills). This learning process is represented in the macro model through the use of primary and secondary skill chains, as shown in Figure 1. New hires receive their primary training and are considered rookies until they become certified. Once certified, a worker is eligible for cross-training on one or more secondary products. Secondary skills involve the same processes of training, certification, and on-the-job learning as primary skills. Cross-trained workers must divide their time among the two or more products on which they are skilled. Consequently, it generally takes longer to climb the secondary skill chain than the primary skill chain. When people leave the workforce, they take away with them all of the skills they have acquired.

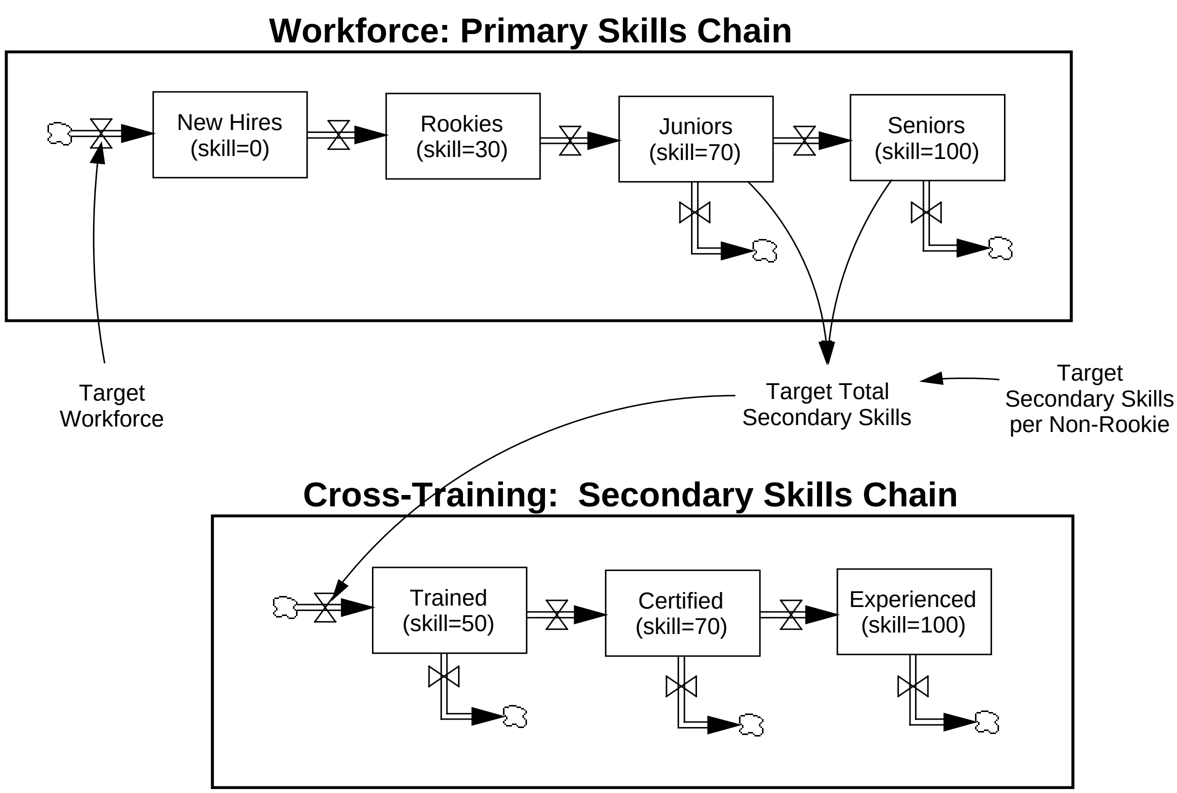

Figure 1. Field service workforce skill chains

The mean skill applied to repair jobs is a weighted average of the skill of rookies (ex. 30/100) and the mean applied skill of non-rookies (juniors and seniors). The mean applied skill of non-rookies, in turn, is a weighted average of their mean primary skill (ex. 90/100) and their mean secondary skill (ex. 80/100).[1] An increased level of cross-training (more secondary skills per non-rookie) will pull the mean applied skill of non-rookies downward in the direction of the mean secondary skill (ex. from 90/100 toward 80/100). This negative consequence of cross-training is referred to as dilution of skills. Another negative consequence is that each new cross-training takes a valuable non-rookie out of the workforce for a period of several weeks, thereby reducing the service hours available during that time.

Hiring is driven by the target size of the field workforce (see Figure 1), which, in turn, grows with the installed base according to the equation:

Target workforce = Projected installed base / Target machines per worker.

