The Efficiency-Quality Trade-Off of Cross-Trained Workers

Edieal J. Pinker • Robert A. Shumsky
William E. Simon Graduate School of Business Administration
University of Rochester
Rochester, New York

Manufacturing & Service Operations Management
Winter 2000

Presented by: Dan McCarthy

Motivation

Generalists vs. Specialists in Medicine
Labor and Delivery Story

• Why wasn’t there an obstetrician on duty?
• Staffing flexible servers (i.e. generalists) is more efficient given heterogeneous customers, all else being equal
• What about cost?
• What about speed of service?
• What about quality of service?

Call Centers
Other fields?
Agenda

- Overview
- Model Formulation
  - Service Process Model
  - Tenure Process
  - Service Quality and the Value to the Firm of Worker Experience
- Service Process Approximation Method
- Numerical Experiments
  - General Model Testing and Insights
  - Case Study
- Conclusions
- Critique
- Questions / Discussion
Overview

**Goal:** Study the trade-off between the cost efficiency provided by cross-trained (or generalist) workers and the experience based quality provided by specialists

Develop a general model that integrates:
- Queuing system model that includes multiple server types
- Model of an individual worker’s career path (tenure)
- Model of experience-based learning

• Output of system = Revenue (varies with the quality of service)
• System Performance = Gross Profits (varies with both revenue and costs)

**Links managerial decisions about staffing policies and worker specialization with worker learning curves, system costs and service quality**
Model Formulation

• **The Service Process Model**
  – Who gets served?
  – By whom?

• **Employee Tenure Model** (*Tenure Process*)
  – Experience of Servers

• **Experience-Based Learning Model of Service Quality**
  – Experience $\rightarrow$ Service Quality
  – Service Quality $\rightarrow$ Value to the Firm (Revenue)

• **Objective Function**
  – Expected Profit of the Firm
The Service Process Model

(See Figure 1 on page 34 of the Pinker and Shumsky paper)

• **Focuses on quality of service** \[ f(\text{server experience}) \]
  – Traditional focus is on waiting time or time in system

• **Assumes that service standards** (e.g. % customers served) **are set exogenously**
  – Treated as constraints in the model

• **Models the SP structure as a “loss system”** (i.e. queuing not allowed)
  – Above routing scheme achieves the highest server utilization
The Service Process Model

Statistics

Throughput:

\[ R = R_{AA} + R_{BB} + R_{AF} + R_{BF} \]

(Erlang’s Loss Formula, Approximation Method)

Server Utilization:

\[ \rho_{AA}, \rho_{BB}, \rho_{AF}, \rho_{BF} \]

\[ \rho_{AF} = \frac{R_{AF}}{(\mu_F N_F)} \]

(Little’s Law)
Tenure Process

- **Tenure** defined
- Model tenure as a random variable drawn from a *mixed exponential probability distribution*

(See Figure 2 on page 35 of the Pinker and Shumsky paper)

\[ \lambda_2 > \lambda_3 \quad \quad \quad p = \frac{\lambda_1}{\lambda_1 + \lambda_2} \]

- Career path model – can be modeled as states of a continuous time Markov chain
- Time a worker stays in a given stage is exponentially distributed
- What is the *expected length of tenure*?
Tenure Process Statistics

- $x = $ worker tenure of a worker

\[
\Pr\{x > t\} = G_x(t) = \frac{\lambda_2 - \lambda_3}{\lambda_1 + \lambda_2 - \lambda_3} e^{-(\lambda_1 + \lambda_2)t} + \frac{\lambda_1}{\lambda_1 + \lambda_2 - \lambda_3} e^{-\lambda_3 t}
\]

- $y = $ time worked

\[
E(x) = \frac{1}{\lambda_1 + \lambda_2} \left( 1 + \frac{\lambda_1}{\lambda_3} \right)
\]

\[
g_y(t) = \frac{G_x(t)}{E[x]}
\]
Service Quality and the Value to the Firm of Worker Experience

How does worker experience translate into monetary value to the firm?

Worker Experience \rightarrow Service Quality \rightarrow Firm Revenue

Experience-based Service Quality

\[ Q(b) = Lb^{n_1} \]
Service Quality and the Value to the Firm of Worker Experience

How does worker experience translate into monetary value to the firm?

Worker Experience → Service Quality → Firm Revenue

Value of Increased Quality

\[ V(q) = Mq^{n_2} \]
Service Quality and the Value to the Firm of Worker Experience

How does worker experience translate into monetary value to the firm?

Worker Experience $\rightarrow$ Service Quality $\rightarrow$ Firm Revenue

Experience-based Service Quality

$$Q(b) = Lb^{n_1}$$

Value of Increased Quality

$$V(q) = Mq^{n_2}$$

Revenue generated by a worker with $b$ time units of experience in a particular task

$$W(b) = K b^n$$

$$K = L^{n_2} M , \quad n = n_1 n_2$$
Objective Function

Maximization of Expected Gross Profits

Revenue = \sum_i \sum_j R_{ij} E[W(b_{ij})]

Where:  
\begin{align*}
  i &= A \text{ or } B \\
  j &= A, B \text{ or } F
\end{align*}

Cost = c_A N_A + c_B N_B + c_F N_F

Expected Profit of the firm:  \[ Z = \text{Revenue} - \text{Cost} \]
Approximating the Average Quality

Link the models for tenure, learning and service

\[
E \left[ W(b_{ij}) \right] = K_{ij} E_y \left[ E \left[ b^{n}_{ij} \mid y \right] \right]
\]

conditioning the probability distribution of \( b_{ij} \) on \( y \)

\[
E \left[ W(b_{ij}) \right] \approx \int_{0}^{\infty} K_{ij} (\rho_{ij}t)^n g_y(t) dt
\]

\[
= \frac{\lambda_3 (\lambda_1 + \lambda_2) K_{ij} \Gamma(n+1) \rho_{ij}^n}{(\lambda_2 - \lambda_3) \lambda_3 + \lambda_1 (\lambda_1 + \lambda_2)} \left\{ \frac{(\lambda_2 - \lambda_3)}{(\lambda_1 + \lambda_2)^{n+1}} + \frac{\lambda_1}{\lambda_3^{n+1}} \right\}
\]
Approximation Error

\[ \varepsilon(Y) = \frac{E[b_{ij}^n | Y] - (\rho_{ij} Y)^n}{(\rho_{ij} Y)^n} = E\left[ \left( \frac{b_{ij}}{\rho_{ij} Y} \right)^n \right] - 1 \]

**Lemma 1:**

\[
\lim_{Y \to \infty} E\left[ \left( \frac{b}{\rho Y} \right)^n \right] = 1
\]

**Implies:** \( \lim_{Y \to \infty} \varepsilon(Y) = 0 \)

How long does it take for the true expectation and the approximation to converge?

(i.e. How long does a server have to be on the job for the approximation to be close enough?)
Simulation Results

(See Figure 3 on page 39 of the Pinker and Shumsky paper)
Impact of Staff Mix on **Cost**

\[
\frac{c_A N_A + c_B N_B + c_F N_F}{\lambda_A + \lambda_B}
\]

(See Figure 4 on page 41 of the Pinker and Shumsky paper)
Impact of Staff Mix on Quality

Quality / Customer:

\[ \frac{\sum_i \sum_j R_{ij} E[W(b_{ij})]}{\lambda_A + \lambda_B} \]

(See Figure 5 on page 41 of the Pinker and Shumsky paper)
What is the **Right** Staff Mix?

- Trade-off between efficiency and quality
- Only generalists, only specialists, or **OPTIMAL mix**

(See Table 1 on page 42 of the Pinker and Shumsky paper)

- Suggests that there is an optimal mix
- *When is it important to determine the optimal mix?*
When is the Optimal Staff Mix Important?

(See Figure 6 on page 42 of the Pinker and Shumsky paper)
Which Extreme System is *Best*?

<table>
<thead>
<tr>
<th>Optimal</th>
<th>Specialized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexible</td>
<td></td>
</tr>
</tbody>
</table>

(This graphic can be superimposed over Figure 7 on page 43 of the Pinker and Shumsky paper)
Optimal Staffing Configurations

(See Figure 8 on page 44 of the Pinker and Shumsky paper)
System Performance - Profit

(See Figure 9 on page 45 of the Pinker and Shumsky paper)
System Performance - Quality

(See Table 3 on page 45 of the Pinker and Shumsky paper)
Conclusions

• “Flexible workers provide more throughput while using fewer workers” – from the Pinker and Shumsky paper, page 46.
• (For other conclusions, see page 46 of the Pinker and Shumsky paper)
Shortcomings / Future Work

- Service systems in practice are substantially more complex than the system described in the paper
- Many service systems allow customers to queue for service
  - Could extend the model to reflect the impact of waiting time on perceived service quality in the performance measure Z
- In addition to learning, many times forgetting is an important phenomenon
  - Could incorporate forgetting into the learning curve model to allow quality to drop when utilization dips below some threshold
- Extensions to the model to represent a more elaborate service facility:
  - More customer and service classes
  - Heterogeneous service rates and tenure processes among server classes
- Other uses for the Model:
  - Assess the benefits of training programs and IT that transfers knowledge between specialists and flexible workers
  - Assess personnel assignment decisions (e.g. job rotation to prevent burnout)
Critique

- **Main Contribution:** First work to integrate the study of service process systems, learning curves, and the modeling of turnover and career paths.
- **Provides a fair amount of practical insight that could be useful**
  - When is it important to have the optimal mix vs. using an extreme solution
  - All-specialists seems to dominate all-flexible except in the extreme case of a small system and little learning
- **Provides a simple framework for expanding the study to include exploring other options**
- **I would like to see how the single-overflow configuration would have performed relative to the others (for quality) if optimized for Z (instead of min Cost)**
- **The paper is well written with the right amount of mathematical detail for the intended purpose**
Discussion