University of St. Thomas, Minnesota
UST Research Online

Operations and Supply Chain Management Faculty Publications

Operations and Supply Chain Management

2012

An Experimentally Confirmed Resource Planning Model of Services Under Production Function Uncertainties

Sheneeta W. White

University of St. Thomas, Minnesota, whit6237@stthomas.edu

Follow this and additional works at: http://ir.stthomas.edu/ocbopmtpub

Part of the Business Administration, Management, and Operations Commons, and the Management Sciences and Quantitative Methods Commons

An experimentally confirmed resource planning model of services under production function uncertainties

Sheneeta W. White*

University of St. Thomas, Operations and Supply Chain Management Department, 2115 Summit Ave (MCH 329), St. Paul, MN 55105, USA

ARTICLE INFO

ABSTRACT

The way in which service firms transform inputs into outputs is typically uncertain or unknown. Consequently decision makers, at best, can only make estimates of the underlying production function. The purpose of this study is to offer policy recommendations to service providers and to examine experimentally the sensitivity of estimates of production functions. The primary result of this study is a foundation for modeling manpower planning decisions for co-produced services when production functions are mis-estimated and/or mis-specified.

1. Introduction

Managing resources is a challenge in service operations planning for two reasons. First, having the client as a resource introduces additional variation into the service process. Client variability is in the forms of knowledge, abilities, and motivation (Frei, 2006). Second, complex services must be flexible enough to deal with resource requirement changes. In complex services, there is a high-degree of customer contact and therefore requirement changes are common. These changes can range from the number of resources required, to the desired capability of each resource (Dietrich, 2006). Therefore, practical resource planning models must account for client variability and must be capable of handling resource requirement changes.

A production function is the mathematical function that finds the maximum output possible from a given level of input. In service organizations, production functions are difficult to estimate because the conversion process is unclear. Lack of clarity results in inaccurate production function estimates that can greatly affect resource planning decisions. In manufacturing, the model of the conversion of inputs to outputs is typically known. The production function must be well defined in manufacturing capacity planning decisions. In contrast, there exist many sources of uncertainty in the transformation process for services, see Dietrich (2006). Inherit variability in the transformation process is a source of uncertainty.

The goal of this study is to offer policy recommendations to service providers regarding how to best allocate resources when the mathematical function that maps inputs to outputs is unclear. We measure sensitivity of a resource planning model to uncontrollable inputs, mis-estimation of function parameters, and mis-specification of the function form that defines the transformation process. When a service provider is unsure of the structure of the function (mis-specification) or unsure of the parameter values of the function (mis-estimation) or the uncontrollable inputs the resource plan may be inefficient.

The author develops a deterministic linear model for the purposes of determining benchmark service inputs levels for given target output quantities and a stochastic model for resource planning. This paper provides experimental results based trade-offs between inefficiencies, risk, and costs versus resource allocations. Badinelli (2010) provides a theoretical base for a similar model. Using a theoretical-based approach to service resource planning is limited because it does not thoroughly account for the dynamic nature of services. Service system modeling should investigate the decision problem through what-if scenarios. The model presented in this paper builds upon Badinelli (2010) using experimental analysis to illuminate model performance under varying conditions.

The contributions of this study are as follows:

- This research uses a stochastic production function for resource planning. Some service-based models use Chance-Constrained DEA, see Land et al. (1993), to incorporate stochastic considerations in inputs and output measures, but these models are not used for resource allocation. Mula et al. (2006) give a review of stochastic production planning models, but not all of them are specific to service systems.
- We have a foundation for modeling resource allocation decisions for high-value adding service systems when production functions are mis-estimated and/or mis-specified. The field of

* Tel.: 651 962 5438; fax: 651 962 5093.
 E-mail address: whit6237@stthomas.edu

service science is still in its infancy, therefore any policy recommendations that can be generated regarding resource-planning for service systems is greatly needed.

- This model that behaves intuitively. This is a well-constructed model that accurately captures the effect of resource allocations on service output levels.

In the next section, we discuss a few pieces of work relevant to the study at hand, followed by a discussion of the motivating decision problem. Then we present our model for resource allocation followed by the results of the experimental analysis. The final section discusses the contribution of the research and future research directions.

2. Discussion of related work

There are many service-based planning models in the literature. In this section, we will only discuss those resource-planning models that define explicit production functions. Many of these models use a Data Envelopment Analysis (DEA)-based approach. DEA, by its design, was not intended for resource allocation but for measuring relative efficiency of service units. In DEA no assumptions are made about the underlying transformation process. However, there are a few papers that define an explicit production function. Golany and Tamir (1995) use DEA to build a resource allocation model. The authors use an empirical, multiplicative production function. Athanassopoulos (1998) develops a model for resource allocation and target setting using DEA and goal programming. The author uses the principal-agent paradigm and a nonlinear production function to achieve resource allocation. Lastly, Golany et al. (2006) develop a DEA based model where the efficiency of the subsystems and the aggregate system is determined. The authors develop a deterministic linear acquisition model for the purposes of resource allocation among subsystems and a deterministic linear model for measuring efficiency of an aggregate system. This paper used a Cobb–Douglas production function.

