Dynamics of Project Screening in a Product Development Pipeline

Paulo Figueiredo and Nitin Joglekar*
Boston University School of Management
595 Commonwealth Avenue
Boston MA 02215

Abstract
Management of a product development pipeline involves starting and steering several promising projects through a sequence of screens known as stages/gates. Only projects with payoffs above a predetermined threshold survive each screen. We model a two-stage product development pipeline as an aging chain with a co-flow. The co-flow structure tracks the number of projects *and* the corresponding net present value (NPV) of payoff. Managers at each stage must decide on capacity utilization, subject to a trade-off between throughput and value creation rate. Our simulation study mimics a range of relevant decision scenarios by varying the number of starts, screen thresholds, and managerial biases while adjusting utilization. Results illustrate that screening can eliminate the backlog bullwhip effect in the pipeline. Allied statistical analysis indicates a non-linear relationship between the number of starts and the value created at end of the pipeline. An increase in the screening threshold, in either stage, increases the average value of the projects but reduces the total value created. We also show that a managerial bias towards reducing backlog, instead of improving utilization, affects the average NPV negatively but does not affect the total value created at the end of the pipeline.
Keywords: Bullwhip, Decision Biases, Product Development, Screening, Stage/Gate

* Please direct correspondence to: joglekar@bu.edu

1. Introduction

The importance of improving product development performance cannot be understated in a competitive business environment (Wheelwright and Clark 1992, Griffin 1997). Almost all product development organizations put their new product development (NPD) projects through a series of screens (a.k.a. stages/gates) so that only the best performing projects are released into the market place (Cooper *et al*. 1998). The term product pipeline management (PPM) alludes to the practice of starting and steering several promising projects through this sequence of screens. For example, Girotra *et al.* (2005) document development processes for pharmaceutical drugs that follow a number of well defined stages/gates. In each stage, potential value is created through development tasks, information is gathered to evaluate the technical adequacy of the project and the future market performance of the end product is forecasted. In each of the corresponding gates, managers examine projected technical and market performance and use that data to assign net present value (NPV) to the projects, and then determine whether to proceed or terminate a fraction of the projects based on a predetermined threshold. The goal of this paper is to provide policy choices for PPM decision making by analyzing the underlying dynamics and determining which screening strategy allows a higher performance. For instance, when is a policy of starting *and* terminating many projects superior to starting and rejecting few projects?

The structure of stocks and flows in PPM can be compared to the structure of a service supply chain model (Anderson *et al*. 2005) as shown in figure 1 and 2. In both situations, the processing flow-time and capacity constraints determine the throughput. Anderson *et al*. have shown that, depending on the relative magnitudes of the processing time and the capacity adjustment time, the service supply chain exhibits a backlog bullwhip, i.e. swings in stage 2 backlogs are larger than swings in stage 1 backlogs in response to perturbations in the customer demand.

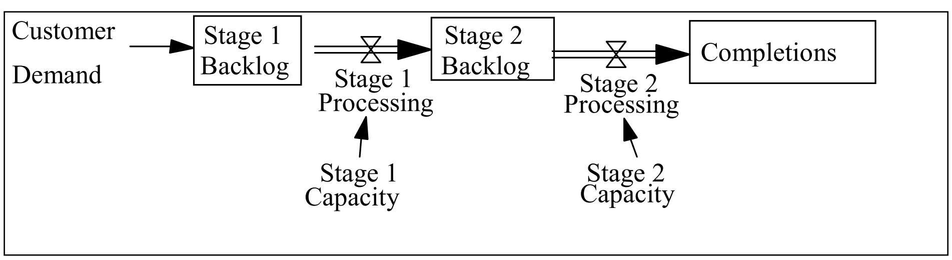


Figure 1: A Service supply chain (Anderson *et al*. 2005)

The PPM problem is a special case of service supply chains where some projects are terminated across stages based on their NPV. We study the behavior of the PPM problem by formulating a System Dynamics model that tracks the number of projects and their NPV at each stage within a co-flow structure. Such a model allows us to explore the following questions: is such a structure susceptible to backlog bullwhip? What is the sensitivity of total NPV created vis-à-vis to the screening thresholds across successive stages? Given these thresholds and assuming fixed resources, how should the capacity be utilized? Also, how does a manager's bias towards reducing backlog instead of improving NPV, a proxy for quality, affect the value created at the end of the pipeline?

