Optimal Execution of Real-Options in Illiquid and Incomplete Markets

By Wajahat H. Gilani

A dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

in partial fulfillment of requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Dr. M.N. Katehakis

and approved by

_____

_____

_____

_____

_____

Newark, New Jersey

May, 2016

ABSTRACT OF THE DISSERTATION

Optimal Execution of Real-Options in Illiquid and Incomplete Markets

By Wajahat H. Gilani

Dissertation Director:

M.N.Katehakis

This dissertation, consists of three essays on the problem of quantifying optimal stopping policies for a multi-period investment, where transition probabilities and the investment value itself are uncertain. These models are applicable to entrepreneurs in the technology sector and any investment where option based approach can be taken.

In the first chapter, I convert the multi-period investment into a partially observable Markov decision process model with bayesian learning. I assume that the core process of the investment value is not observable during the multi-period investment process but can be observed only in its final state if the decision to exploit the investment is made. I assume that the probability distribution between the observed demand levels and the underlying value is known. Since this POMDP model is difficult to solve with dynamic programming because of the size of the possible states, we introduce

a heuristic based on marginal profit gains at each state. With the marginal profit heuristic we can calculate the minimum probability threshold of the unobservable state, in a 2-state model, that is the optimal stopping for the process.

In the second chapter, I drop the assumption of knowing the probability distribution between the observable demand and unobservable underlying value of the state to the investment, and replace it with a second type of demand level that when observed together with the first demand level imply certain values of the underlying investment. I introduce an algebraic logistic function that has the characteristics of a sigmoid distribution, to serve as an approximation of the probability of the underlying state, based on the observations of the two demand levels but the ratio between them quantify the probability, not a known distribution. Since this model has no defined transition matrix, I develop a best case heuristic, for the 2-state model, that finds a local optimal range, without the use of the Lambert function, and therefore optimal stopping point when a local optimal range does not exist. For the n-state model we define least-case heuristic, similar to the best-case heuristic, except m-local optimal ranges are defined, where m<n and corresponds to the number of states with a positive return.

In the third chapter, using the algebraic sigmoid function from the second chapter, I develop a policy approximation problem for the N-state model, where I define an optimal policy that maps the probability of the states of the underlying value of the investments, to an action at each period. In addition, I apply the best-case heuristic from chapter 2 in aggregating the N-state into a M state decision problem.

Preface

This Ph.D. dissertation entitled "Real-Option Execution in Illiquid and Incomplete Markets" has been prepared by Wajahat H. Gilani during the period September 2008 to May 2016 at the department of Management Science and Information Systems at Rutgers University, Newark and New Brunswick.

The Ph.D. project has been completed under the supervision of my advisor Professor Michael N. Katehakis. The dissertation is submitted as a partial fulfillment of the requirement for obtaining the Ph.D. degree at the Rutgers University. The project was supported by a Instructor Position at the Rutgers Business School, Rutgers University.

I would like to first express my sincere gratitude to my advisor, Dr. Michael N. Katehakis. I am very fortunate to be his student during my Ph.D work at Rutgers. I truly appreciated the continuous support, keen research incite, everlasting patience, and encouragement he gave me to pursue my Ph.D. degree. His advices are invaluable. Without his guidance this dissertation would not have been possible. I am also indebted to him for introducing me to the world of research and showing me how to take complicated theories and break them down into simpler models that can be programmed and analyzed. I want to also thank the committee for being patient with me during the dissertation process.

I would also like to thank my father, mother, brother, sister-in-law without whom, everything I have accomplished would not have been possible, and the countless hours they spent taking care of my daughters so that I could finish my research. I would like to thank my wife and my daughters for their immense patience and gratitude for the

iv

Contents

List of Figures

Introduction

The 7th century Arabian general Khalid ibn al-Walid, is respected as one of the greatest military minds of the ancient world. Given the title, "The Sword of God", by the Islamic prophet Muhammad himself, he was known for his superior tactics and strategies and was able to lead the smaller forces of the recently created Muslim nation in Arabia to victory over the vastly superior armies of the Sassanid-Persian Empire and Byzantine-Roman Empire Akram (2004). Today historians and researchers debate and analyze the many factors that lead to such remarkable victories, some of which can directly be applied to the subject of investing in highly uncertain businesses. First, he created an elite light cavalry utilizing fast camels and horses, known as the mobile guard (Tulay'a mutaharikka) Malik (1968). Second, he equipped his cavalry with lances, allowing the mostly skirmish oriented Arabs, who had little experience in open field warfare, to engage and disengage with their adversaries with relatively little loss. Third, having created the fast infantry with lances, he employed hit and run tactics, attacking and then retreating to attack the flanks and rears. Lastly, he created a core group of advisors of intelligent men from the various regions he had been engaged in battle that served as a simple version of what a modern military today would refer to as an intelligence staff Akram (2004).