The models presented in this study differ from the aforementioned models using stochastic production functions for resource planning and by the usage of controllable and uncontrollable inputs in the models.

3. Motivating resource planning decision

The service process has inputs from both the service provider and the client. The service provider and the client must determine the values of the decision variables of this resource-planning problem, which are the quantities of inputs to allocate to each service process. Managers want to allocate resources in order to improve their day-to-day operations and ultimately to better position themselves in the market place.

This study focuses on resource planning for a particular service type within a particular service firm. A service type is a well-defined, value-adding experience that is offered by the provider to the client, such as processing a mortgage loan application, admitting a patient to a hospital, or teaching a class.

A set of different process types is needed to deliver a service type. For example, the service type of processing a mortgage loan application includes the process types of interviewing the applicant, gathering required documents, and submitting the application for underwriting. Each process type uses multiple inputs to generate multiple outputs. A service engagement represents an instance of a service type. An example of the service system is depicted in Fig. 1. The outputs should reflect all useful outcomes on which we wish to assess the service engagement.
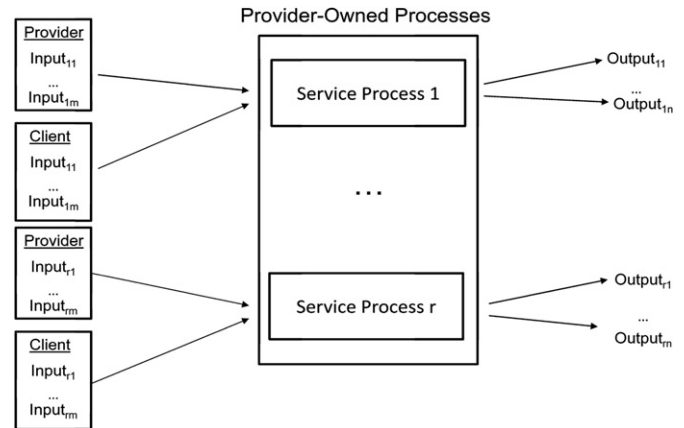

Fig. 1. Service system.

In this model, the degree of client involvement is discretionary. Client involvement becomes discretionary when the service provider has the ability to limit or increase the amount of client contact in some way. For example, consultants can decide how many times of day they want to meet with a particular client. In mathematical modeling, when the degree of client involvement is discretionary its identifier changes from a parameter to a decision variable.

Utility theory assumes that every decision maker has preferences towards risks and return and that the decision maker will choose the alternative that maximizes his/her utility. Based on the decision makers' preferences, there are different weights placed on the loss functions and the weights reflect the relative value of each loss function. Service providers and clients want to ensure that they are allocating resources in the most efficient manner based on what they do know about the production function in an effort to minimize the costs of underproduction and over allocation.

4. The model

Two optimization models are developed in order to determine the optimal resource allocation plan. We take a two-tiered approach in order to avoid arbitrarily setting benchmark levels of resources needed to maximize output. Our goal is to find the most efficient production function mapping of inputs to outputs. Therefore, results from the first model are passed to the second model.

The first optimization model is a deterministic optimization model. The results from the first model are benchmark levels of input resources from the service provider and client. The benchmark production function, which is comprised of given usage and yield rates, is what the service provider believes to be the most efficient function mapping inputs to outputs.

The second optimization model is a stochastic, resource-planning model. This model takes the benchmark input levels from the first model and develops the resource allocation plan for the service engagement. The stochastic element of this model is represented by a probability density function that captures the deviation from the benchmark output levels reflecting both inefficiencies and uncertainties.

4.1. Model assumptions

We measure the output level of each process independently. We assume each process has its own production function. There are no setup costs associated with re-allocating resources to a service engagement. It is assumed that the provider and the client each have fixed capacities per input to be allocated to all service processes. There is a linear cost for each input.

The benchmark usage and yield rates are estimated using historical data. Target output levels for each process are provided by the client to each optimization model.

We assume there is a density function of each output level. The input levels are parameters of this density function. As resource allocations change the shape of this density function will change. We assume that the density function represents a service engagement's deviation from the benchmark output level. This deviation reflects both inefficiency and uncertainty in the production function form and parameters.