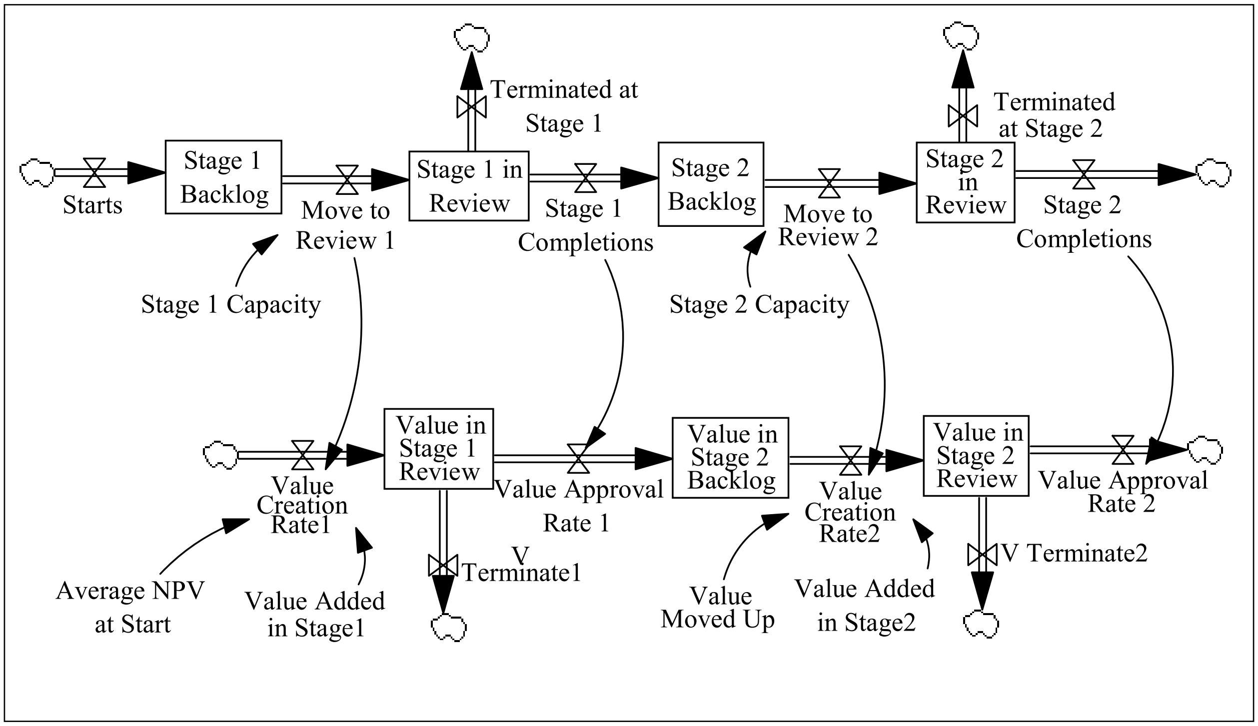


Figure 2: A Multi Stage Product Development Pipeline

We explore the effects of relevant managerial decisions on the development pipeline performance by varying the number of starts, screening thresholds and the biases in adjusting capacity. Our simulation results show that the PPM model behavior can differ remarkably from conventional service supply chains. For instance, we illustrate that the PPM process can eliminate the backlog bullwhip. Statistical analysis of these data illustrates that there is a non-linear relationship between the number of starts and *both* the average and the total NPV created at the end of the pipeline. An increase in the screen threshold, in either stage, increases the average value of the projects but reduces the total value created. We also show that a managerial bias towards reducing backlog affects the average NPV negatively but does not affect the total value created. Implications of these findings for improving PPM management policies are discussed in section 6.

2. Research Setup

An established body of literature characterizes product pipeline decisions as a dynamic problem that is often beset with congestion effects (Griffin 1997, Ulrich and Eppinger 2004). For example, Adler *et al.* (1995) modeled the project development organization by setting engineering resources as “workstations” and projects as “jobs” that flow between the workstations. At any given time, a job is either receiving service or queuing for access to a resource. The authors investigated the development performance, as measured by the cycle time. At any one stage of the pipeline, PPM decisions can be studied as a portfolio management problem. For instance, Banerjee and Hopp (2001) formulated a stochastic dynamic problem where limited resources must be allocated among a set of candidate projects over time so as to maximize expected net present value. The optimal solution is to create an index policy, so that projects are sequenced according to a simple ratio and then resources are allocated up to each project’s practical limit in the order given by this sequence. Building on the Banerjee and Hopp formulation, Gino and Pisano (2005) have taken a behavioral approach to this problem and explored the application of such policies to test heuristics for resource allocation across multiple stages of a pharmaceutical R&D process.

NPD managers are often endowed with limited resources. However, their focus is not limited to efficient resource allocation under these situations. They are also interested in the trade-offs between value creation and throughput involving the product pipeline management decisions. A similar type of trade-off between quality and throughput has been studied in other settings such as the service industry. For instance, Oliva and Sterman (2001) identify "time per order," as a key construct that drives the service quality dynamics in a single stage model calibrated for a lending center at a UK bank. The applicability of service quality trade-offs has not been explored across an entire service supply chain and/or in service profit chains (Heskett *et al.* 1997). However, capacity utilization has been identified as a key construct that drives the performance of service supply chains (Anderson and Morrice 2005).

A relevant assumption for our study is that the relationship between capacity utilization and value created in each gate has an inverted U shape, with the peak value being observed at nominal value of utilization. We set the nominal value of capacity utilization to be at unity (as shown in the appendix.).This assumption follows field observations by Wheelwright and Clark (1992, pg. 91) and by Girotra *at al.* (2005). The latter authors have pointed out that total development costs can be thought of as the sum of opportunity costs and the cost of capacity, resulting in a convex function of capacity utilization. There exists an interior utilization level that maximizes firm profit. Utilization therefore affects the dynamics across multiple stages in a product development pipeline: after introducing a very large number of projects, utilization goes up above its nominal value and reduces the relative amount of NPV added for each project. Utilization goes down when a stage is starved of input projects, and that too reduces the relative amount of NPV added to each project. Hence we hypothesize that a negative quadratic relationship will exist between the number of starts and both the average NPV and the total value created. Total NPV is the product of average NPV and the number of projects approved. We also hypothesize that the linear term of the relation between number of projects started and the performance variables will be positive, because utilization on stage 2 tends to be lower due to the experimental conditions and the elimination of projects in the first stage. A low utilization means that starting more projects will increase value creation (see U-shaped curve in the appendix).

Hypothesis 1: *An increase in the number of projects started increases the average NPV of the projects but increases total NPV created at the end of the pipe line*.

Hypothesis 2: *There is a negative quadratic relationship between number of projects started and both the average NPV and the total NPV created. As the squared term of number of projects increases, average NPV project and total NPV decrease at the end of the pipeline.*

In a product development pipeline, projects are either approved or taken to the next stage or they are terminated. Managers set thresholds (or minimum acceptable NPV values) in order to decide which projects are going to be approved. We argue that by setting a higher threshold, more projects will be terminated and their value lost, but the average value of the remaining projects should be higher (Dahan and Mendelson 2001). On the other hand, the total value created will be adversely affected.