The innovations and tactics employed by the general are in themselves impressive, but when looked at as a whole we claim that a particular framework starts to become evident; an option-based framework combined with observational learning. An option based framework, is a multi-stage decision tree, where at any given stage, there is a certain or a probabilistic current state of the process. From the current state, an

agent must decide to take an action, defer an action, or decide to make the current state the final one. If an action or deferment is taken, then an exogenous force is applied in tandem possibly causing the state to move to another state or stay in the current state, in the next stage. The stages of the decision tree continue to either a predefined number of stages or until the agent chooses to make the current state the final state. The multistage methodology is favored for situations where there is a great deal of uncertainty. At every stage of the decision process the agent can make observations and collect data. This allows the agent to make incremental gains in knowledge and allows for flexible and adaptive responses. The large unknown decision can be made into smaller decisions that increase with certainty.

The general faced a large amount of uncertainty about the capability and strategies of the opposing army. To deal with such uncertainties he took what was a one-time large bet, a full frontal assault, and transformed it into many smaller bets, i.e. the hit and run tactics. This not only slowly weakened his adversaries, but also allowed him to probe for weaker flanks while allowing him to quickly "learn", via his intelligence staff, the best possible points of attacks in a manner that was relatively the safest way for this cavalry to attack and fight. Whenever he engaged in battle, he gave himself the option to continue an attack, change the direction of the attack, or retreat. All the while collecting better information with every decision, leading to better decisions and options during the battle, in the cheapest manner (minimizing his army's casualties).

An entrepreneur, like the general Khalid ibn al-Walid, is faced with a large amount of uncertainty when he or she is looking to develop a new product or service. In the face of such uncertainty, the entrepreneur needs to quickly gather information relevant to his or her venture in a cost-efficient manner while simultaneously being flexible enough to respond to the incremental information. Just like the general, the
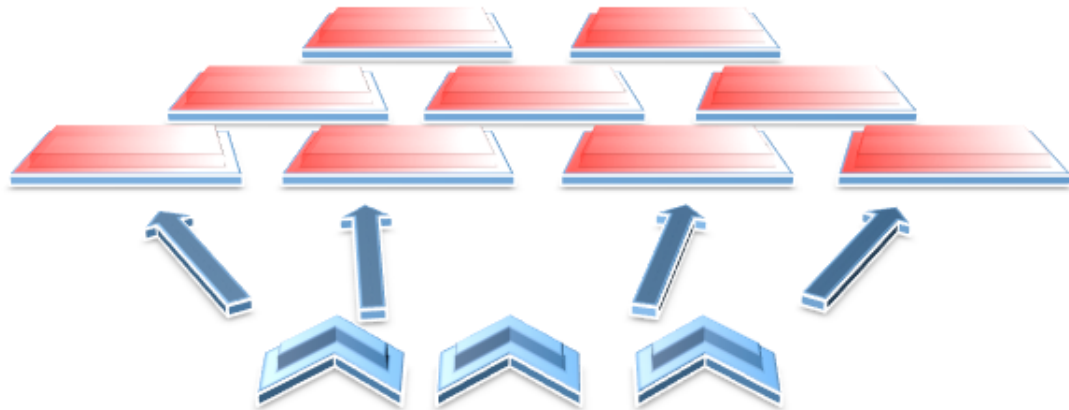
Figure 1: Standard Warfare formation

ideal scenario for the entrepreneur is an option-based framework with observational learning. According to Schwartz and Trigeorgis (2004), real options are essentially opportunities that are irreversible and allow the entrepreneur the right but not the obligation to exploit the opportunity. The authors frame the option-based framewoark as a series of compound options where the value of each individual option or decision depends on other options or decisions. This framework implies the entrepreneur takes the one potentially large bet and transforms it into many smaller, faster and cheaper bets, which in turn allows for more information gathering opportunities.

In today's startup world, this philosophy is best captured by the lean methodology, