MANAGING ACCURACY OF PROJECT DATA IN A DISTRIBUTED PROJECT SETTING

(Completed Research)

G. Shankaranarayanan
Babson College, USA
gshankar@babson.edu

Nitin R. Joglekar
Boston University, USA
joglekar@bu.edu

Edward G. Anderson Jr.
University of Texas-Austin, USA
edward.anderson@mccombs.utexas.edu

Abstract: Organizations (principals) manage projects by outsourcing tasks to partners. Coordinating and managing such projects requires sharing project-data, status data on the work-in-progress residing with the partners and estimates of completion time. Project data is rarely accurate due to errors in estimation, errors in aggregating data across partners and projects, and gaming by the partners. While managers are aware of the inaccuracies, they are forced to make decisions regarding outsourcing the tasks (how much, to whom, and when). In this paper, we develop a control theoretic model that analyzes utilization of capacity of both the principal and partners. This model also permits corruption of project-data regarding progress status. We use this model to compute the costs of using perfect project-data versus inaccurate project-data and show that these costs can be significant. We propose a control policy, using filters, to correct inaccurate project-data and generate an estimate of true progress status. We show that this policy, depending on the relative magnitude of variances in demand and measurement error, can minimize cost-penalties due to inaccurate project-data.

Key Words: Aggregate project planning; distributed projects; data quality; Kalman filtering; value of data.

INTRODUCTION

In today's global environment projects are typically carried out by multiple teams in distributed settings. There is a primary project management and coordination center (we will refer to this as the *principal*) and several partners (called *partners*) who may be geographically distributed. The principal will decide to outsource some tasks to one or more partners depending on the progress data (number of tasks the partner has in hold and time to complete each) available about each partner. If we extend this to a principal dealing with multiple partners and multiple projects, we can understand why the distribution of tasks is so critical to the success of the project and the economic welfare of the principal. The economic impact of distributing the load is significant [21]. However, such arrangements require that a majority of the task is routine [37], i.e., the majority of the work is well-specified and can be executed not only by the *principal* but also by one or more partners. The *principal* must however maintain a certain fraction of the work in-house to protect themselves against rising costs associated with *partners*. The *principal* must also manage the coordination and execution of the project tasks effectively [33].

A requirement for effectively managing and executing a project involving distributed tasks is the availability of perfect data about the project (such as progress, available resources and status-of-task). We use "project-data" to represent the data exchanged between the partners and principal for coordination and project execution. Literature on distributed projects assumes that perfect project-data about the progress made by the partners executing the tasks, is available. "Perfect" project-data is free of any error due to incorrect estimation of time to complete the work or errors in aggregating the data in information systems.

This assumption is critical for planning as the principal and partners use up-to-date data on project-progress to make resource-allocation decisions.

We present an example to illustrate decisions, project-data and quality issues in distributed project planning using the aggregate planning and oversight process in an insurance firm. Most product offerings in this firm are developed as software features such as graphical interfaces for the generation of insurance quotations and tools for assessment of damages due to natural disasters. Many of these features start as “bugs” in existing features. Planners can ignore bugs, fix them, and/or enhance the software by adding new features. New features are also added based on market needs and customer feedback. Regardless of the origin, each fix, enhancement, or new addition is bundled into a finite-duration project (project has a set of tasks, possibly with predecessor/successor relationships, and tasks can be assigned to partners).

A planning team (in the principal firm) assigns the tasks either to a software development group within the firm (core group within the principal) or outsources it to a different software development group that is or resides within *another firm* (a partner). The core group has a fixed number of personnel who work on several tasks concurrently and does not outsource any task. The partner has flexible capacity (and is referred to as the *flex* group in the figure below) as it can further outsources excessive workload. Some of the tasks, when finished by the partner, are transferred back to the core group because these tasks need to go through integration and/or testing processes. Figure 1 shows a time series of the work-in-progress (WIP)—measured in number of tasks—for the core group and partner (the flex group) over a period of two hundred days. The average size of WIP for the core group during the period is 1.7 (std.dev. = 0.61) tasks. The average WIP for the partner during the period is 7.4 (std. dev. = 2.80) tasks.

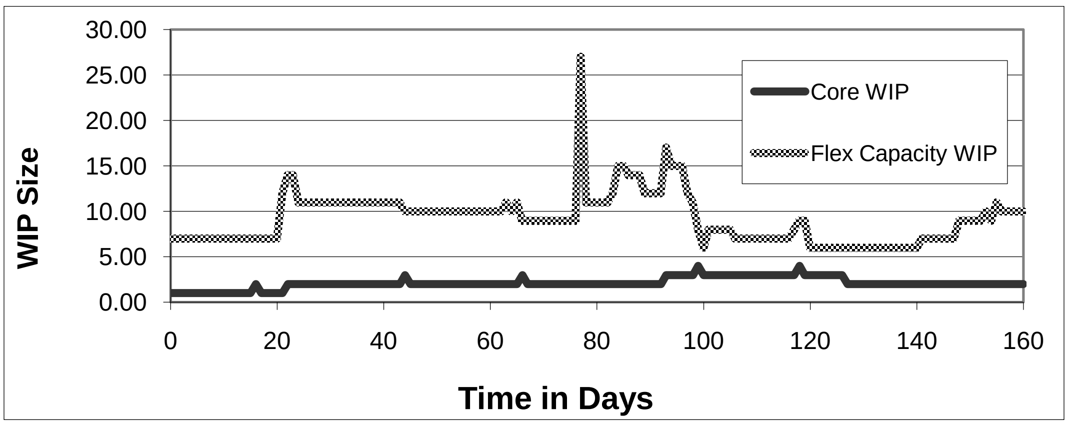

Figure 1: Amounts of WIP with Core group and Partner (Flex group)

We interviewed the principal’s planning team about their project assignment practices. The project managers to whom a task is assigned are primarily concerned completing the task as quickly as possible (minimize completion time). Alternately, the planning managers for the overall (aggregate) process are concerned about capacity utilization and smoothing the workload. These managers pointed out that they always work with a target WIP level. While the targeted WIP is good for smoothing workload and facilitating multitasking, these managers are also concerned with ensuring that the WIP does not get out of hand. Work done outside the principal, particularly when partners from a different outside firm are involved, requires a higher coordination and has the probability of requiring some rework, and hence—for this firm—*all of the project can never be outsourced in its entirety*. Another issue is highlighted by the nature of the data in Figure 1. Since each task’s completion time, *ex ante,* is a random distribution, the actual hours of work *ex post* required to complete the WIP tasks depicted will not be exactly proportional to the series depicted in the table. This fact draws attention to another difficulty in this planning and oversight process highlighted in standard texts on management of development-projects (see [23]): the uncertainty associated with the measurement of WIP and the poor quality project-data generated. While it is relatively easy to count the number of unfinished tasks, the planners cannot fully observe the progress of individual tasks, nor can they perfectly estimate any given task's required work-hours (i.e., how much more time is needed to finish a task in progress). Hence the status of the aggregate WIP at any point in time is, at best, estimated with uncertainty and, is hence, not accurate. However, this inaccurate project-data on WIP is used for assignment decisions of additional incoming work. This is the focus of this paper.