Hypothesis 3: *An increase in the thresholds for minimum acceptable NPV in either stage will increase the average value of the projects, but reduce the total value created.*

In a product development pipeline, the available capacity of the development teams is adjusted in order to either adapt to the work demand of each stage of the chain with or to keep the utilization level around its nominal value. If more weight is given to the objective of achieving the fastest rate instead of the nominal utilization level, the increase in the value of the projects as they pass through the gates should be proportionally smaller, because the capacity utilization will be above or below its nominal levels (Girotra *et al.* 2005).

Hypothesis 4: An *increase in the bias towards reducing backlogs through the adjustment of capacity will decrease the average value of the projects, and vice-versa.*

3. Model Description

Most firms use multiple, typically four to six, gates in their pipelines (Ulrich and Eppinger 2004). For parsimony, our model incorporates only two gates as shown in figure 2. Outcome variables of interest are the total value created, and the average value created at the end of the pipeline. The independent variables in our model are number of projects introduced into the pipeline, minimum acceptable NPV in each stage (thresholds 1 and 2), and the managerial bias in adjusting capacity. The model structure is comprised of three processes: capacity management, screening and value creation. These are described next.

3.1 Capacity Management Process

A central construct of the model is the utilization of capacity. Figure 3 displays the structure that captures the decision process for adjusting capacity. As pointed out in the previous section, research shows that employee productivity (percent of time spent on *value-adding* tasks) initially increases and then decreases as the number of development projects assigned concurrently to each engineer increases (Wheelwright and Clark 1992, pg. 91). We capture this effect in a table function that links utilization and value created. Allied details are provided in the appendix.

Managers have at each stage a fixed amount of resources (employees). An increase in capacity is only possible by using the existing resources more intensively, there by increasing their utilization. Utilization is calculated according to equation 1. In case of overcapacity, the utilization equals to the demanded capacity based on the backlog.

$$\text{Utilization} = \frac{\text{MIN}(\frac{\text{Stage Backlog}}{\text{Nominal Dev Time}}, \text{Available Capacity})}{\text{Nominal Capacity}} \tag{1}$$

Capacity is adjusted continuously, depending on the value of the target capacity and on the time to adjust capacity. We define target capacity as the weighted average of the nominal capacity (a capacity that yields the peak NPV) and the demanded rate of development in each gate based on the backlog.

$$\text{Change in Capacity} = (\frac{\text{Target Capacity - Available Capacity)}}{\text{Time to Adjust Capacity}} \tag{2}$$

$$\frac{\text{d(Available Capacity}}{\text{dt}}) = \text{Change in Capacity} \tag{3}$$

$$\text{Target Capacity} = \frac{\alpha * \text{Stage Backlog}}{\text{Nominal Dev Time}} + (1 - \alpha) * (2 - \text{Utilization}) * \text{Nominal Capacity} \tag{4}$$

Here α is the manager's bias towards reducing backlog ($0 < \alpha < 1$).

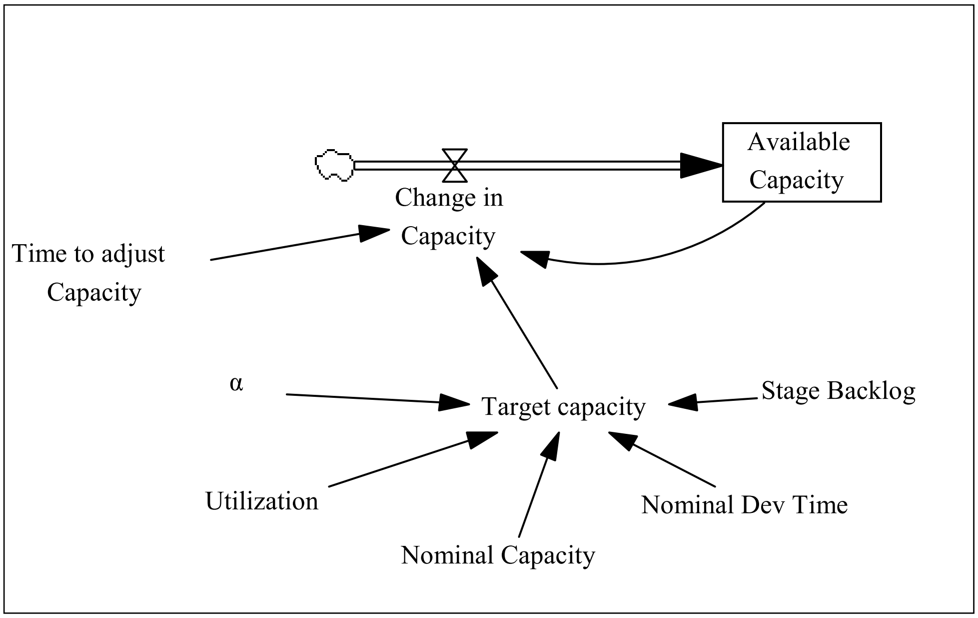


Figure 3: Procedure to Adjust Capacity

3.2 Value Creation Process

The available capacity derived from equation 3 is used within each stage as shown in figure 4 during the process of value creation. A certain number of projects enter stage 1 backlog. The co-flow stocks track the value of the projects along with their number. The average NPV of projects is normalized to unity at start. This value is subsequently multiplied by a factor ranging from 1.35 to 2, depending on the utilization, as the projects that were in the backlog are developed and go to the next phase to be reviewed (see equation 5). The rate “move to review” is equal to the available capacity, unless there is overcapacity (see equation 6). The projects then reach gate 1, or “stage 1 in review”. In this phase projects are reviewed, and depending on the average NPV (see §3.3 for details), some fraction will be terminated and the rest will “follow the flow” to the next stage, the backlog of stage 2.