The projected installed base is just the sum of the current installed base and machines on order. The target machines per worker is a number (which in real life is set separately for each product type) that is subject to annual adjustment, based on recent workforce utilization. If utilization (service hours worked divided by total baseline hours) has been higher than its target value, target machines per worker will be reduced; if utilization has been lower than its target value, target machines per worker will be increased. Because of its central role in workforce sizing, the utilization target is an ongoing subject of discussion in the company and is closely related to the debate over idle time.

Feedback Structure and Behavior of the Strategic Model

Figure 2 is a causal-loop diagram that brings together many of the strategic model relationships described above. Loop B1 is a balancing loop that can generally do a fairly good job of adjusting the field workforce to meet service demand (hours required), based on recent utilization. Service demand grows with the installed base, but is also affected by contract penetration (the fraction of out-of-warranty machines under contract), product reliability and serviceability, technical support, the mean skill level applied, and fatigue. Because these factors may change, so too may utilization and, thus, the target number of machines per worker. Such change, in turn, may cause the workforce to grow at a rate or in a manner different from that of the installed base.

Even under relatively benign conditions, reinforcing Loop R1, involving variations in workforce skill level, can complicate the balancing act of Loop B1 and induce some fluctuations in workforce utilization and hiring. A sudden change in the rate of workforce growth can lead to an imbalance in the primary skills chain (too many or too few rookies) that, in turn, causes a hiring seesaw effect for several years thereafter.

Figure 3 shows a run in which the workforce is initially flat for two years before it starts to grow at about the same rate as the installed base. (The variables are graphed on different scales, but with a minimum value of zero for workforce, installed base, and machines per worker.) The workforce growth pattern reflects the fact that workforce utilization starts below its target value (the horizontal line that bisects the y-axis.) This lower-than-desired utilization induces management to increase the number of machines per worker during the first two years, which succeeds in bringing utilization up to its target value. Utilization fluctuates throughout the run due to both the industry business cycle (which causes contract penetration to move up and down by a few percentage points every two years) and the hiring seesaw effect. Despite this fluctuation, utilization remains close enough to its target value after the first two years that the number of machines per worker stays essentially flat. Loop B1 does its job effectively in this run, and no significant problems are created for the workforce, customers, or company finances.

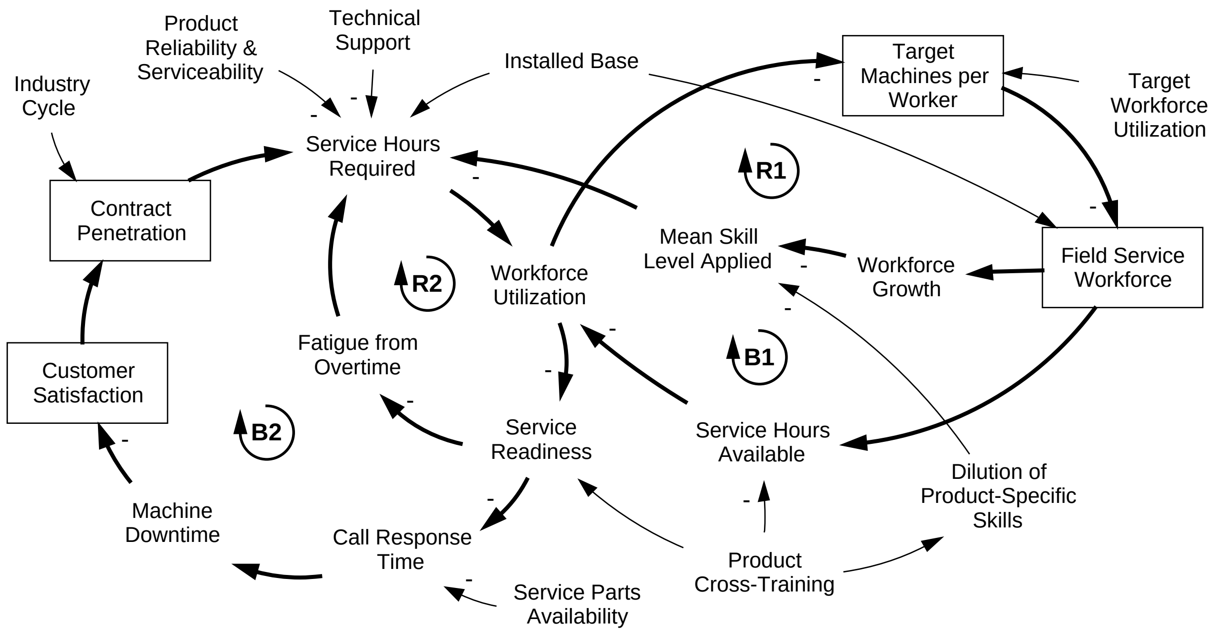

Figure 2. Feedback structure of the strategic model

Under different assumptions, the strategic model may produce results that are not so rosy as those in Figure 3. In particular, the reinforcing loop R2, involving fatigue from overtime, can under certain conditions cause problems of sharply reduced workforce productivity and service readiness. (Though not shown in Figure 2, fatigue may also lead to increased workforce turnover and even further loss of readiness.) Another concern is balancing loop B2, which indicates the loss of customer satisfaction and contract penetration that may occur when machine downtime increases due to lack of readiness. Both of these loops indicate the importance of service readiness to the stability and success of field service operations.