Anecdotal evidence from multiple industries also suggests that project data on aggregate progress are noisy (inaccurate) [9]. For instance, during our fieldwork at a large electronic product firm in Austin, the planning managers raised the issue of accuracy in the progress status data and posed the question: *how much should his firm be willing to spend to clean it?* Accordingly, the central questions explored in this article are: How does the knowledge of progress status affect optimal assignment policies for aggregate planning in distributed work settings? What is the cost impact of developing aggregate plans in such settings with imperfect data generated either from information system or estimation errors? Finally, can measures be taken to ameliorate the impact of imperfect information?

This paper makes several interesting contributions. *First*, we develop a control theoretic model to capture sourcing decisions for work in a set of distributed projects. The model allows us to manipulate project-data to simulate corruption owing to measurement errors. *Second*, we use the model to explore the effect of cost-structure asymmetry on the optimal division of tasks between a principal and its partners. This will illustrate that, with perfect project-data, such a decision structure will limit projects from over-running time/resource constraints. It will also show that project-plans are relatively smooth in their variation compared with random disturbances in demand and reintegration. Third, we determine the value of perfect versus imperfect progress/status data under the previously determined "optimal" control rule. Finally, we examine how filtering the data (and thus improving its accuracy) when it is inaccurate may reduce the aggregate planning cost penalties associated with imperfect data.

In the rest of the paper, we begin with an overview of relevant literature. We then present the model and derive the optimal control rule. We describe the numerical analysis using the model and discuss the implications of our findings for managing data quality of progress-data in distributed project settings. We conclude by presenting the limitations of our approach and point to the extensions needed to employ this model in real-world decision support systems to mitigate problems due to inaccurate project-data.

RELEVANT LITERATURE

Our research covers capacity planning in multi-project settings, sourcing decisions, managing data quality in distributed projects and value of information (or data). Given the context of this research, we will first briefly summarize the first two areas to differentiate this research from other similar work. We will then summarize the literature in data quality and information value to define the scope of this research and its contributions to data quality management.

The field of operations management has built up a body of analytical studies [29],[38] and empirical evidence on the potential benefits of flexible capacity in manufacturing [34] and R&D project settings [18]. The structure of our problem is similar to many queuing models that deploy control theory and address measurement noise (e.g., [14] [19]. While most queuing models minimize the waiting time, our model aims to minimize aggregate costs by smoothing capacity utilization around a targeted WIP value.

Our work builds on the body of knowledge in aggregate job-shop scheduling (e.g., [17]). It views partner capacity as a hedge against demand, completion and measurement uncertainties, and the need to make decisions using poor-quality data. We, however, are not focused on deriving the optimal capacity for the principal and a partner, but rather on the optimal switching of tasks between the capacities of principal and partners. Our model is structurally closest to the two-reservoir problem [29], except that we allow for the work to be completed by either the principal or the partner, each of which is subject to a different rework penalty. Further, some research that explore the dynamics of distributed projects use linear systems models. It is also customary to assume that optimal control policies operate around nominal WIP targets, with small deviations, such that the state variables always, or nearly always, remain positive [25]. However, our treatment differs slightly because we model the aggregate work of all tasks across multiple projects rather just those for any one individual project. The most important implication of this assumption is that our expected aggregate work-in-progress stock will remain constant and positive.

Literature in information economics differentiates between perfect, imperfect, and incomplete information (data) [36]. Perfect data pertains to a non-cooperative setting in which the player whose turn it is to play knows all the actions taken before her turn. With imperfect data, a player does not know what the other player has done so far. On the other hand, incomplete data describes the situation in which a player does not know the other player's precise characteristics. Our work here differs in an important respect, because it follows the dynamic control literature [3] by differentiating between perfect and imperfect data, not on the basis of the structure of a non-cooperative game, but rather on whether or not the data is created without error. Hence, under this definition of imperfect data, a firm cannot measure even the state of its own internal WIP progress without some uncertainty and without errors.

Data has value. Better data can reduce risk and opportunity cost, lead to better decisions and consequently lead to improved efficiency and performance [2]. Inter-organizational information exchange is an essential part of partnerships, particularly in distributed project management settings. In recent years, sharing analytical data has become a new business practice, enabled by the Internet and wireless technologies. In today's knowledge-economy, projects are executed in distributed settings, relying extensively on sharing data. Research has minimally examined data quality issues in settings where data is exchanged across departmental and organizational settings [30][6]. Even these propose high-level frameworks for managing quality. They do not address the economics associated with poor quality data.

Literature in data quality has addressed accuracy, a quality dimension, extensively. Although useful, approaches to improve accuracy such as data cleansing [15], data tracking and statistical process control [28], and data stewardship [10], do not mitigate the effects of inaccurate data and do not examine associated cost implications. In this paper, we propose a solution to correct (smooth out) inaccuracies. We compare the cost of making decisions with inaccurate data *and with corrected* data. We show this in the context of distributed project data and associated decisions.

THE MODEL AND ANALYSIS

Figure 2 models a distributed development system as two reservoirs with reflow. The principal has a constant capacity (number of severs or employees)—consistent with recent practices [22]. The demand for new development work feeds the principal's WIP. The principal can either keep and execute work internally ("in-sourcing") or outsource to a development partner for execution. The partner has a flexible capacity - it can change the number of its employees each period so that it can accept all work given and complete each after a constant flow time – consistent with current practices in software outsourcing. Both the principal and the partner can complete work; however, a fraction of the work completed by the partner must pass back to the principal for integration performed internally, inducing a reintegration penalty. This penalty is typical of real-world outsourcing of product and process development [27].