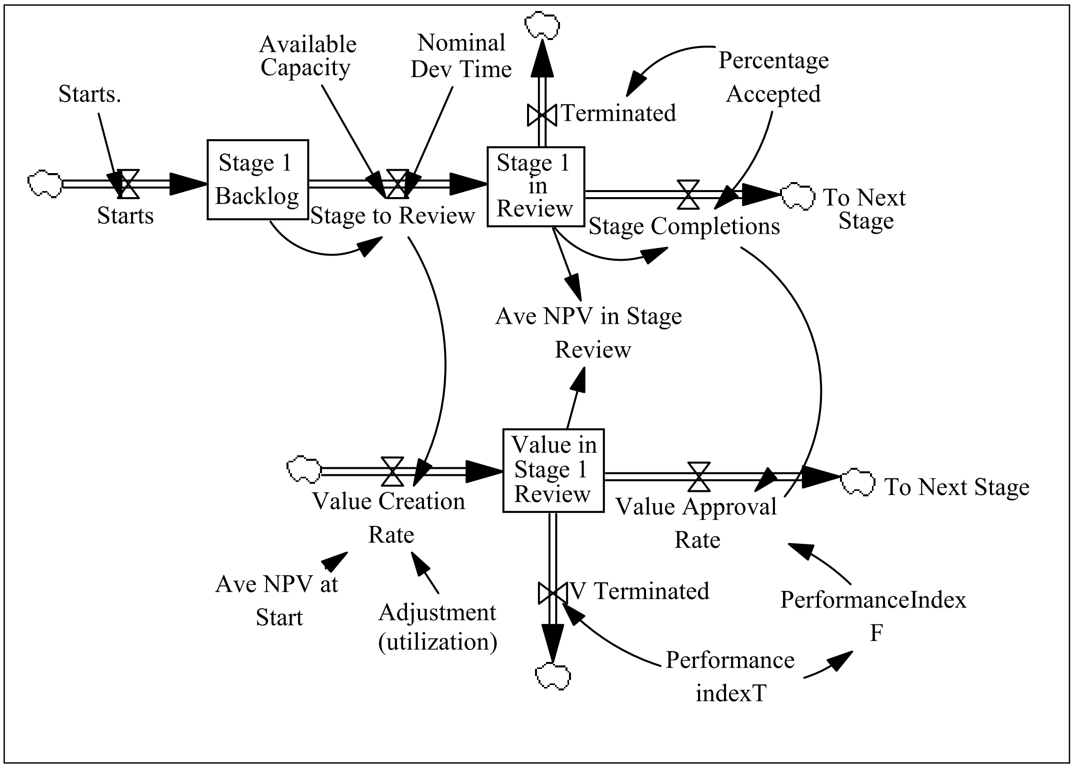


Figure 4 : Stock and Flow Structure of a Typical Gate

$$\text{Value Creation Rate} = \text{Average NPV at Start} * (1 + \text{Adjustment(Utilization)}) * \text{Move to Review} \quad (5)$$

$$\text{Move to Review} = \text{MIN}(\frac{\text{Stage Backlog}}{\text{Nominal Dev Time}}, \text{Available Capacity}) \quad (6)$$

Adjustment is a table function described in the appendix. Equations 7, 8 and 9 represent the rate of change in the stocks of stage backlog, stage in review, and value in stage review. Projects that are approved in the second phase are launched to the market. The values of total NPV created, number of projects and average NPV of finished projects are tracked and used as performance measures. These calculations have been simplified by assuming that the time discounting effect is built into the Ave NPV at Start parameter.

$$\frac{d(\text{Stage Backlog})}{dt} = \text{Starts - Move to Review} \quad (7)$$

$$\frac{d(\text{Stage in Review})}{dt} = \text{Move to Review - Stage Completions} \quad (8)$$

$$\frac{d(\text{Value in Stage Review})}{dt} = \text{Value Creation Rate} - \text{VTerminate} - \text{Value Aproval Rate} \quad (9)$$

Performance indices, T and F, are defined in the appendix.

3.3 Project Screening Process

The average NPV of the projects feeds into the screening process: the decision to proceed or terminate a fraction of project is made depending on the average NPV and a predetermined threshold. The population of NPVs of projects after a review is assumed to follow a Gumbel distribution, because project screening is a search process that selects NPV extreme values (Gumbel 1958, Galambos 1978, Dahan and Mendelson 2001). The Gumbel distribution is the probability distribution for the maximum of multiple draws from exponential-tailed distributions. It applies to NPD problems especially well when there are no specific limits on the potential NPV of a project (Dahan and Mendelson 2001). Appendix A provides a summary of the Gumbel distribution and the formulation of percentage terminated/ accepted, Performance IndexT and Performance IndexF. The latter two are the corrections to the changes in NPV stocks based on percentage accepted, as shown in figure 4.

4. Model Behavior

We have conducted a series of tests to build confidence in the model structure and behavior (Forrester and Senge 1980). Initially, we set number of starts as a step function that introduces 5 projects and step time 12, and chose the “base case” settings (see table 1 and appendix A for details) in order to analyze the behavior of selected variables: the rate of projects that go into and through the pipeline and the backlogs at each stage. Figure 5 illustrates the outcome of this simulation: the number of starts per period and the number of projects completed per period, in stage 1 and 2. Stage 2 completion rate is slightly below stage 1 completion rate, as expected, because some of the projects are terminated during stage 1.