Cross-training is the policy lever that most directly affects service readiness. Along with its beneficial impact, increased cross-training does cause some dilution of skills and loss of availability for non-rookies, and generates additional training costs. Also, there is a common belief among field managers that the service readiness gains from additional cross-training may diminish rapidly beyond a certain level. For example, some may believe there is little readiness to be gained from having field engineers trained on more than two or three products each.

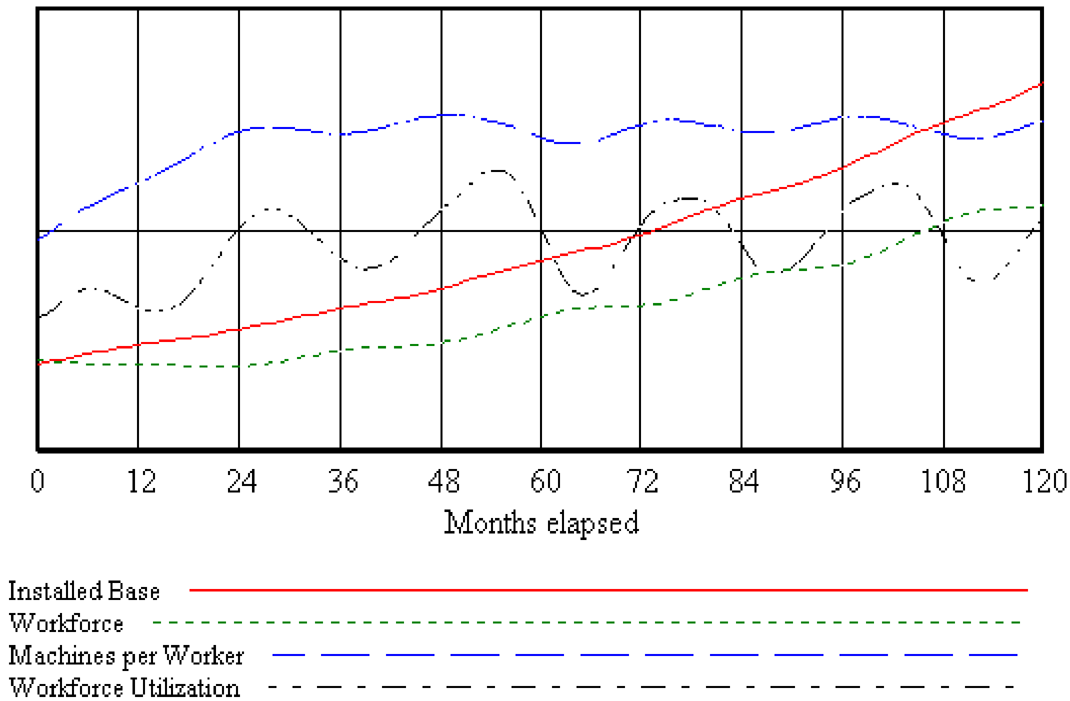

Figure 3. Workforce growth modulated by utilization target
(sample output from strategic model; variables graphed on different scales)

Figure 4 illustrates how cross-training policy can affect field service performance, as reflected in customer satisfaction. The three cross-training scenarios differ only in terms of the target for secondary skills per non-rookie; this target has been set at 1 (baseline), 0.5, and 2, respectively, for the three runs. (The baseline run is the same run that produced the output in Figure 3.) All three runs are initialized with the average number of secondary skills at 1, rather than at the different target values, so that both transient and steady-state impacts of a change in policy may be seen. In the run with a target value of 0.5 ("*Xtrain 0.5*"), the new lower level is achieved after three and a half years of no new cross-training combined with workforce turnover. In the run with a target value of 2 ("*Xtrain 2*"), the new higher level is achieved after a year and a half of intensive cross-training.