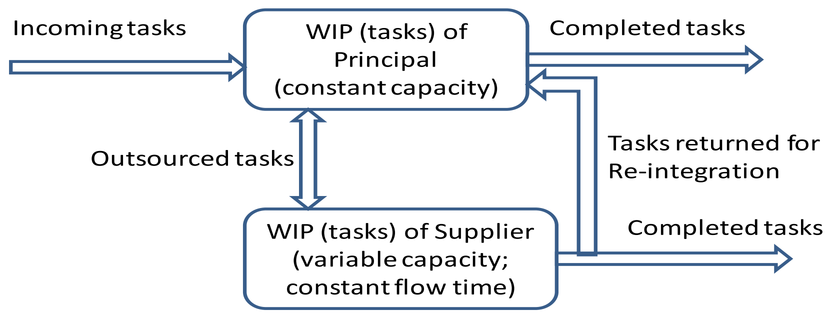

Figure 2: Base Model

Tracking Variations in WIP

To model this system, let $t \geq 0$ represent time and the index $d \geq 0$ represent the time period. Each period is of duration δ, i.e. the time period d begins at time $t = \delta d$ and ends at $t = (\delta+1)\,d$. Since we are modeling at the aggregate level across a large number of different projects, we model the arrival of each period's numerous tasks at the principal as a Poisson process (e.g., [7])—although this assumption can be relaxed without loss of generality. The task arrival process has an average inter-arrival time of α. The number of hours to complete each task is distributed randomly (with independent, identical distributions) with respect to the work required to complete them. Let i and o be indices that denote in-sourced (i.e. principal) and outsourced (i.e. partner) work. Hence, $x_{o,d}$ *and* $x_{i,d}$ will represent the number of work hours in progress stocks at the principal and partner, during period d, respectively. The size of each WIP may or may not be known perfectly (poor-quality project-data) due to how well each task's number of hours can be estimated [23]. We also assume that the flow of tasks outsourced from the principal to the partner is a weighted sum of the two WIPs (which, as we show later, is an optimal policy). We further assume that the completion rate of hours of work at the principal's WIP is constant because it maintains a constant number of employees, c_i. Each employee can complete a number of hours of work each period. There may be additional variation in the net completion rate due to various factors, such as administrative misallocation, distractions, stages in the project lifecycle, etc. that will be represented by a Gaussian random disturbance with constant mean and standard deviation. These variations in employee productivity per period capture distractions at work, holidays, administrative overhead, illnesses etc.

The partner, in contrast, can adjust its number of employees each period. That is, the number of partner employees is proportionate to the number of hours in its WIP x_o, i.e. $c_{o,d} = \kappa\delta x_{o,d}$. This implies that, on average, each unit of work spends an identical amount of time, i.e. κ^{-1} hours, in WIP prior to completion, as required. The partner also experiences a Gaussian random disturbance analogous to that at the principal, but with potentially different mean and standard deviation. The reintegration penalty is modeled as a flow of hours into the principal's WIP per period that is proportionate to the completion of work in the partner's WIP. Throughout this paper, we will also assume that, over time, the average completion rate of work at each WIP stock is equal to the average arrival rate of new and, for the principal, reintegration work. Without this assumption, the WIPs would either go to infinity or repeatedly approach zero. Both cases have been ruled out in practice by Morrice et al. [25]. Note, that in the firms that inspired our model, planning periods are typically a week in duration, e-mail in-boxes fill on a seemingly hourly basis with new tasks, and these tasks are indeed random. Given these time scales, it seems reasonable that tasks accumulate into a WIP stock and that the variation in net flow from this stock of WIP would be governed by the law of large numbers and result in a Gaussian distribution. We present two Lemmas to more formally capture this approximation. We have not included proofs for brevity.

Lemma 1 - Approximation of Net Changes in the WIP Stocks by Gaussian Distributions: Under the assumptions above, as the time interval becomes large relative to the inter-arrival time, i.e. $\delta/\alpha \rightarrow \infty$, the change in the number of hours of work in each stock per period will approach a Gaussian distribution.

Furthermore, that change is a linear function of the two WIPs $x_{i,d}$ *and* $x_{o,d}$ and the disturbances to the two completion rates. The distribution of the amount of stock at each WIP, $x_{i,d}$ *and* $x_{o,d}$, will also be asymptotically Gaussian. This lemma can be easily extended to any multi-state system in which each the change in each stock is a linear function of (1) other stocks in the system; (2) Gaussian disturbances; and (3) exogenous inflows whose inter-arrival times are short compared with the duration of each time period.

Lemma 2 - Asymptotic Approximation of Discrete Time System by a Continuous Time System: For a linear time invariant system consisting of stocks $x_{j,d}$ such that $x_{j,d} \in \{i, o\}$, let $a_{j,k}$ represent the portion of the change in x_j per period attributable to x_k, i.e. $x_{j,d+1} - x_{j,d} = \sum_k a_{j,k} x_{k,d} + \xi_j$ where ξ_j represents the sum of exogenous Gaussian disturbances and/or inflows at stock x_j. Let a_{max} represent the largest of these $a_{j,k}$. If $a_{\max}$ is small with respect to unity, then the distribution of any state variable $x_j(t)$, in the continuous approximation will approach that of the discrete-time state variable $x_{j,d}$ for $\delta d \le t < \delta(d+1)$ This lemma could be extended to an arbitrary number of stocks. For practical purposes, this lemma is appropriate as long as a_{max} does not exceed one quarter of δ, the duration of each time period [33]. For systems in which the rate of change in WIP and the actual size of WIP itself can be approximated by, respectively, a Gaussian white noise process and a continuous state variable with a Gaussian distribution, there are many control theory techniques for analyzing impact of imperfect information (e.g., [32]).

State Equations

Given Lemmas 1 and 2, we pose the outsourcing problem described in the previous section using a continuous time approximation as follows. Let:

r(t) = End-customer demand rate,

c = Completion rate at the principal, a constant,

λ = Fractional completion rate of WIP per unit time at the partner,

$w_i(t)$ = the net variation in the changes in stock of WIP at the principal, $w_i(t) \sim N(0, \sigma_i^2)$ and cov($w_i(t)$, $w_i(t+\varepsilon)$) = 0, for $\varepsilon \neq 0$,

$w_o(t)$ = the net variation in the changes in stock of WIP at the partner, $w_o(t) \sim N(0, \sigma_o^2)$. and cov($w_o(t)$, $w_o(t+\varepsilon)$) = 0, for $\varepsilon \neq 0$,

$x_i(t)$ = the work in progress, measured in number of hours of work, at the principal at time *t*,

$x_o(t)$ = the work in progress, measured in number of hours of work, at the partner at time *t*,

ϕ = The fraction of tasks completed at the partner requiring re-integration at the principal $(0 \le \phi < 1)$

u(t) = The number of tasks transferred from the principal to the partner at time *t*. Note that when $u(t)<0$, re-in-sourcing of previously outsourced tasks occurs.

The system is updated based on the following task balance relationships:

$$\frac{dx_i(t)}{dt} = r - c + w_i(t) - u(t) + \phi\left[\lambda x_o(t) + w_o(t)\right] \qquad (3.1)$$

$$\frac{dx_o(t)}{dt} = u(t) - \left[\lambda x_o(t) + w_o(t)\right] \qquad (3.2)$$

Equation (3.1) tracks the rate of change of the in-sourced work in process, whereas (3.2) tracks the rate of change of the outsourced work in process. Eq. 3.1 and 3.2 are linked via the decision variable *u(t)* and the amount of reintegration work needed determined by the outflow of rework, defined by the fraction ϕ.

Control with Perfect Data

To obtain a reasonable control policy for *u(t)*, we look forward from the present time $t = 0$ seeking to balance WIP, capacity, and capacity adjustment costs. Let:

$\beta_i^2(t)$ = penalty on insourced backlog (square is merely for analytic convenience), $\beta_{\mathrm{i}} > 0$,

$\beta_o^2(t)$ = penalty on outsourced backlog (square is merely for analytic convenience), $\beta_{\mathrm{o}} > 0$,

$\tilde{x}_i$ = Target insourced WIP, $\tilde{x}_i \geq 0$,

$\tilde{x}_o$ = Target outsourced WIP, $\tilde{x}_o \geq 0$,

J = The total penalty function.