We assume that the service provider delivers a service through a service system consisting of linked processes—each process delivering one type of service component. We also assume that the precedence constraints among the processes that are required for a particular service define a network of these processes—a service supply chain. Hence, for each service, there is a known network of processes. Each process of the network requires a certain number of "cycles" per unit of the service component that is delivered to the client.

4.2. Descriptive models

Indices (*general to both models*)

p	process type index
i	input index
j	output index
r	number of process types; $p=1,\ldots,r$
m_w	number of inputs from the service provider; $i=1,\ldots,m_w$
m_c	number of inputs from the client; $i=1,\ldots,m_c$
n	number of outputs; $j=1,\ldots,n$

Optimization Model #1

The first optimization model applies a benchmark production function to obtain benchmark input levels for given target output quantities. The model minimizes the input levels from the service provider and the client across all processes. Constraints (M1.1) ensure that the number of process cycles that can be generated by each service provider resource is greater than or equal to the number of completed process cycles. Constraints (M1.2) ensure that the number process cycles that can be generated by each client resource is greater than or equal to the number of completed process cycles. Constraints (M1.3) ensure that the number process cycles needed to generate a particular given output target is less than or equal to the number of completed process cycles. Constraints ((M1.1)–(M1.3)) balance the resources received by each service process with that process' number of generated cycles, (i.e., inflow=outflow). Constraint (M1.4) and (M1.5) are the service provider and client capacity constraints, respectively.

Decision variables

$\hat{x}_{ip}^{w}$	benchmark quantity of resource i allocated to process p by the service provider
$\hat{x}_{ip}^{c}$	benchmark quantity of resource i allocated to process p by the client
ν_p	the number of benchmark process cycles of process p, which are completed

Parameters

$\hat{y}_{jp}$	target quantity of output j by process p
a_i^{w}	available service provider capacity of resource i
a_i^{c}	available client capacity of resource i
β_{pi}	benchmark generation rate of process p that is supported by resource i (cycles/unit of resource)

α_{pj}	benchmark generation rate of process p that is required by output j (cycles/unit of resource)
$\mu_{pi}=1/\beta_{pi}$	benchmark usage rate of resource i per cycle of process p (units of resource/cycle)
$\gamma_{pj}=1/\alpha_{pj}$	benchmark generation rate of output j per cycle of process p (units of resource/cycle)

$$\text{Minimize} \sum_{i,p}\hat{x}_{ip}^{w} + \sum_{i,p}\hat{x}_{ip}^{c}$$

Subjected to

$$\frac{\hat{x}_{ip}^{w}}{\mu_{pi}} \geq \nu_p \quad \text{for all} \ \ i=1,\ldots,m_w, \quad p=1,\ldots,r \tag{M1.1}$$

$$\frac{\hat{x}_{ip}^{c}}{\mu_{pi}} \geq \nu_p \quad \text{for all} \ \ i=1,\ldots,m_c, \quad p=1,\ldots,r \tag{M1.2}$$

$$\frac{\hat{y}_{jp}}{\gamma_{pj}} \leq \nu_p \quad \text{for all} \ \ j=1,\ldots,n, \quad p=1,\ldots,r \tag{M1.3}$$

$$a_i^{w} - \sum_p \hat{x}_{ip}^{w} \geq 0 \quad \text{for all} \ \ i=1,\ldots,m_w, p=1,\ldots,r \tag{M1.4}$$

$$a_i^{c} - \sum_p \hat{x}_{ip}^{c} \geq 0 \quad \text{for all} \ \ i=1,\ldots,m_c, \quad p=1,\ldots,r \tag{M1.5}$$

all variables are nonnegative.

Production Function

For the production functions applied in the model, the resources must be procured according to usage rates of a process cycle and outputs are generated according to yield rates of a process cycle. The way in which we represent a production function is a special case of the linear constant returns-to-scale (CRS) production function used by Athanassopoulos (1998). Therefore, the input–output correspondence per process cycle forces all outputs and inputs to be in fixed proportions with respect to one another.

The linear CRS production function can be written as

$$T\,\vec{x}_p = \vec{y}_p \tag{1}$$

where, $T=\lfloor t_{jip}\rfloor_{n\times m}$. We define the component of the production matrix for row j and column i as, $t_{jip}=y_{jp}/x_{ip}$.

$$t_{jip} = \frac{y_{jp}}{x_{ip}} = \frac{\beta_{pi}}{\alpha_{pj}} = \frac{\gamma_{pj}}{\mu_{pi}} = \gamma_{pj}\beta_{pi} \tag{2}$$

Optimization Model #2

Optimization Model #2 is a stochastic, resource-planning model. The first part of the objective function, the integral, is similar to the stock-out loss function in a newsboy model. This integral captures the effect of a service engagement not generating output at the target output levels. There are penalties/weights placed on underproduction. The weights on underproduction and the target output levels are determined by the service provider and the client. The second part of the objective function captures the costs of allocating more than the benchmark input quantities of service provider and client resources to a service engagement. The constraints (M2.1) and (M2.2) are service provider and client capacity constraints, respectively. Optimization can be done via standard search routines.