Under all three scenarios, customer satisfaction rises transiently for the first two years of the run.[2] During this early period, the short-term disadvantages of intensive cross-training (non-rookie unavailability and skills dilution) take some toll on service readiness and customer satisfaction in *Xtrain 2*. But after the early period, it is the beneficial impact of cross-training that dominates, and that leads, relative to the base run, to somewhat higher and more stable customer satisfaction under *Xtrain 2*, and significantly lower and less stable customer satisfaction under *Xtrain 0.5*. The reduced level of service readiness due to lack of secondary skills in *Xtrain 0.5* is sufficient to trigger the dangerous fatigue loop (R2 in Figure 2), which drives up the number of service hours required. This increase in work hours, in turn, leads to a boost in the hiring rate, thereby amplifying the hiring seesaw and destabilizing the primary skills chain for a number of years thereafter.

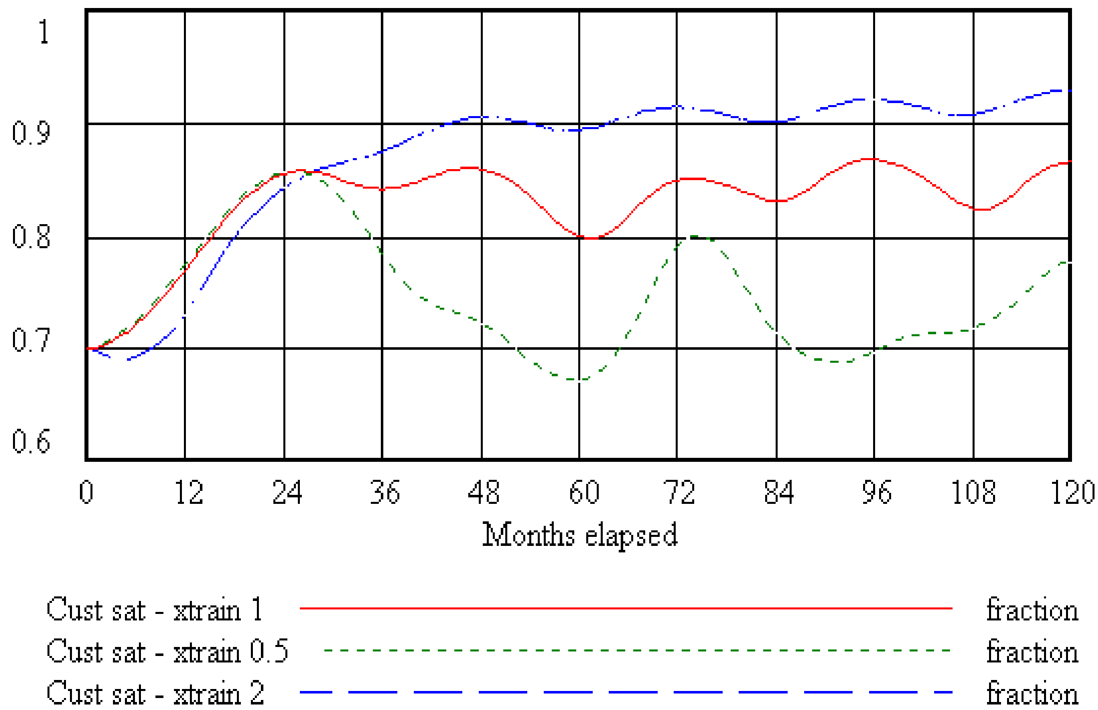

Figure 4. Impact of cross-training policy on customer satisfaction (sample output from strategic model)

Having confirmed the importance of cross-training to field service performance through such simulations, company decision-makers naturally wanted to know how large the target for secondary skills should be. The model includes metrics, such as operating margin, that may be used to compare simulation results for analyzing this decision. But to have confidence in these results, the decision-makers first required assurance that the model had accurately captured the key nonlinear relationship between the average number of secondary skills and service readiness. (The table function used in generating the runs in Figures 3 and 4 is presented in the next section as Figure 7.) Because of the sensitivity of model results to this relationship, and because of its underlying complexity in real life, such accuracy could not be assured by simply asking the company's field service managers to estimate the table function. These managers, whose expertise was critical for formulating and calibrating much of the strategic model, acknowledged that their gut feels about the table function were nothing more than that, and lacked supporting evidence. So we developed a micro-level model to answer the question.

A Model of Service Job Queueing and Assignment

Queueing models have been used for many years by operations researchers to address questions of how much and what configuration of service capacity will best meet randomly arriving service demand. The range of applications is very broad, including commercial services (ex. bank tellers, grocery checkout, auto repair), transportation services (ex. traffic lights, railroads, fire trucks), social services (ex. health care, judicial, legislative), and industrial services (ex. materials handling, equipment maintenance, quality control) (Hillier and Lieberman 1974).