As discussed in the illustrative example, the objective for aggregate planning is to smooth out the WIP, while accounting for dissimilar cost and integration penalties. Thus our objective is to:

$$\min_{u(t)} J = \lim_{t_f \to \infty} \frac{1}{t_f} \int_0^{t_f} \left\{ u^2(t) + \beta_i^2 \left[X_i(t) - \tilde{X}_i \right]^2 + \beta_o^2 \left[X_o(t) - \tilde{X}_o \right]^2 \right\} dt \qquad (3.3)$$

The first term reflects the quadratic cost associated with the control effort, and the second and third terms account for the quadratic costs associated with deviations from the target in-sourced and outsourced WIP. Fundamentally, the objective function in (3.3) performs an infinite horizon trade-off between smoothing each stage's work in process (WIP) and smoothing the task switching costs. In the operations literature, smoothing is often seen as an inherently desirable goal, e.g. Graves (1986) set up objective function to emphasize production smoothing [13]. Equation (3.3) quadratically weights the difference between actual WIP and a nonnegative target. The form of this penalty is *exactly analogous* to that of the HMMS model [16] or Sethi and Thompson's penalty for inventory fluctuations [29]. As in both models, the WIP target is typically greater than zero, albeit often small. In the models, the penalty for undershooting the WIP "target" existed to avoid idling workers through WIP starvation. Somewhat more complex, yet analogous issues apply to WIP in professional services. Because the opportunity cost of lost revenue relative to salary costs is on the order of a factor of twenty, generally development firms consider zero-WIP (which implies idling capacity) unacceptable. For these reasons, and because all the aforementioned costs are convex, we follow literature in penalizing deviations from "target" WIP levels in a quadratic manner [24].

For brevity, we assume that determining target WIP levels is done prior to assessing the penalty on the deviation from these targets. The cost of switching tasks is also quadratically weighted—as is typical in the capacity planning (e.g. [16], [29])—for a number of reasons. Moving tasks from one site to another often involves some sort of set-up cost [1]. Further, transferring five tasks per month from the principal to a partner is logically much simpler than fifty. Hence, the coordination problem tends to escalate in a convex manner, suggesting the quadratic form. In summary, the objective function is designed to balance task switching costs against service capacity utilization at both principal and partner.

Control With Inaccurate Data

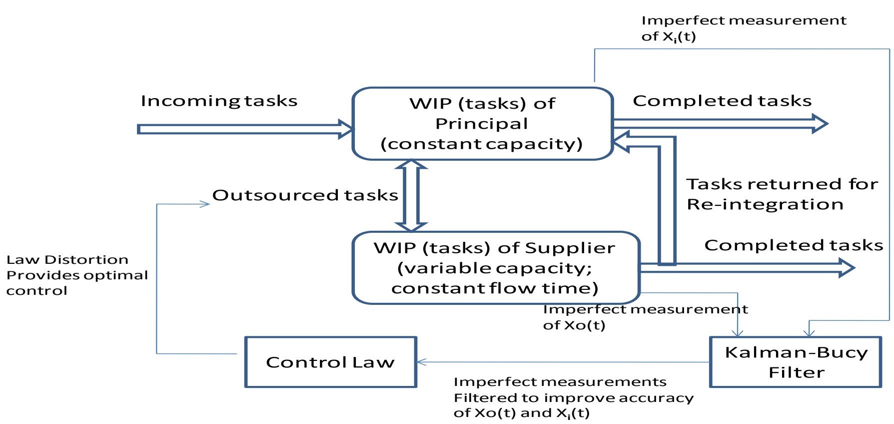

Figure 3: Closed loop control with inaccurate data

The model in figure 3 builds on that in figure 2 by allowing inaccuracies in project-data. The data are filtered through a mechanism to control the outsourcing decision using a control law derived assuming perfect data. The managers keep aggregate statistics (means and variance) on the estimation and measurement errors. These measurement errors are assumed to be independent, white noise processes.

Analysis – Deriving the Optimal Control Law

We derive the optimal control law based on the objective function in (3.3). We then select the filter characteristics that provide minimal cost penalty. For convenience, let:

$$\mathbf{x}(t)=\begin{bmatrix} x_i(t) \\ x_o(t) \end{bmatrix}, \mathbf{w}(t)=\begin{bmatrix} w_i(t) \\ w_o(t) \end{bmatrix}, \mathbf{z}(t)=\begin{bmatrix} z_i(t) \\ z_o(t) \end{bmatrix}, \mathbf{n}(t)=\begin{bmatrix} n_i(t) \\ n_o(t) \end{bmatrix}, \tag{4.1}$$

and

$$\mathbf{F}=\begin{bmatrix} 0 & \phi\lambda \\ 0 & -\lambda \end{bmatrix}, \mathbf{G}=\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \mathbf{Q}=\begin{bmatrix} \beta_i^2 & 0 \\ 0 & \beta_o^2 \end{bmatrix}, \mathbf{R}=[1], \mathbf{W}=\begin{bmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_o^2 \end{bmatrix}, \mathbf{N}=\begin{bmatrix} \nu_i^2 & 0 \\ 0 & \nu_o^2 \end{bmatrix} \tag{4.2}$$

Here $n_i(t)$ = the variation in the demand and completion rate at the lead firm. $n_i(t) \sim N(0, \upsilon_i^2)$ and corr($n_i(t)$, $n_i(t+\varepsilon)$) = 0, for $\varepsilon \neq 0$.

$n_o(t)$ = the variation demand and completion rate at the supplier. $n_o(t) \sim N(0, \upsilon_o^2)$ and corr($n_o(t)$, $n_o(t+\varepsilon)$) = 0, for $\varepsilon \neq 0$.

We write the evolution of the state variable, i.e. number of in-sourced and outsourced tasks, as:

$$\begin{bmatrix} x_i'(t) \\ x_o'(t) \end{bmatrix} = \mathbf{F}\left[\mathbf{x}(t)-\tilde{\mathbf{x}}(t)\right]+\mathbf{G}u(t)+\mathbf{w}(t) \tag{4.3}$$

If the information is imperfect due to measurement error, the observed state variables are **z**(t):

$$\mathbf{z}(t)=\mathbf{x}(t)+\mathbf{n}(t)\,. \tag{4.3b}$$

Theorem 1: The optimal control policy $u^*(t)$ with perfect measurement (perfect data) is given by:

$$u^*(t)=\mathbf{k}_c\left(\mathbf{x}(t)-\tilde{\mathbf{x}}\right) \text{ where } \mathbf{k}_c=\begin{bmatrix} \beta_i \\ \beta_i+\lambda-\sqrt{\beta_o^2+\beta_i^2+\lambda^2+2\lambda\beta_i(1-\phi)} \end{bmatrix}^T = [k_{c,in} \;\; k_{c,out}] \tag{4.4}$$

Superscript T represents a transpose operation.