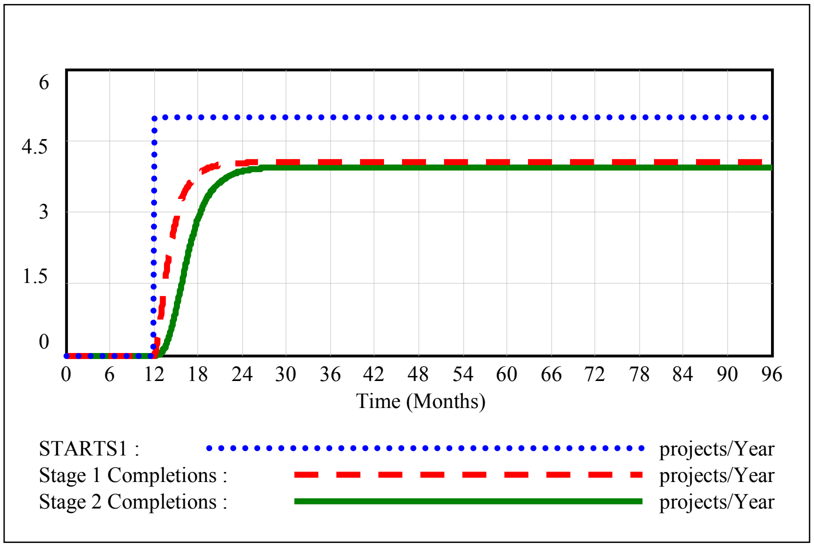


Figure 5: Starts and Completion Rates versus Elapsed Time

Next, we describe the effect of high number of starts, with respect to the nominal development capacity, on the backlogs. We set the number of starts at 7.5 projects per month, and a policy bias towards improving capacity utilization ($\alpha 1 = \alpha 2 = 0.05$). In effect, the pipeline capacity is blocking some of the projects. We term this situation an *untouched backlog* at the front end (figure 6).

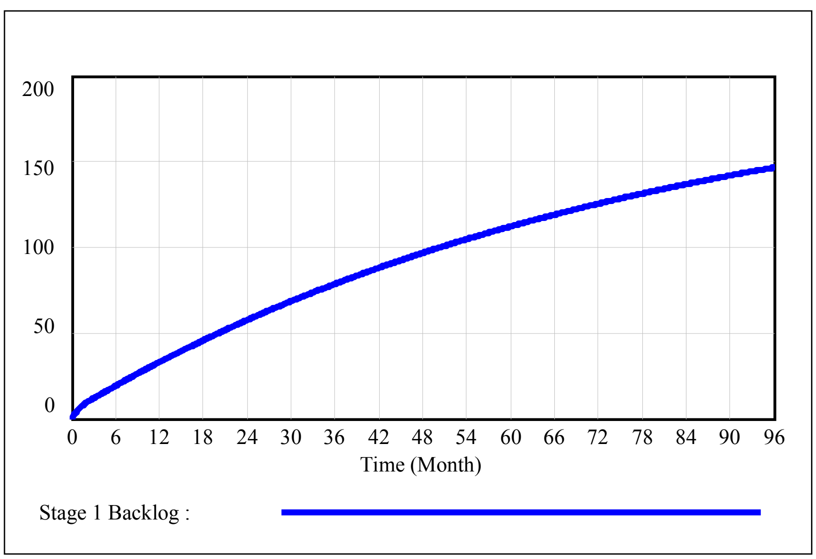


Figure 6: Build up of Untouched Backlog at the Front of the Development Pipeline

In order to test for the presence of a bullwhip effect, we set parameters to their extreme conditions: both thresholds are set to their lowest settings (threshold1=threshold2=0) so that all projects are approved. We set the managerial bias towards extreme values for reducing the backlog ($\alpha 1 = \alpha 2 = 0.99$) and introduce a constant number of starts (5 projects per month) beginning with the twelfth month. We also allow for the conditions that enable a backlog bullwhip effect (Anderson and Morrice 2005) by setting up the time to adjust capacity at 8 months and the nominal development time at 2 months. The result is presented on the left hand side of figure 7. Indeed, a backlog bullwhip effect is observed. The right hand side of figure 7 shows the model's behavior with the same parameters, but with screens in place (threshold1=1.68 and threshold2=3.09). Screening eliminates the backlog bullwhip.