Though they treat service events stochastically, conventional queueing models are essentially static, in the sense that the basic parameters of service capacity and demand, such as the number of servers and the number of customers, are considered constant. Thus, these models are not built to explore the evolution of a service system over a period of years, but rather to understand its statistical characteristics under different parametric assumptions over a period of hours, days, or weeks. These statistical characteristics include means and variances for such metrics as queue length, waiting time, and percentage of servers busy.

Figure 5 presents the basic structure of the field service job-queueing model. The focus is on a single regional hub with a complement of field engineers who work on service jobs of three types: repair, ECO (bug fixes), and preventive maintenance (PM). Time is measured in work shifts of eight hours. The hub's installed base is disaggregated by product type, and for each combination of product type and service job type there is a job arrival probability per shift, based on mean time between service. The nature of the service job types is such that, at any given time, a machine may be awaiting at most a single repair job or a single PM job, and any number of ECO jobs (corresponding to different bug fixes.) Thus, the inflow of job arrivals by type in the model is a random variable (binomially distributed) whose mean is determined jointly by the installed base, the number of jobs already in the queue, and the job arrival probability.

The remainder of the model is devoted to calculating the flow of job completions by type out of the queue. This outflow is a binomial random variable whose mean is determined jointly by the number of jobs successfully assigned to engineers and a job completion probability per shift. The job completion probability, in turn, is based on a mean service time reflecting normal preparation and travel time plus the time spent at the customer site. The on-site time itself is a function of two things which may vary by product-job type: a specified minimum time based on the performance of fully-skilled engineers, and the mean skill level of engineers assigned to the product-job type during the current work shift.

The effective assignment of engineers to service jobs is an intricate matter in real life, and the same is true in the model. On any given day, field service managers and dispatchers must attempt to achieve an effective match between the jobs in the queue and the skills of their available engineers, also taking into consideration any response-time agreements that exist with specific customers. A good approach generally is to match the more skilled engineers to the more difficult and urgent jobs in the queue, so as to minimize machine downtime and satisfy customer expectations. This is the approach taken in the model, where it has been implemented in the form of a network optimization algorithm called upon each simulated work shift to assign individual field engineers to product-job types for which they have the necessary skill.[3]

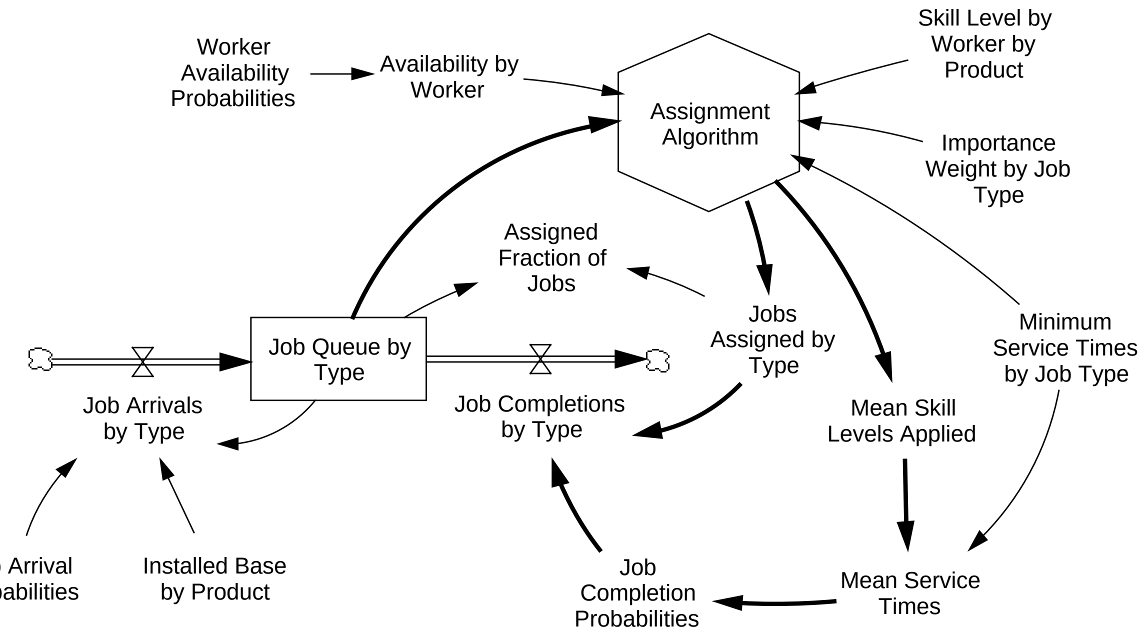

Figure 5. Structure of the job-queueing model

Figure 5 shows the ingredients that go into the assignment algorithm. Field engineers are identified individually, and a matrix specifies each engineer's skill level (from zero to one) by product type. Engineers also differ in regard to their days available to perform service jobs, due to differences in number of personal days off, training days, and days assigned to installation and upgrade teams. An engineer's availability (zero or one) during a given work shift is determined by a binomial random variable whose mean is the engineer's fraction of days available. The other required ingredients include size of the queue by product-job type, as well as the minimum service times and importance weights associated with each product-job type. Repair jobs are given a greater weight than ECO or PM jobs; for repair jobs, every hour in the queue is another hour of machine downtime, whereas for ECO and PM jobs downtime occurs only when the job is being performed. Also, certain product types may have greater importance weights than others to reflect response-time agreements the company may have with customers owning those products.

The job-queueing model was calibrated to represent two of the company's actual service hubs and in both cases accurately recreates their real-life workload situations in detail. One of these hubs has been troubled by a loss of engineers and has been overwhelmed with workload. The model shows how the overload in this hub has led to virtual abandonment of ECO and PM jobs, and it identifies those product types that have been hardest hit by the shortage of engineers. The second hub is in much better shape, with all service jobs getting attended to, although not always straightaway. For example, the model suggests that while 90% of repair jobs get an engineer assigned immediately upon arrival in the queue, the numbers are closer to 50% for ECO jobs and 20% for PM jobs. Figure 6 shows graphically how the repair jobs queue is kept under rather good control in this hub, primarily due to the high rate of job assignment. The efficient completion of jobs is also facilitated by a relatively high average skill in this hub.

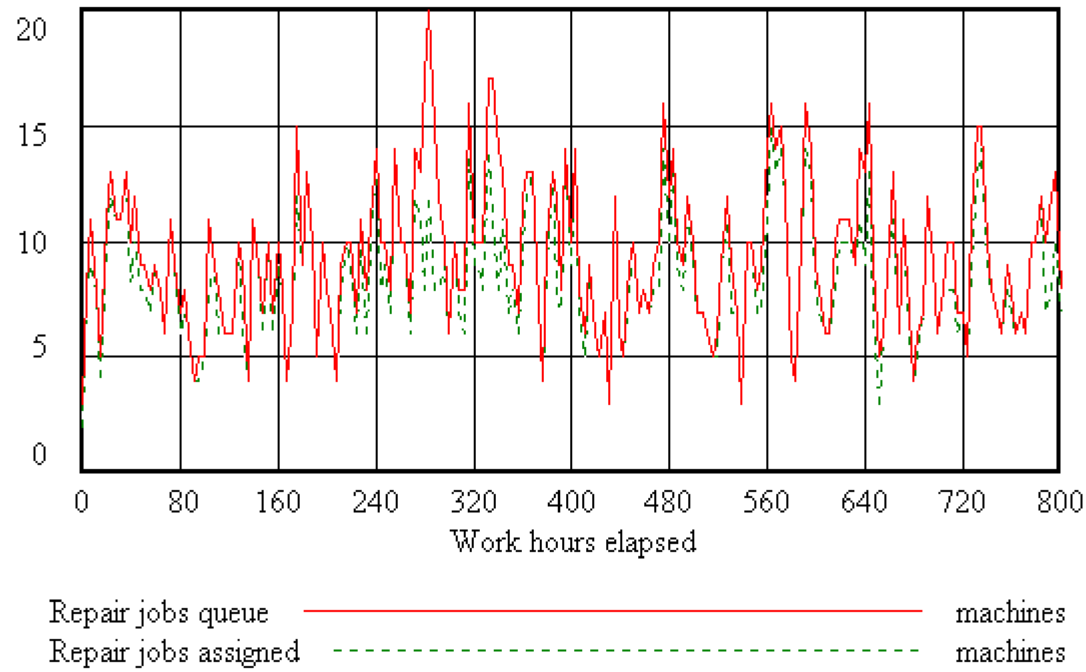

Figure 6. Randomly arriving repair jobs kept under control (sample output from job-queueing model)

Having validated the model's detailed behavior, we proceeded to do some what-if analysis for the two calibrated hubs. For example, for the troubled hub, one possible solution that had been discussed was to teach customers how to do their own preventive maintenance on certain products and thereby relieve the field workforce of the responsibility. This idea could be easily tested with the model simply by setting the PM job arrival probability to zero for those product types. This sort of tactical analysis proved quite illuminating and useful, especially for the two members of the team who happened to be the service managers of the hubs in question.