An important managerial test that this policy must pass is whether it results in a "stable" system, i.e. neither of the WIPs can "run away" by growing without bound towards infinity under a finite demand. Such behavior is clearly undesirable. We would also like to determine whether or not the optimal policy creates oscillations over and above the disturbances inherent in the demand and reintegration rates themselves. Smith and Eppinger (1997) tested for this in their project WIP model by borrowing the concept of "over-damping" from control theory [31]. Specifically, "over-damped" indicates that a system's state variables do not overshoot their final values in response to a one-time increase in demand; whereas "under-damped" indicates that they do ([26], pp. 245). Hence, because any oscillations from the input disturbances are attenuated, the state variables in an over-damped system will be, in some sense, "smoother" than the input, whereas in an under-damped system, they will be "rougher." We now seek to determine whether the optimal policy creates a stable system and, if so, whether it is over or under-damped. In the discussion below, we have omitted proofs for brevity.

Corollary 1: The two Eigen-values for the system using the optimal control policy $u^*(t)$ are

$$\tfrac{1}{2}\left[-\sqrt{\beta_o^2+(\beta_i+\lambda)^2-2\beta_i\lambda\phi} \pm \sqrt{\beta_o^2+(\beta_i-\lambda)^2+2\beta_i\lambda\phi}\right] \tag{4.4b}$$

Proposition 1: The state variables under the optimal control policy are stable. That is any bounded demand will generate bounded completion rates. Furthermore, the control policy will create an over-damped system. Thus, runaway project WIPs are impossible under the optimal policy and furthermore the oscillation of the two WIP is, in some sense, strictly a result of demand and reintegration disturbances. The optimal policy does not increase that oscillation.

Proposition 2: The magnitude of the control law weight on *in-sourced* tasks increases in the cost β_i of deviating from the target in-sourced tasks. The magnitude of the control law weight on *outsourced* tasks increases in the cost β_i of deviating from the target in-sourced tasks and the rate at which outsourced tasks are completed λ. It decreases in cost β_o of exceeding the target outsourced tasks and the reintegration fraction ϕ of outsourced tasks.

Lemma 3 - Certainty Equivalence Property of Linear Quadratic Gaussian Filters: The same control law $\mathbf{k}_c$ that minimizes the objective function *J(t)* when applied to the state vector $\mathbf{x}$(t) in the presence of *perfect* state information will also, when applied to $\hat{\mathbf{x}}(t)$, the minimum mean square error estimate of $\mathbf{x}$(t), minimize the objective function *J(t)* in the presence of *imperfect* state information.

The minimum mean square estimate of $\mathbf{x}$(t) is $\hat{\mathbf{x}}(t)=\left[\hat{\mathbf{x}}_i(t) \quad \hat{\mathbf{x}}_o(t)\right]^T$.

Theorem 2: Evolution of minimum mean square estimate of $\mathbf{x}$(t) is governed by a Kalman-Bucy filter

$$\hat{\mathbf{x}}'(t)=\mathbf{F}\left[\hat{\mathbf{x}}(t)-\tilde{\mathbf{x}}(t)\right]+\mathbf{G}u(t)+\mathbf{K}_e\left[\mathbf{z}(t)-\hat{\mathbf{x}}(t)\right] \tag{4.5}$$

where $\mathbf{K}_e=\mathbf{PN}^{-1}$ such that $\mathbf{PN}^{-1}\mathbf{P}=\mathbf{FP}+\mathbf{PF}^T+\mathbf{W}$ and $\mathbf{P}$ is positive definite (4.6)

Corollary 2: If the fraction of reintegration work as a result of outsourcing $\phi = 0$, then

$$\mathbf{K}_e=\begin{bmatrix} \dfrac{\sigma_i}{\nu_i} & 0 \\ 0 & -\lambda+\sqrt{\lambda^2+\dfrac{\sigma_o^2}{\nu_o^2}} \end{bmatrix} \tag{4.7}$$

Corollary 3: Assume that the fractional rate of outsourced task reintegration $\phi = 0$. Then the four Eigen-values for the system using the optimal control policy *u*(t)* in Corollary 2 are:

$$\tfrac{1}{2}\left[-\sqrt{\beta_o^2+\left(\beta_i+\lambda\right)^2}\pm\sqrt{\beta_o^2+\left(\beta_i-\lambda\right)^2}\right],-\frac{\sigma_i}{\nu_i}\ \text{and}\ -\sqrt{\frac{\sigma_o^2}{\nu_o^2}+\lambda^2} \tag{4.8}$$

Proposition 3: Assume that the fractional rate of outsourced task reintegration $\phi = 0$. Then, the rate at which the estimate of in-sourced tasks $\hat{x}_i(t)$ is updated increases in the variance of the demand and in-sourced completion rates σ_i and decreases in the variance of the measurement error of in-sourced tasks ν_i. The rate at which the estimate of outsourced tasks $\hat{x}_o(t)$ is updated increases in the variance of the outsourced completion rates σ_o and decreases in the fractional outsourced WIP completion rate λ and the variance of the measurement error of in-sourced tasks ν_i.

Proposition 4: Assume that the fractional rate of outsourced task reintegration $\phi = 0$. Under imperfect information, the state variables under a combination of the optimal control policy and the optimal Kalman-Bucy filter are stable. That is, any bounded demand will generate bounded completion rates. Furthermore, the control policy is guaranteed to create an over-damped system only when the reintegration rework fraction $\phi = 0$. (Note: from numerical studies, if the re-integration rework fraction $\phi > 0$, then the system may be under-damped.) Hence, there can be no "runaway" portfolio of projects or tasks at an aggregate level even with imperfect information. However, if outsourced work requires integration, the WIPS may no longer necessarily be smoother than their input disturbances.

Theorem 3: The penalty function for the case with perfect information is

$$J_p=\mathbf{SW}. \tag{4.9}$$

Where $\mathbf{S} = \dfrac{1}{\lambda(\phi-1)} \begin{bmatrix} \beta_i\left(\beta_i - \tilde{\beta}\right) & \beta_i\left[\beta_i + \lambda(1-\phi) - \tilde{\beta}\right] \\ \beta_i\left[\beta_i + \lambda(1-\phi) - \tilde{\beta}\right] & (\beta_i+\lambda)^2 - \lambda\phi(2\beta_i+\lambda) - \left[\beta_i + \lambda(1-\phi)\right]\tilde{\beta} \end{bmatrix}$ (4.9a)

And where $\tilde{\beta} = \sqrt{\beta_o^2 + \beta_i^2 + \lambda^2 + 2\lambda\beta_i(1-\phi)}$

The penalty function for the case with imperfect information under the Kalman-Bucy filter and control is:
$J_k = J_{ce} + J_s$ where $J_{ce} = \mathbf{SK}_e(\mathbf{N}+\mathbf{P})\mathbf{K}_e^T$ and $J_s = \mathbf{QP}$. (4.10)

The penalty function for the case with imperfect information (and $\phi = 0$) when the control law $u^*(t)$ is used directly upon the corrupted measurements $\mathbf{z}$ as if they were perfect measurements of the state variables $\mathbf{x}$ is

$$J_m = J_p + J_n \quad \text{where} \quad J_n = \frac{\left[(k_{ci}-k_{co})^2 + k_{ci}\lambda\right]\left(k_{ci}^2 v_i^2 + k_{co}^2 v_o^2\right)}{2(k_{ci}-k_{co}+\lambda)}. \qquad (4.11)$$

MODEL INSTANTIATION

To explore the behavior of the system to a range of plausible inputs, we present a series of numerical analyses using the control law and Kalman-Bucy filter derived in the previous section. The performance with perfect data is discussed first and that with inaccurate data is discussed later. The perfect data results serve as a benchmark for exploring the efficacy of the filter.