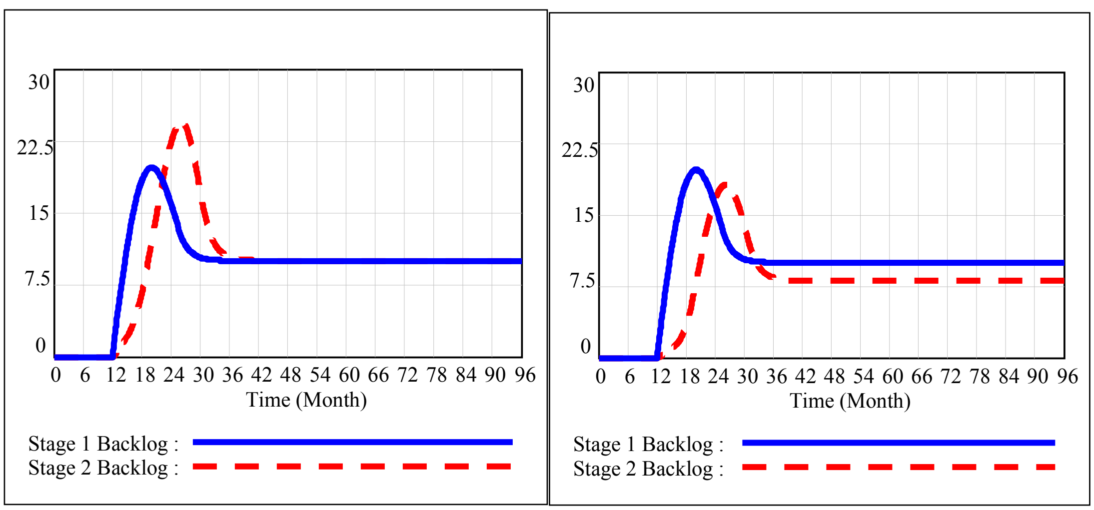


Figure 7: Effect of Screening on the Backlog Bullwhip

5. Statistical Analysis

Our hypotheses establish a direct relationship between the performance variables (average NPV and Total NPV created) and the number of projects introduced into the pipeline, threshold in stage 1, threshold in stage 2 and policy regarding the calculation of target capacity (captured by the parameter α). In order to test our hypotheses, we employ an experimental design consisting of simulating three levels for each model parameter as shown in Table1. The values of the other parameters remain constant (as shown in Appendix A). Number of starts per month is a assumed to be a random variable, normally distributed with mean λ (set at 2.5, 5 or 7.5) and standard deviation σ=0.83. [1]

Table 1: Independent Variable in the Experimental Design

Factor Level	**Low**	**Medium**	**High**
Number of Starts (λ)	2.5	5	7.5
Threshold1	1.51	1.68	1.84
Threshold2	2.77	3.09	3.39
α1	0.05	0.5	0.95
α2	0.05	0.5	0.95

These parameters have been selected to span the entire range of settings for the capacity adjustment policy (α), target thresholds and number of projects introduced into the pipeline (Starts), while satisfying the constraints imposed by the table function used to calculate NPV creation rate based on utilization. In each case, we select levels for threshold 2 to be larger than threshold 1 because projects in stage 2 will be expected to have higher average NPV than in stage 1. The range of average gain in NPV is set at 68% (ranging from 35% to 100%). Similarly, we set medium threshold for stage 2 as 3.09. These settings, and their relative values, are consistent with recent empirical studies that document thresholds for different stages of the product development pipeline (Schmidt *et al.* 2006). The high and low conditions are calculated by adding or subtracting 50% to the starts and 10% to the threshold. The nominal capacity is set to 5 projects per month. These parameters ensure that we sample each segment of the utilization curve presented in the appendix.

[1] We have also tested the model for other values of σ, but do not report these settings because their results are materially identical.

We ran a total of 105 simulations, based on a 3X5X7 design, setting 5 combinations for the values of threshold 1 and 2, and 7 combinations for the values of $\alpha1$ and $\alpha2$ (see table 2). Each simulation ran for 96 months with the first 12 months truncated in order to eliminate initialization effects (Law and Kelton 2000).

Table 2: Experimental Design

Variables	Combinations (L=Low, M=Medium, H=High)
Starts	H, M, L
Threshold 1 and Threshold 2	MM, MH, ML, HM, LM
$\alpha1$ and α 2	MM, MH, HM, HH, ML, LM, LL

Table 3 and 4 contain the regression results for the 2 dependent variables, average NPV and total value created, as a function of the independent factors listed in table 1. We also included a squared term of the "starts" variable as an independent variable. The regression model was constructed using SPSS statistical software and a stepwise regression approach in which the regression model is built progressively by adding variables to the model. The first factor to enter the model has the highest R^2 and subsequent factors enter the model in case they provide an increment in the R^2 regression statistic.

Table 3: Results with Average NPV as Dependent Variable (N=105, Adj R^2=0.931).

Independent Variables	Coefficient	Standard Error	T	P value
(Constant)	1.288	.120	10.763	.000
Starts (λ)	.388	.017	22.462	.000
StartsSquared (λ^2)	-.032	.002	-18.661	.000
Threshold2	.394	.026	15.367	.000
Threshold1	.283	.048	5.875	.000
$\alpha1$	-.038	.015	-2.541	.013

Table 4: Results with Total NPV as Dependent Variable (N=105, Adj R^2=0.976)

Independent Variables	Coefficient	Standard Error	T	P value
(Constant)	1331.360	192.597	6.913	.000

Starts (λ)	937.211	27.830	33.676	.000
StartsSquared (λ^2)	-70.988	2.754	-25.772	.000
Threshold1	-1321.863	77.747	-17.002	.000
Threshold2	-221.383	41.380	-5.350	.000

Hypothesis 1: We have hypothesized that an increase in the number of starts increases the total value created and the average value of the projects. This hypothesis is supported. The number of starts has a positive effect on both average NPV and total Value Created. These results can be explained by the fact that utilization on stage 2 was governed by queuing physics and did not achieve high values (e.g. even when number of starts was set to 7.5 projects per month, the average utilization on stage 2 was 0.89).

Hypothesis 2 and 3: The negative coefficient on the "Starts Squared" construct in both models confirm our second hypothesis: there is a negative quadratic relationship between number of projects started and the variables average value per project and total value created. As the squared term of number of projects increases, average NPV and total value created decrease. The third hypothesis was also confirmed: both thresholds have a positive effect on average value of projects, and negative effect on total value created.