T density function of actual output is a function of inputs. As resource allocations change, the shape of the density function of y_{jp} will change. This is because the resource levels are incorporated into the parameters of the distribution; see Eqs. (10) and (11) below.

Decision Variables

x_{ip}^{w}	quantity of input i allocated to process p provided by the service provider

x_{ip}^c quantity of input i allocated to process p provided by the client

Performance Measures

y_{jp} actual quantity of output j achieved by process p

Parameters

$\hat{y}_{jp}$ target quantity of output j by process p

$f_{y_{jp}}(y_{jp}; x_{ip})$ the probability density function of output j for process p

$\hat{x}_{ip}^w$ benchmark quantity of input i allocated to process p provided by the service provider; this quantity is obtained from the solution of Optimization Model #1

$\hat{x}_{ip}^c$ quantity of input i allocated to process p provided by the client; this quantity is obtained from the solution of Optimization Model #1

c_{jp}^u weight applied to under-production of output j from process p

c_{ip}^o cost of over allocation of service provider input i for process p

c_{ip}^o cost of over allocation of client input i for process p

a_i^w available service provider capacity of input i

a_i^c available client capacity of input i

$$\underset{\{x_{ip}^w\},\{x_{ip}^c\}}{\text{Minimize}} \sum_{p=1}^{r} \sum_{j=1}^{n} c_{jp}^u \int_0^{\hat{y}_{jp}} (\hat{y}_{jp}-y_{jp}) f_{y_{jp}}(y_{jp}; x_{ip}) dy_{jp}$$
$$+ \sum_{i \in S_w} c_{ip}^o(x_{ip}^w - \hat{x}_{ip}^w) + \sum_{i \in S_c} c_{ip}^o(x_{ip}^c - \hat{x}_{ip}^c)$$

subjected to

$$a_i^w - \sum_p x_{ip}^w \geq 0 \quad \text{for all } i,p \tag{M2.1}$$

$$a_i^c - \sum_p x_{ip}^c \geq 0 \quad \text{for all } i,p \tag{M2.2}$$

all variables are nonnegative. There exists a unique solution to Optimization Model #2; see Badinelli (2010). Consider random variation in the elements of the matrix $\vec{\gamma}_p \vec{\beta}_p^T$. Define

$$\vec{g}_p = \vec{\gamma}_p - \vec{z}_{gp} \tag{5}$$

$$\vec{b}_p = \vec{\beta}_p - \vec{z}_{bp} \tag{6}$$

where $\vec{z}_{gp}, \vec{z}_{bp}$ are non-negative random variables. Approximation:

$$(\vec{z}_{gp}\vec{\beta}_p^T + \vec{\gamma}_p \vec{z}_{bp}^T + \vec{z}_{gp}\vec{z}_{bp}^T) \approx [\tau_{jip} + \varepsilon_{jip}] \tag{7}$$

where τ_{jip}=constant, $\varepsilon_{jip} \sim N(0,\sigma_{ji})$ and $(\tau_{jip}/\sigma_{jip}) > 3$. $\tau_{jip}+\varepsilon_{jip}$ is an nxm matrix of normal random variates with positive mean values and negligible probabilities of negatives values. The constant τ_{jip} represents the overall level of inefficiency of the process type. We considered deviations from the benchmark recipe as inefficiencies. Therefore

$$\vec{y}_p = \frac{\vec{\gamma}_p \vec{\beta}_p^T}{m} \vec{x}_p - [\tau_{jip} + \epsilon_{jip}]\vec{x}_p \tag{8}$$

or

$$\vec{y}_p = T \vec{x}_p - [\tau_{jip} + \epsilon_{jip}]\vec{x}_p \tag{9}$$

which implies, $y_{jp} \sim N(\mu_{yjp}, \sigma_{yjp})$, where

$$\mu_{yjp} = \sum_{i=1}^{m} (T_{jip} - \tau_{jip}) x_{ip} \tag{10}$$

$$\sigma_{y_{jp}}^2 = \sum_{i=1}^{m} x_{ip}^2 \sigma_{jip}^2 \tag{11}$$

5. Experimental results

A series of cases were run via Microsoft Excel Solver®. The experiments show plots of optimal solutions for different parameter settings, such as capacities, benchmark output levels, risk levels, inefficiency levels, and penalty costs. For illustrative purposes we chose eight input types (four from the service provider and four from the client), two output types, and three process types. The base case is specified in Table 1 and the benchmark provider and client resource quantities are specified in Table 2. The output targets are specified in Table 3. The parameter values for the base case were chosen based on reasonable assumptions. We acknowledge that an empirical study or case study needs to be performed in order to have more accurate estimates of the parameter values.