Performance With Perfect Data

The numerical studies for a system with perfect information are conducted in the following sequence. The response of the system with asymmetric costs at the principal and partners is explored initially, and the response of the system to changes in the rework fraction ϕ and the task completion rate at partners λ is presented afterwards. (For ease of exposition, these costs are assumed to be symmetric in the second set of studies.) Three types of performance measures are included in both these studies:

i. The optimal control policy $u(t)$ coefficients of $\mathbf{k_c}$ on in-sourced and outsourced WIP, which for the ease of description shall be termed as $k_{c,in}$ and $k_{c,out}$ respectively.
ii. Decay rates (δ_1 and δ_2) for the natural response of the system (these values are computed based the eigen-values in equation 4.4b)
iii. The penalty function for the case with perfect data J_p (based on equation 4.9)

Table 1 illustrates the impact of an asymmetric cost structure, assuming that the rework fractions ($\phi = 0.5$) and the task completion rate at partners ($\lambda = 1$) are kept constant. Standard deviation of disturbances to the demand and completion rates at each stage are $\sigma_i = 1$, $\sigma_o = 1$.

Table 1: Impact of Cost Structure on Optimal Assignment

β_i	β_o	$k_{c,in}$	$k_{c,out}$	δ_1	δ_2	J_p
1.00	**0.10**	-1.00	0.63	0.36	1.37	1.50
1.00	**1.00**	-1.00	0.00	0.29	1.71	2.44
1.00	**10.00**	-1.00	-0.85	0.05	10.10	31.01

The variation in the values of the control law coefficient $k_{c,out}$ is logical given the model. When the principal's internal WIP deviation penalty is larger than the partner's ($\beta_i > \beta_o$), a larger fraction of tasks from the principal will be moved to (or from) the partner (i.e. $k_{c,in} < k_{c,out}$), which in effect transfers more WIP oscillation to the partner**.** In contrast, when the internal WIP deviation penalty is *smaller* than the partner's, the reverse is true. We conjecture that the first case ($k_{c,in} < k_{c,out}$) will prevail in practice because most organizations, with which we have contact, seem to outsource in order to minimize variation in the utilization of their own resources, but we include the second for completeness. We also note that the two coefficients, based on our choice of control parameters ($\phi = 0.5$ and $\lambda = 1$), show some asymmetry.

The decay rates δ_1 and δ_2 are a measure of system stability. The numerical values are inversely related to the square root of the negative of the eigen-values in Corollaries 1 and 3. Runaway work in process occurs whenever either value of $\boldsymbol{\delta}$ is negative. The values of J_p in the table confirm Proposition 1 in that the positive real values of all the decay rates in a system indicate that it is both stable and over-damped, i.e. for these parameter sets, the optimal policy will create neither runaway WIP's nor increase the oscillation of the WIPs above that of demand. Moreover, the system will react more quickly to changes in demand as the magnitudes of the $\boldsymbol{\delta}$'s increase. Additionally, while the decay rates will vary based on the value of control parameters, one can see that the cost structure will also affect the rate of task completion. Finally, consistent with modeling intuition, outsourcing reduces the overall cost penalty J_p when the deviating from internal WIP targets costs more than deviating from those at the partner.

Table 2: Impact of Partner Characteristics on Optimal Assignment

λ	ϕ	$k_{c,in}$	$k_{c,out}$	δ_1	Δ_2	J_p
0.10	**0.00**	-1.00	0.39	0.07	1.42	9.18
1.00	**0.00**	-1.00	0.24	0.62	1.62	1.71
10.00	**0.00**	-1.00	0.05	0.99	10.05	1.05
0.10	**0.50**	-1.00	0.35	0.04	1.42	17.5
1.00	**0.50**	-1.00	0.00	0.29	1.71	3.00
10.00	**0.50**	-1.00	-0.42	0.50	10.09	2.42

In the next set of tests, the effect of partner characteristics are explored by assuming that the cost structure is symmetric and constant as are the disturbances (($\beta_i = \beta_o = 1.0$; $\sigma_i = \sigma_o = 1$), while the rework fraction ϕ and the task completion rate at partners λ vary in a systematic manner as shown in Table 2. The numerical values of the control law coefficients, $k_{c,in}$ and $k_{c,out}$ are behave reasonably. When the fractional task completion rate at partner, λ, increases, more oscillation in work is transferred to the partners. Similarly, when the reintegration penalty ϕ rises, less oscillation in work is transferred to the partner. Again, all the $\boldsymbol{\delta}$ values are positive, which corroborates that the system is over-damped and stable. Finally, the task completion rate at partners λ reduces the overall penalty cost J_p, and the reintegration penalty ϕ increases it. We have explored, but not shown interaction effects between these tables in the interest of saving space. However, data from Tables 1 and 2 should suffice to show that the system cost structure and the characteristics of partners affect performance measures. In the next sections we will assume that the cost and system characteristics are invariant and appropriate values of overall penalty costs J_p will be used as a benchmark for studying the performance with imperfect information in the next section.

Performance with Inaccurate Data

Figure 4 illustrates the system behavior with imperfect data considering two policies. One is to use the optimal control law from *Theorem 1* assuming that the measurement is perfect, even though it is not. The other is to assume the data is inaccurate and use a filter, a Kalman-Bucy filter, to obtain improved estimates (more accurate data) of the number of tasks in the in-sourced and outsourced WIP stocks.