Despite its attractions as a stand-alone model for tactical decision making, the primary purpose of the job-queueing model was to support calibration of the strategic model's table function relating secondary skills and service readiness. For this purpose, a battery of sensitivity tests was performed. The team decided that the more stable of the two hubs was representative of most hubs across the U.S.A., the global region that the strategic model had been calibrated to represent. In this hub, there is an average of 0.9 secondary skills per engineer. The hub's service manager was asked to describe how he would go about changing the skills matrix if he were required to increase or reduce significantly the number of secondary skills in his workforce. (Presumably, a reduction in secondary skills would correspond to a decision to stop assigning some engineers to products for which they had been cross-trained.) In order to insure as pure as possible a test of the secondary skills effect, each alternative skills matrix had to be constructed so that the secondary skills were distributed across the product types in a balanced way, as in the baseline matrix. Also, the mean secondary skill level (ex. 0.75) had to be approximately the same in each alternative matrix as it was in the baseline matrix. Five alternative skills matrices were built in this way, with secondary skills per engineer varying from a low of 0.1 to a high of 2.5. (No proposals under consideration at the time called for more than an average of two secondary skills per engineer.)

The goal of sensitivity testing was to find, for each alternative skills matrix, a workload factor (affecting all job arrival probabilities equally) that would offset the change in secondary skills and cause the queueing system to achieve approximately the same level of performance seen in the base case. These workload factors would tell us exactly what we needed to know: the impact of a change in cross-training policy on service readiness.[4] We determined that the best performance metric to use for this purpose was the assigned fraction of jobs (ex. 40%), averaged across all product-job types and over 800 simulated work hours. For example, when testing the skills matrix with only 0.1 secondary skills per engineer, the workload had to be reduced using a factor 0.66 to bring the assigned fraction of jobs back up to its base case value. At the other extreme, when testing the matrix with 2.5 secondary skills per engineer, a workload factor of 1.65 had to be used to get the assigned fraction of jobs back down its base case level.

The five workload factors found through sensitivity testing told us how changes in the number of secondary skills would affect service readiness, relative to the base case. These relative values, combined with a base case value for the service readiness effect, would provide six points for the desired table function in the strategy model. The base case value (the impact on service readiness corresponding to 0.9 secondary skills per engineer) was adjusted in the strategy model until the initial value of the model's "current-shift assignment fraction" for repair jobs was equal to the same 90% seen in real life. This tuning process produced a base case service readiness effect of 0.5, meaning that the level of service readiness with 0.9 secondary skills per engineer is half of what it would be if all engineers were trained on all products. At a level of 2.5 secondary skills per engineer (where the relative value is 1.65), the service readiness effect rises to .825 (=0.5 x 1.65). The complete table function is shown in Figure 7.

Conclusion

It is a well-known maxim in system dynamics that model parameters should be estimated using data below the level of aggregation of model variables wherever possible. Such detailed data characterize the individual members of a level, or the individual events within a rate. For example, the best way to estimate the average lifetime of housing in an urban area may be to use official records detailing the dates of construction and demolition of individual houses (Graham 1980). However, it is not always easy or straightforward to make the leap from individual-level data to the aggregate concepts used in a strategic system dynamics model. This paper has presented a case in which a key table function affecting service readiness could only be properly estimated by analyzing a separate micro-level model, a model that mimics the daily queueing of service jobs and their assignment to individual field engineers.

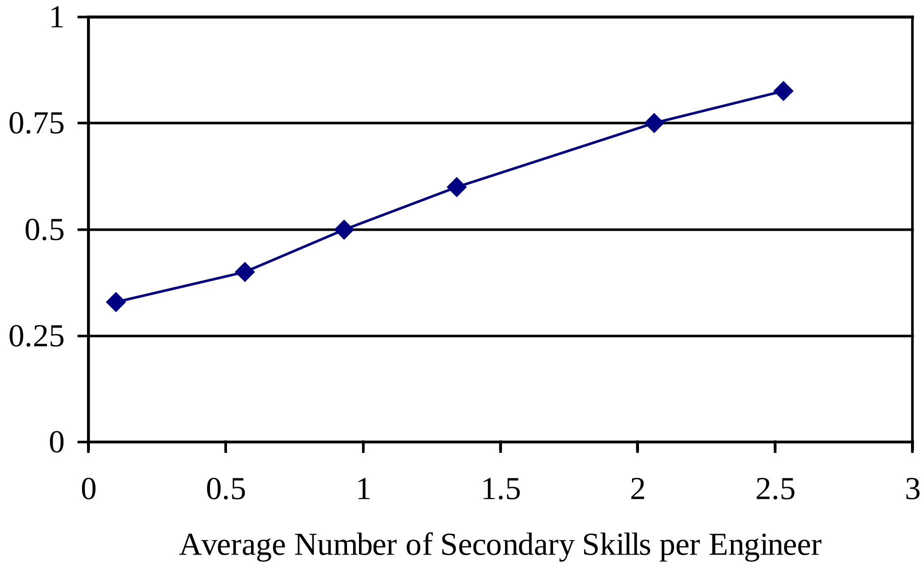

Figure 7. Service readiness multiplier from cross-training (derived from sensitivity testing of job-queueing model and used as table function in strategic model)