Each case/experiment was designed to illuminate model behavior under specific conditions. We want to highlight the elements of the model that distinguish this study from others in literature and gain managerial insights. For example, we designed cases to show the behavior of the loss function under certain conditions. We also designed cases highlighting production function uncertainties. Other cases show the effects of the two-tiered modeling approach by examining how parameters of the first optimization model effect resource allocations in the second optimization (e.g., changes in benchmark output rates). Note: we averaged the service provider and client resource quantities when displaying the results.

Case 1. Resource Allocation vs. Output Target

In this experiment we varied the output target of Output 1, while keeping fixed the output target of Output 2 at the base case value.

Table 1
Base case parameters.

a_i^w and a_i^c	c_{ip}^o	c_{jp}^u	β_{pi}	α_{pj}	γ_{pj}	τ_{jip}	σ_{jip}
150	20	50	10	1	1	3	1

Table 2
Benchmark resource quantities.

	Process 1	Process 2	Process 3
$\hat{x}_{ip}^w$			
Resource 1	40	50	40
Resource 2	40	35	35
Resource 3	80	75	70
Resource 4	10	10	10
$\hat{x}_{ip}^c$			
Resource 1	40	40	30
Resource 2	10	20	12
Resource 3	30	30	20
Resource 4	50	25	53

Table 3
Benchmark output targets.

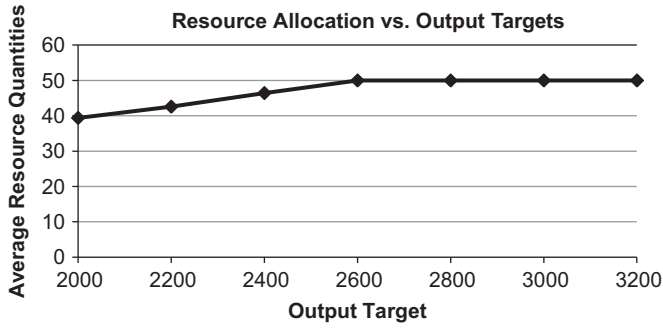| $\hat{y}_{jp}$	Process 1	Process 2	Process 3
Output target 1	2000	2000	2000
Output target 2	2000	2000	2000

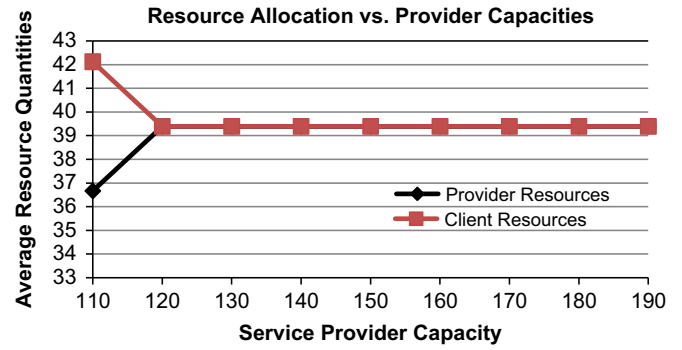Fig. 2. Resource allocation vs. output target.

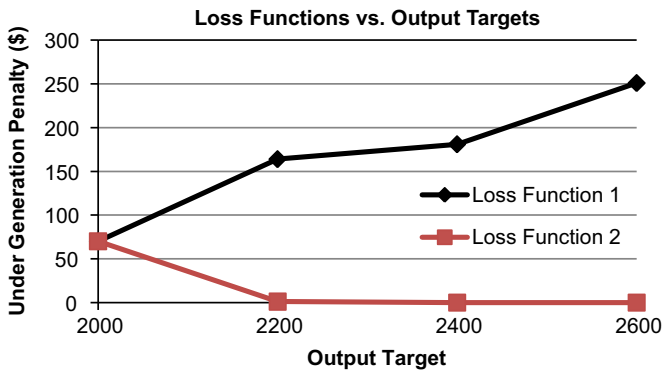Fig. 4. Resource allocation vs. service provider capacities.