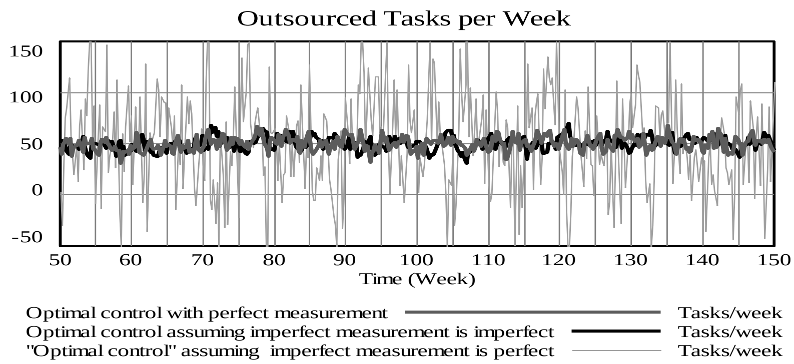

Figure 4: Evolution of Outsourcing Policy in due to Variation in Incoming Tasks
$\beta_I=1, \beta_O=2, \sigma_I=\sigma_O=10, \nu_I=\nu_O=20, \phi=0,$ **and** $\lambda=0.2$

Table 3: Performance under Varying Amounts of Noise

ν_i	ν_o	J_p	J_m	J_{CE}	J_S	J_K
0.01	**0.01**	1.71	1.71	171.12	0.02	171.14
0.10	**0.10**	1.71	1.72	17.48	0.19	17.63
0.10	**1.00**	1.71	1.79	13.71	0.51	14.22
1.00	**0.10**	1.71	2.94	6.36	1.09	7.45
1.00	**1.00**	1.71	3.01	2.59	1.41	4.00
1.00	**10.00**	1.71	9.83	2.47	1.00	3.97
10.00	**1.00**	1.71	125.38	1.50	10.00	11.89
10.10	**10.00**	1.71	132.20	1.36	10.52	11.86

Notice that the deployment of the optimal control policy, while using the filtered data, results in a response that tracks the system response with perfect data closely. On the other hand, using the optimal control law without filtering the imperfectly measured data leads to a much larger variance in the number of outsourced tasks per week. The results are consistent with intuition—absence of erroneous estimates when using an optimal policy attenuates task switching, because the error terms are denominators in the filter multipliers (see equation 4.7). In Table 3, we explore the contribution of individual system parameters to this attenuation process. Control parameters for these simulations have been set at $\beta_I=\beta_O=1$, $\omega_I=\omega_O=1$, $\phi=0$, and $\lambda=1$. Three performance measures are explored for varying levels of measurement (ν_i and ν_o):

i. J_p : penalty function for the case with perfect data
ii. J_m: penalty for optimal control assuming that the inaccurate data is perfect.
iii. J_k : penalty for optimal control assuming that the inaccurate data is imperfect and hence has filtered before use. There are two components of this penalty: $J_k = J_{CE} + J_S$. Here J_{CE} is the penalty associated with the estimate error and J_S is the penalty associated with the measurement error.

These results show that inaccurate data always creates a performance penalty. In the most benign conditions (for high signal to noise ratios, $\omega/\nu \to 0$), and assuming that the optimal control policy is known and but measurements are not filtered, the performance degradation from J_p to J_m is not significant. As the error-level increases, however, the performance of the system without using the filter gets progressively worse.

When the variances in measurement errors are much smaller than those in random disturbances to demand and to task-completion rates, the Kalman-Bucy filter results in a decidedly worse performance,

J_k, It improves rapidly and becomes very significant as the measurement errors increase. The behavior of Kalman-Bucy type data filters, depending on system characteristics, has been known to exhibit singularities at very high signal to noise ratios. In such situations enhancements to the derived filter have been advocated [12]. On the other hand, even the currently filter can improve performance in noisy environments. Further, these data show an interesting trade-off between J_{ce} versus J_s as function of the "signal-to-noise" ratio ω/v. Recall that J_s corresponds to the penalty associated with the measurement error, as shown by the co-variation term $\mathbf{P}$ in (4.10). J_{ce}, on the other hand, is linked with the certainty equivalence, and its value is dependent on the filter gain matrix $\mathbf{K}_e$. Equation 4.7 confirms that the filter gains $\mathbf{K_e}$ are inversely related to the noise level. Hence, the J_{ce}-J_s trade-off can be explained as follows: at high signal to noise ratio, the penalty due to J_s is minimal; however, because the signal drowns out the noise and the filter gains are very large, the system becomes highly sensitive to small perturbations. With a low signal-to-noise ratio, in contrast, the measurement errors create the need for more adjustments and consequently, increase the overall cost penalty.

DISCUSSIONS AND CONCLUSION

We discuss the managerial implications of our results by commenting on the extant practices for collecting aggregate project-data on progress-status and the implications of our results for data quality management in distributed projects. Collecting precise progress status data (project-data) is not trivial. Browning et al. (2002) discuss several reasons in the context of developing an Uninhibited Combat Aircraft Vehicle (UCAV) to illustrate the lack of precision in monitoring the progress status. They argue that this inaccuracy leads to a lag between value creation and value determination [5]. We account for these difficulties while selecting key parameters that are needed to set up our model: the costs structure parameters β_i and β_o, the completion rate at the partner site (λ). Users may estimate these parameters using a regression model adopted by [25]. Similarly, the noise (error) associated with the progress data can be estimated by mining progress status data [4]. Managers are aware of the inaccuracies in progress status data. The reader will recall from the sensitivity analyses presented in Table 3 that the performance penalties associated with poor data quality rise with the level of noise in the progress status data.

There are three potential sources of errors in data associated with distributed projects: technical limitations within information processing systems, data collection errors, and gaming.

(a) Information system errors: Information systems are not error-free. However, we assume that only a fraction of information system errors can be attributed to computational errors. A much larger fraction of the errors may be due to exchanging data across multiple systems, combining data from multiple sources and combining data from legacy applications.

(b) Data collection errors: The second set of concerns revolves around an organization's ability to collect project-data in a timely manner. Some attention has been paid in the information systems literature, for instance the need to understand the value of information and frequency of updates in a supply chain management context [8]. This leads to project data being collected only on a periodic basis and their aggregation remains a point of concern [1]. It is quite difficult to nail down the amount of measurement error generated by such processes. Moreover, in some instances, the distributed nature of product development work creates what has been described as "project oscillatory behavior" ([24], [39]) over and above what has been discussed in this paper. Such oscillations can compound the data quality problems.

(c) Gaming: It is not a surprise that a big concern for data quality is the possibility of gaming of project-data. Ford and Sterman (2003) report that progress status data are routinely misreported under what they term as the "liar's poker syndrome," wherein the nature of reintegration burdens makes it rational for managers to provide erroneous data [11]. Of course, savvy project managers typically adjust for this by discounting such reports by a fixed “fudge factor.” However, the likelihood that the “fudge factor” exactly balances out under-reporting at every reporting period remains low.

In this paper, we develop a model for managing projects in a distributed setting. In the presence of inaccuracies in demand and progress status data, we explore the impact of cost structure asymmetry on optimal division of tasks between the principal and its partners. Our analysis shows that with perfect data about progress status, such a system will not create runaway project-WIP and, in fact, will smooth it. Further, we determine the value of obtaining perfect progress status data and the value of filtering the data when data is inaccurate. These ideas are particularly interesting in light of our results that show significant degradation in performance due to the relative differences in the noise levels for the principal and the partner.

There are several limitations to this work. One is our choice of a quadratic cost structure, which yields a linear control rule. Some managers may want to run their operations under linear control because such rules are easy to understand. However, such rules are typically not used for systems with extremely variable demand or significant peak capacity utilization or idling. Another concern is the performance of the filter under conditions of a high signal-to-noise ratio. Other practical questions include: exactly how should a manager obtain the data to set up the control law and the filter? And how much should one pay, either within the lead firm or to the supplier, for a metering process to reduce the level of noise in the progress status data below a specified threshold? These questions suggest that the data quality in distributed project setting offers a rich source for research and investigation.