System dynamics modeling has certainly been used in the past for looking in a detailed way at work processes, and for drawing conclusions about the relationship of operational decisions to key process metrics. For example, Paich (1994) describes a model of semiconductor manufacturing that allowed the client to analyze the relationship between the target for finished inventory and on-time delivery percentage. The difference in the case presented here is that the micro model was not only useful in its own right for local decision making, but also supported calibration of the strategic macro model. Although the development and testing of the micro model was a time-consuming affair, it allowed the modeling team to draw important conclusions about the appropriate level of cross-training. The modeling team and senior management told me directly that these conclusions would have been viewed with skepticism had the micro model not been developed to assure the accuracy of the underlying numerical assumptions.

There are undoubtedly other opportunities for supporting high-level system dynamics models with micro models that put key aggregate relationships under the microscope of individual-level analysis. Such opportunities may lie wherever the underlying process has a stochastic, queueing aspect to it. Not only may clients be happier with a combination of macro- and micro-analysis in such cases, but the worlds of operations research and system dynamics may perhaps be brought a little closer in the process.

Notes

[1] It is assumed that a cross-trained engineer is assigned with roughly equal frequency to each of the products for which he has been trained. Under this assumption, the weight (W) on secondary skills for non-rookies is related to the average number of secondary skills per non-rookie (S) as follows: W = S/(1+S).

[2] The transient increase in customer satisfaction is related to a significant decline in the fraction of the installed base under warranty. For historical reasons, the model is initialized with an unusually large number of new machines under warranty, a condition that dissipates during the first two years of the simulation. Customers tend to have higher expectations for new products than for older ones, and are invariably disappointed when the new products are found to have bugs that affect product performance and require time to fix.

[3] The optimization routine was designed and developed by Bob Brooks of RBA Consultants, Los Angeles, California. It is compiled into a Windows Dynamic Link Library and called from Vensim as an external function each computation interval.

[4] Service readiness in the strategic model is represented by the aggregate ratio of repair hours available to repair hours required, as modified by the cross-training or secondary skills effect. One may thus say that the secondary skills effect is equivalent to a workload factor (less than one) that would exactly offset the reduction in service readiness caused by the fact that not all engineers are trained on all products. The fewer secondary skills the workforce has, the smaller the offsetting workload factor must be, and thus, the smaller the multiplier on readiness.

References

Carroll, J.S., J.D. Sterman, and A.A. Marcus. 1998. Playing the Maintenance Game: How Mental Models Drive Organizational Decisions. In *Debating Rationality: Nonrational Elements of Organizational Decision Making*, ed. J.J. Halpern and R.N. Stern. Ithaca, NY: Cornell University Press.

Graham, A.K. 1980. Parameter Estimation in System Dynamics Modeling. *In Elements of the System Dynamics Method*, ed. J. Randers. Cambridge, MA: MIT Press. Reprinted by Productivity Press, Portland, OR.

Graham, A.K., J.D.W. Morecroft, P.M. Senge, and J.D. Sterman. 1992. Model-Supported Case Studies for Management Education. *European Journal of Operational Research* 59(1): 151-166.

Hillier, F.S. and G.J. Lieberman. 1974. *Operations Research* (2nd Edition). San Francisco, CA: Holden-Day.

Levin G., E.B. Roberts, G.B. Hirsch, et al. 1976. *The Dynamics of Human Service Delivery*. Cambridge, MA: Ballinger.

Oliva, R. 1996. *A Dynamic Theory of Service Delivery: Implications for Managing Service Quality*. Ph.D. Thesis, Sloan School of Management, Massachusetts Institute of Technology.

Paich, M. 1994. Re-engineering: The Need for Understanding the Interdependencies Between Process Metrics. In *ithink® Business Applications Guide*. Hanover, NH: High Performance Systems.

Richmond, B. 1994. Classic Core Business Infrastructures: Inadvertent Evolution and Fighting the Physics. In *ithink® Business Applications Guide*. Hanover, NH: High Performance Systems.

Senge, P.M. and J.D. Sterman. 1992. Systems Thinking and Organizational Learning: Acting Locally and Thinking Globally in the Organization of the Future. *European Journal of Operational Research* 59(1): 137-150.