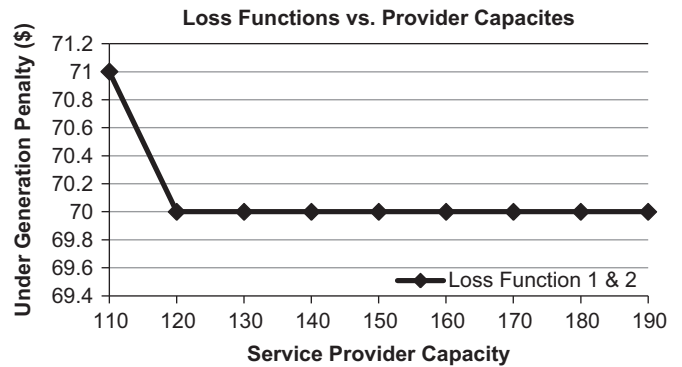Fig. 3. Loss functions vs. output targets.

Fig. 5. Loss functions vs. service provider capacities.

The results of this experiment show as the output target increased for Output 1, the average allocation of service provider and client resources increased in order to meet the output target; see Fig. 2. The resource quantities leveled off once all capacities had been used. Since there is no hiring included in this model, managers should understand that they will incur penalty costs if the workforce level is not sufficient to meet changes to output targets.

There are other interesting results of this experiment. If we compare the two loss functions, we see that, as the output target level increases, the penalty of the loss function for Output 1 also increases; see Fig. 3. The loss function for Output 1 is increasing despite the increase in resources. The loss function representing Output 2 decreases because resource quantities are increasing and it is easier to meet targets with more resources. We only show output targets up to 2600 in Fig. 3, so that we can highlight the difference between the two loss functions.

Case 2. Resource Allocation vs. Worker Capacities

In this experiment we varied the service provider capacities across all inputs and held fixed the client capacities at the base-case value.

The results of this experiment show that when service provider capacities are low there is a need to use more client resources in order to meet output target levels; see Fig. 4. As service provider capacities are increased, fewer client resources were needed. This behavior levels off after output target levels are achieved.

Service capacity constraints are not like those in manufacturing. When service provider capacity constraints are binding, client resources are allocated in order to keep the costs of missing the

output targets low. As service provider capacities are increased, there is no longer a need to allocate more client resources because the cost of adding more resources outweighs the costs of missing the output targets. This is a behavior that is not typical in manufacturing. There is more flexibility in services. The production function matrix allows for the exchange between service provider and client resources. There are variations in the "recipe"—more than one way to achieve the output.

In Fig. 5, the values of loss function for both output types are equal. When the service provider's capacity level increases from 110 to 120 there in an increase in provider resources and as a result there is a decrease in both loss function penalties. It is easier to meet output targets with more resources. The penalties of missing output targets level off as provider and client resource allocations level off.

Case 3. Resource Allocation vs. Underproduction Penalty

In this experiment we varied the weight on the loss function for Output 1, while keeping fixed the weight on the loss function of Output 2 fixed at the base case value.

The results of this experiment show that as the weight on the first loss functions is increased, more service provider and client resources are allocated in order to meet the target output levels and keep costs low.

The concavity seen in Fig. 6 is due to the nonlinearity in the objective function. When the plot of the average resource quantities is concave, it exhibits diminishing marginal investment as the loss function weights increase. The resource costs and the loss make up the objective function. At optimality, there is a trade-off between resource costs and loss. The decision model finds an
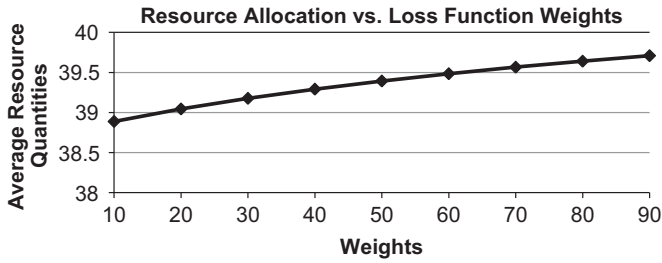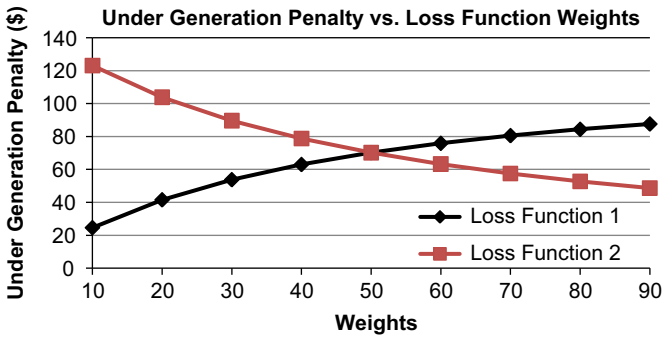
Fig. 6. Resource allocation vs. loss function weights.