Hypothesis 4: This hypothesis was partially supported. Only $\alpha 1$ (i.e. manager's bias towards reducing backlog in stage 1) entered the final model, and its coefficient has a negative value as expected. The bias toward reducing backlog (increasing α) will decrease the average value of the projects. However, Variable $\alpha 2$ did not enter the first model and both α's did not enter the second model.[2]

[2] There is multicollinearity between the starts and starts squared (VIF=49), however this is an endemic problem in regression models that contain linear and squared terms of the same variable. However, we argue that a potential collinearity problem can be neglected, since the usual consequences of multicollinearity (i.e. overall significance of regression without significance of individual coefficients) are not present (Deeds and Rothaermel 2003). Rather, the individual coefficients are significant. Therefore,

6. Discussion

For the parameter set used in our tests, we have shown that an increase in the thresholds in either stage will increase the average value of the projects, but reduce the total value created. Our results show that if the number of projects introduced into the pipeline is relatively high when compared to the nominal capacity, a policy bias towards improving capacity utilization will result in an accumulation of projects in the first stage backlog. A policy bias in the other direction, towards reducing backlog, will decrease average NPV and also increase the number of projects terminated. These findings, along with additional simulations to adjust for *in situ* parameters, can be used to design screening, nominal capacity and utilization policies in a real project.

Although there are many studies that describe stage/gate processes (Griffin 1997), there are few guidelines available on how to set the target screen levels in a product development pipeline. The structure of the screening problem is like a queue with built in quality adjustment mechanisms. In such a queue, the front end can block the next stage. NPD literature recognizes the importance of the "fuzzy front end" (Khurana and Rosenthal 1997) and recommends a front loading strategy (Thomke and Fujimoto 2000) for a development process. These NPD studies have been empirical in their orientation and based on their observations, their authors argue for the need to focus on the front end owing either to the high amount of uncertainty or the ability to generate early information. Consistent with these NPD findings, *but based on pipeline management and throughput-quality considerations*, our analysis also finds that management of "front end" is more critical than subsequent stages in governing the overall pipeline performance.

---

any existing multicollinearity did not cause a type II error as it potentially can. Moreover, any existing multicollinearity does not bias the estimates (Greene 1997).

Moreover, our model provides a mechanism to compute the elasticity of the outcomes with respect to intermediate thresholds and managerial choices.

The manner in which our model has been set up differs from inventory/ service supply chain models (Sterman 1989, Anderson and Morrice 2005) both in terms of stock/flow and policy structures. One structural difference is that inventory and service supply chain models do not have exit flows (aka screens). Also, in our model, the procedure to adjust capacity is different from the one suggested in the literature. For instance, in the Anderson and Morrice (2005) study service chain managers consider not only the size of the backlog, but also the end-customer demand *and* the local demand while making their capacity decisions. The nominal capacity is not taken into account and there is no adjustment for utilization. Thus, for the set of parameters under which we have conducted our analyses, the PPM model does not exhibit any bullwhip when the screens are in place. However, under extreme conditions (Forrester and Senge 1980), when the screens are eliminated (thresholds=0), together with a bias towards reducing backlog and a specific choice of nominal development times and time to adjust capacity, the model reverts back to a bullwhip effect similar to that of a service supply chain.

Ours is a highly stylized model that comes with several limitations. For instance, we do not account for dependencies among projects, such as sharing of resources and sub-additive pay-offs. It is known that shared resources affect the relationship between development costs and project development time (Girotra *et al*. 2005). Sub-additive pay-offs occur when a firm launches many products that are related (such as derivatives of a product family), but could generate the same revenues if it had developed only one product. Another limitation of this formulation is that the number of employees is fixed, therefore an increase in capacity is automatically translated into an increase in utilization. In our model, the fixed resources are evenly distributed among stages (nominal_capacity_1 = nominal_capacity_2 = 5), however in a many situations, these resources may be allocated in a centralized manner according to the necessity of each stage. The total capacity could then be shared by both stages unevenly. This constitutes one possible improvement to the model.

The limitation of managers' ability to account for the supply line and backlogs has been documented extensively in the inventory/services management context (Sterman 1989, Anderson and Morrice 2005). A related avenue for research, within the product innovation context, is to generate policy guidelines about the dynamics of capacity, resource utilization and backlog management while accounting for behavioral biases related to product innovation (Schmidt and Calantone 2002, Gino and Pisano 2005). Admitting behavioral biases raises a number of pertinent questions: will managers overweigh information that comes from companies or projects whose names/themes are salient (availability bias)? Do managers react to financial success/failures of previous products by increasing/reducing the amount of projects entering the pipeline? Will managers set different requirements for successive NPV thresholds depending on the degree of innovativeness of their portfolio?

In conclusion, PPM is a well established business process within the NPD community. However, extant literature aimed at managerial insights in this realm has been based on descriptive and empirical analyses. We offer the PPM model as a complementary tool for providing simulation based insights into the dynamics of project screening in a product development pipeline.

7. References

Adler, P. S., Mandelbaum, A., Nguyen, V., & Schwerer, E. (1995). From project to process management - An empirically-based framework for analyzing product development time. *Management Science, 41*(3), 458-484.

Anderson, E. G., Morrice, D. J., & Lundeen, G. (2005). The "physics"' of capacity and backlog management in service and custom manufacturing supply chains. *System Dynamics Review, 21*(3), 217-247.

Banerjee, S. , Hopp, W. J. (2001). The Project Portfolio Management Problem. Department of Industrial Engineering and Management Sciences, Northwestern University, June.

Cooper, R.G., Edgett S.J., Kleinschmidt E.J. (1998): *Portfolio Management for New Products.* 2nd Ed. Perseus Publishing, MA.

Dahan, E., & Mendelson, H. (2001). An extreme-value model of concept testing. *Management Science, 47*(1), 102-116.

Deeds, D. L., & Rothaermel, F. T. (2003). Honeymoons and liabilities: The relationship between age and performance in research and development alliances. *Journal of Product Innovation Management, 20*(6), 468-484.

Forrester, J. W. and Senge, P. M. (1980). Tests for building confidence in system dynamics models. *TIME Studies in the Management Science* 14, 209-228.