REFERENCES

[1]. Anderson, E. G., and Joglekar, N. R. 2005. A Hierarchical Product Development Planning Framework. *Production and Operations Management* **14**(2).

[2]. Banker, R. D. and Kaufmann R. J. 2004. The Evolution of Research on Information Systems: A Fiftieth-Year Survey of the Literature in Management Science. *Management Science* **50**(3) 281-289.

[3]. Bertsekas, D. 2001. *Dynamic Programming and Optimal Control, 2nd edition.* Athena Scientific: Belmont, MA.

[4]. Braha, D. 2001. *Data Mining for Design and Manufacturing: Methods and Applications*, Kluwer Academic Publication, Norwell, MA.

[5]. Browning, T. R, Devst, J. J., Eppinger, S. D. and Whitney, D. E. 2002. Adding Value in Product Development by Creating Information and Reducing Risk. *IEEE Transactions on Engineering Management* **49**(4) 443-458.

[6]. Cai, Y. and Shankaranarayanan, G. (2007) *Managing Data Quality in Inter-Organizational Data Networks,* International Journal of Information Quality, Vol. 1, No. 3, 2007, pp. 254-271.

[7]. Carrascosa, M., Eppinger, S. D. and Whitney, D. E. 1998.Using the Design Structure Matrix to Estimate Time to Market in a Product Development Process. *Proceedings of ASME Design Theory and Methodology Conference* Atlanta 98-6013.

[8]. DeHoratius, N. 2004. In Pursuit of Information Quality. *Cutter IT Journal.*

[9]. Dominguez, A. 2006. Project Management in Noisy Environments. *Proc. of POM Annual Meeting*, Boston.

[10]. English, L. P. (1999). *Improving Data Warehouse and Business Information Quality- Methods for Reducing Costs and Increasing Profits*. New York, NY: John Wiley & Sons, Inc.

[11]. Ford, D. and Sterman J. 2003. The Liar's Club: Impacts of concealment in concurrent development projects. *Concurrent Engineering Research and Applications*.

[12]. Gerlach, H., Dahlhaus, D., Pesce, M. and Xu, W. 2004.Constrained Chip-Level Kalman Filter Equalization for the UMTS-FDD Downlink. *Proc. World Wireless Congress (WWC).*

[13]. Graves, S.1986. A Tactical Planning Model for a Job Shop. *Operations Research* **34** 522-533.

[14]. Harrison, J.M. and Wein, L. M. 1990. Scheduling Networks of Queues: Heavy Traffic Analysis of a Two-Station Closed Queues *Operations Research* **38**(6) 1052-1064.

[15]. Hernadez, M. A., & Stolfo, S. J. (1998). Real-world Data is Dirty: Data Cleansing and the Merge/Purge Problem. *Journal of Data Mining and Knowledge Discovery, 1*(2)

[16]. Holt, C. C., Modigliani, F., Muth, J. F. and Simon, H.A. 1960. *Planning Production, Inventories, and Work Force*, Prentice-Hall, Englewood Cliffs, NJ.
[17]. Khouja, M. and Goyal, S. 2008. A review of the joint replenishment problem literature: 1989–2005 European Journal of Operational Research **186 (1)** 1-16
[18]. Krishnan, V. and Bhattacharya, S. 2002. Technology Selection and Commitment: The Role of Uncertainty and Design Flexibility. *Management Science* **48** (3) 313-327.
[19]. Kumar, S. and Kumar. P. R. 1996. Fluctuation Smoothing Policies are Stable for Stochastic Reentrant Lines. *Discrete Event Dynamic Systems* **6** 361-370.
[20]. Lee, H. L., So, K. C. and Tang, C. S. 2000. The Value of Information Sharing in a Two-Level Supply Chain. *Man. Sci.* **46** 626-643.
[21]. Mandel, M. and Engardio, P. 2007. The Real Cost Of Off-shoring Business Week June 18.
[22]. Mcgrath, M. 2001. How to Boost R&D Productivity by 50%? *Insights Magazine*, www.prtm.com, Summer/Fall Issue.
[23]. Meredith, J.R. and Mantel, S. J. 2005. *Project Management: A Managerial Approach.* New York: Wiley.
[24]. Mihm, J., Loch, C. H. and Huchzermeier, A.. 2003. Problem -Solving Oscillations in Complex Projects. *Management Science* **49**(6) 733-750.
[25]. Morrice, D. J., Anderson, E. G. and Bharadwaj, S. 2004. A Simulation Study to assess the efficacy of linear Control Theory Models For the coordination of a Two-Stage Customized Service Supply Chain. *Proc. of Winter Simulation Conference*, R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds).
[26]. Oppenheim, A., Willsky, A. S. and Young, I.T. 1983. Signals and Systems. Englewood Cliffs, NJ: Pren.-Hall.
[27]. Parker and Anderson. 2003. From Buyer to Integrator: The Transformation of the Supply Chain Manager in the Vertically Disintegrating Firm. *Production and Operations Management* **11** (1): 75-91.
[28]. Redman, T. C. (1996) "Data Quality for the Information Age," Boston, MA: Artech House
[29]. Sethi, S. P. and Thompson, G. L. 2000. *Optimal Control Theory: Applications to Management Science and Economics.* Boston, MA: Kluwer Academic Publishers.
[30]. Shankaranarayanan, G., Ziad, M. and Wang, R. Y. (2003) "Managing Data Quality in Dynamic Decision Environment: An Information Product Approach", *Journal of Database Management*, Vol. 14, 14-32.
[31]. Smith, R. and Eppinger, S. D. 1997. Identifying Controlling Features of Engineering Design Iteration. *Management Science* **43** (3) 276-293.
[32]. Stengel, R. F. 1994. *Optimal Control and Estimation*. New York: Dover.
[33]. Sterman, J. D. 1989. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. Man. Sci. **35** 321-339.
[34]. Suarez, F. F., Cusumano, M. A. and Fine., C. H. 1996. An Empirical Study of Manufacturing Flexibility in Printed Circuit Board Assembly. *Operations Research* **44**(1) 223-240.
[35]. Tam, P. and Range, J. 2007. Second Thoughts: Some in Silicon Valley Begin to Sour on India. The Wall Street Journal.
[36]. Tirole, J. 1993. *The Theory of Indutrial Organization*. MIT Press, Cambridge, MA.
[37]. Upton, D. M. and Staats, B. R. 2006. Lean at Wipro Technologies. HBS Case 9-607-032.Ulrich, K.T., S.D. Eppinger 2000. Product Design and Development. McGraw-Hill, New York.
[38]. Van Mieghem, J. A. 2003. Capacity Management, Investment and Hedging: Review and Recent Developments. *Manufacturing & Service Operations Management* 5(4): 269-302.
[39]. Yassine, A. Joglekar, N. R., Braha, D., Eppinger, S. D. and Whitney, D. 2003. Information Hiding in Product Development: The Design Churn Effect. *Research in Engineering Design* **14** 145-161.