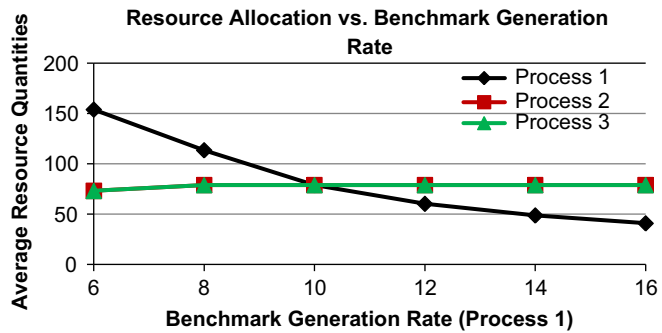**Fig. 7.** Under generation costs vs. loss function weights.

Fig. 8. Resource allocation vs. benchmark generation rate.

optimal solution that is a compromise between the increasing resource costs versus the increase in loss.

A plot of the two loss functions for this experiment is shown in Fig. 7. As we increased the weights on Output 1 for missing output targets, the costs for Output 1 of the missing output targets also increased. Due to the increase in resource allocations as seen in Fig. 6, output targets for Output 2 are easily met. Hence the decrease in the costs of loss function 2.

Case 4. Resource Allocation vs. Benchmark Generation rate

In this experiment we varied the benchmark generation rate for Process 1, while keeping fixed the benchmark generation rates of Processes 2 and 3. By increasing the benchmark generation rate for Process 1, we made the process more efficient. The process is more efficient because the number of process cycles generated per unit of service provider and client resource increased.

The results of this experiment show that as the number of process cycles per unit input increases the service firm received more "bang for the buck", because fewer resources are needed to meet target output levels. Fig. 8 also shows when the benchmark generation rate=6, we are stealing from Processes 2 and 3 because we have hit our provider and client capacity constraints.

Case 5. Resource Allocation vs. Inefficiency Level

In this experiment we increased the inefficiency level, τ_{jip}, of Process 2 while the inefficiency level of other processes was held fixed at the base case value.

The results of this experiment show that as Process 2 becomes more inefficient, more resources are allocated to that process; see Fig. 9. When the inefficiency level ≤ 3, Process 2 is more efficient than Processes 1 and 3 and when the inefficiency level > 3 Process 2 is less efficient. As Process 2 becomes more inefficient, more resources are allocated to that process in order to minimize penalty costs.

In Fig. 10 we see that as Process 2 becomes more inefficient the penalty cost of not meeting output targets increases because it is harder for Process 2 to meet its target output level. Processes 1 and 3 penalty costs of not meeting output targets are equal and constant.

The convexity seen in Figs. 9 and 10 is due to the nonlinearity in the objective function. There is an increasing marginal investment in resource costs and loss as processes become more inefficient. The significance of an increasing marginal investment of average resource quantities and loss is that the service provider and the client should ensure processes are as efficient as possible in order to keep resource costs and loss low.

Case 6. Resource Allocation vs. Risk Level

In this experiment we increased the randomness (risk) of Process 1, while keeping fixed the randomness's of Processes 2 and 3.

The results of this experiment show that as the risk level of Process 1 is increased, more service provider and client resources were allocated to that process in order to keep penalty costs low. The resource allocations for Processes 2 and 3 were evenly allocated.

In Fig. 12 we see that as the risk (randomness) of Process 1 is increased, the penalty cost of not meeting output targets increases because it is harder for Process 1 to meet its target output level.
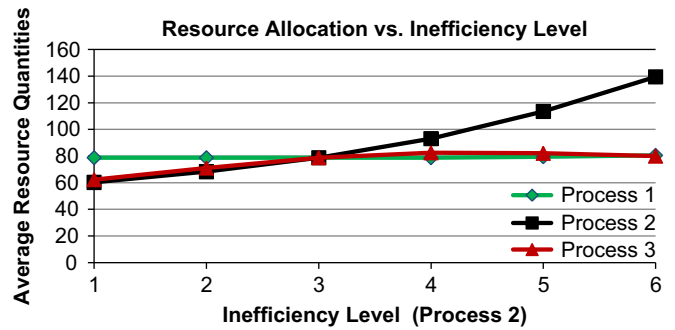

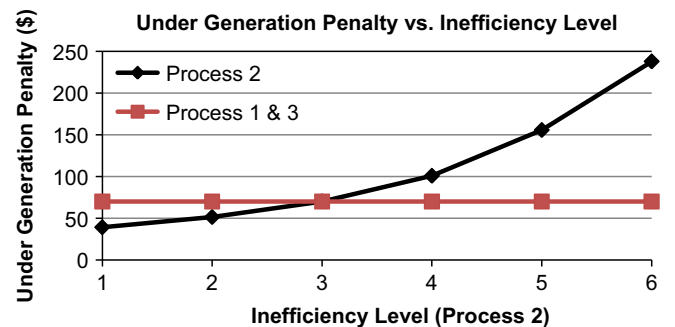**Fig. 9.** Resource allocation vs. inefficiency level.