Galambos, Janos (1978). *The asymptotic theory of Extreme Order Statistics*. John Wyley and Sons.

Gino, F., Pisano, G. (2005). Do Managers' Heuristics Affect R&D Performance Volatility? A Simulation Informed by the Pharmaceutical Industry. Harvard Business School Working Paper.

Girotra, K., Terwisch, C., Ulrich, K.T. (2005). Managing the Risk of Development Failures: A Study of Late-Stage failures in the Pharmaceutical Industry. The Wharton School, University of Pennsylvania, January.

Greene, W.H. (1997). *Econometric Analysis*, 3rd Edition. Upper Saddle River, NJ: Prentice-Hall.

Griffin, A. (1997). PDMA research on new product development practices: Updating trends and benchmarking best practices. *Journal of Product Innovation Management, 14*(6), 429-458.

Gumbel, E. J. (1958): *Statistics of Extremes*. Columbia University Press, New York.

Heskett, J.L., Sasser, W.E., Schlesinger, L.A. (1997). *The Service Profit Chain*. New York: Free Press.

Khurana, A., & Rosenthal, S. R. (1997). Integrating the fuzzy front end of new product development. *Sloan Management Review, 38*(2), 103-120.

Law, A.M., Kelton, W.D. (2000). *Simulation Modeling and Analysis* (3rd ed.) McGraw-Hill: New York.

Oliva, R., & Sterman, J. D. (2001). Cutting corners and working overtime: Quality erosion in the service industry. *Management Science, 47*(7), 894-914.

Schmidt, J.B., Sarangee, K., Montoya-weiss, M.M. (2006). Exploring New Product Development Project Review Practices and Performance. Draft, being revised at *Journal of Product Innovation Management*.

Schmidt, J. B., & Calantone, R. J. (2002). Escalation of commitment during new product development. *Journal of the Academy of Marketing Science, 30*(2), 103-118.

Sterman, J. D. (1989). Modeling managerial behavior - Misperceptions of feedback in a dynamic decision-making experiment. *Management Science, 35*(3), 321-339.

Thomke, S., & Fujimoto, T. (2000). The effect of "front-loading" problem-solving on product development performance. *Journal of Product Innovation Management, 17*(2), 128-142.

Ulrich,K.T., Eppinger, S.D. (2004). *Product Design and Development*. Third Edition, McGraw-Hill, New York.

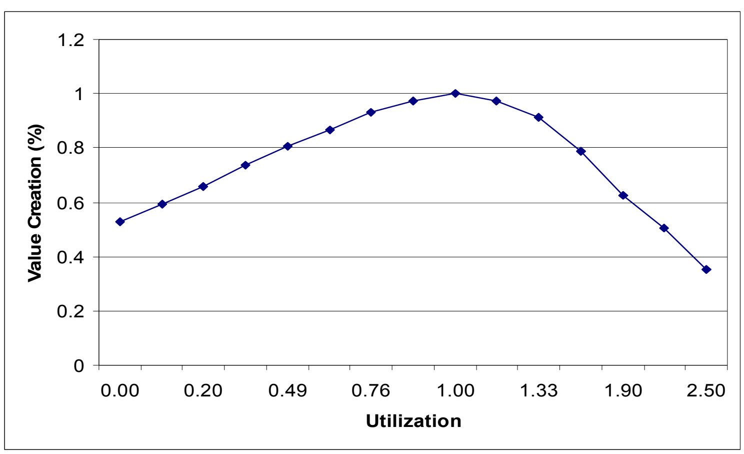
Wheelwright, S.C. and Clark, K. B. (1992). *Revolutionizing Product Development: Quantum Leaps in Speed, Efficiency and quality*. The Free Press.

Appendix

A.1 Model Parameters and Utilization Function

Time to adjust capacity	0.5 month
Nominal development time	2 months
Nominal Capacity	5 projects/month
Time to review a project	1 month

Default Parameters



Relationship between Resource Utilization and NPV Creation Multiplier

The shape of this function follows Wheelwright and Clark (1992, pg. 91), with peak normalized to 1.0 when utilization is set at unity.

A.2 Screening using Gumbel Distribution

Number of projects that are terminated or approved, depending on the net value, is calculated by assuming that the NPVs follow a Gumbel probability distribution, with a mean equal to "average NPV in stage review" and a selected standard deviation of 0.38 in stage 1 and 0.64 in stage 2. We establish the total value that is lost and the total value that is transferred to the next stage by calculating the average NPV of the terminated projects and the average NPV of the approved projects. The same process is repeated for the second phase. The probability density function of the Gumbel (maximum) distribution is

$$f(x) = \frac{1}{\beta} e^{-\frac{(x-\mu)}{\beta}} e^{-e^{-\frac{(x-\mu)}{\beta}}}$$

Here μ is the location and β is the scale parameter. The mean is equal to $\mu + 0.5772\beta$ and the standard deviation is equal to 1.2825β. Calculation of termination criteria (P, or percentage of terminated projects) is a table function computed from the following integral:

$$\text{Termination Criteria} = \int_{-\infty}^{Y} f(x)\, dx$$

The percentage accepted is percentage complement of the termination criteria. If Y is the termination threshold, then the equation for setting up a table function for correcting the Average NPV of the terminated projects is:

$$\text{Performance IndexT} = \frac{1}{P} \int_{-\infty}^{Y} x * f(x)\, dx$$

The equation that calculates the index for average NPV of the approved projects is:

$$\text{Performance IndexF} = \frac{\text{Ave NPV in Stage Review} - \text{Performance IndexT} * (1 - \text{Percentage Accepted})}{\text{Percentage Accepted}}$$