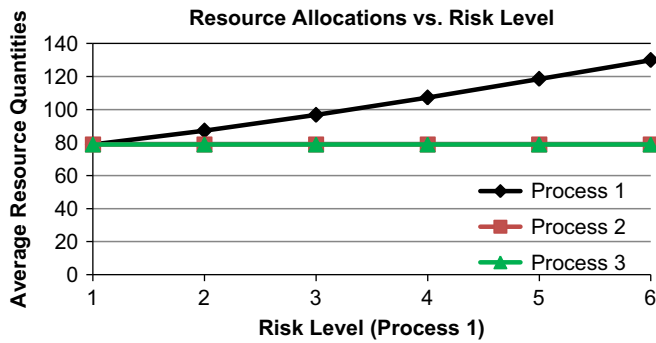**Fig. 10.** Under generation penalty vs. inefficiency level.

Resource Allocations vs. Risk Level

Fig. 11. Resource allocation vs. risk level.

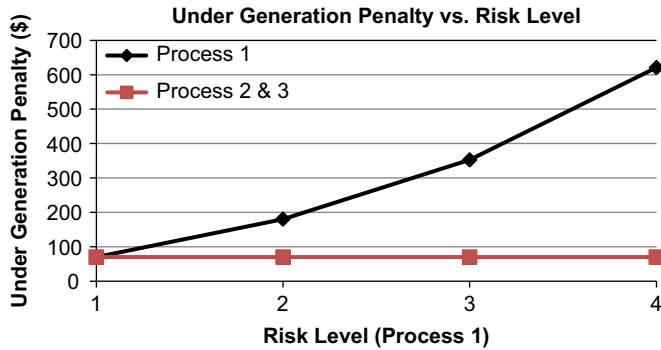## Under Generation Penalty vs. Risk Level

Fig. 12. Under generation penalty vs. risk level.

Processes 2 and 3 penalty costs of not meeting output targets are equal and constant.

The convexity seen in Figs. 11 and 12 is due to the nonlinearity in the objective function. There is an increasing marginal investment in average resource quantities and loss as process risk increases. The significance of this experiment is that the service provider and the client now have more insights into the effects of increased process risk on resource costs and loss.

6. Conclusion and directions for future research

This paper has examined the effects of uncertainties of the structure of the production function (mis-specification), uncertainties of the parameter values of the production function (mis-estimation), and the uncontrollable inputs on resource plans for a service engagement.

Two resource-planning models were developed. The first model is a deterministic, resource-planning model that applies a benchmark production function to obtain benchmark input levels for given target output quantities. We apply linear production functions in which the inputs must be procured according to usage rates of a process cycle and outputs are generated according to yield rates of a process cycle.

The second model is a stochastic, resource-planning model. This model receives benchmark input levels and target output levels as parameters from the first model and allocates resources at a minimal cost. In the objective function of this model, we incorporate a probability density function that captures the deviation from the benchmark output levels reflecting both inefficiencies and uncertainties.

We prove that in the presence of production function uncertainties, service firms will compensate by allocating resources so that penalty costs are minimized. The service provider and the client should put forth every effort to minimize uncertainties. Although there are uncontrollable inputs, efforts should be made to improve production function parameter inaccuracies and production function form specification.

In the future, we will examine the effects of production function uncertainties on a multi-period service supply chain. We will also extend this research using various distributions for representing inefficiency randomness. Additionally, it would be beneficial to the field of service research to perform a case study to add further validity to our model.

References

Athanassopoulos, A.D., 1998. Decision support for target-based resource allocation of public services in multiunit and multilevel systems. Management Science 44 (2), 173.

Badinelli, R., 2010. A stochastic model of resource allocation for service systems. Service Science 2 (1/2), 76.

Dietrich, B., 2006. Resource planning for business services. Communication of the ACM 49 (7), 62–64.

Frei, F.X., 2006. Customer-Introduced Variability in Service Operations. Harvard Business Review.

Golany, B., Hackman, S.T., Passy, U., 2006. An efficiency measurement framework for multi-stage production systems. Annals of Operations Research 145 (1), 51.

Golany, B., Tamir, E., 1995. Evaluating efficiency-effectiveness-equality trade-offs: a data envelopment analysis approach. Management Science 41 (7), 1172.

Land, K., Lovell, C.A.Knox, Thore, S., 1993. Chance-constrained data envelopment analysis. Managerial and Decision Economics 14 (6), 541–554.

Mula, J., Poler, R., Garcia-Sabater, J.D., Lario, F.C., 2006. Models for production planning under uncertainty: a review. International Journal of Production Econmonics 103 (1), 271–285.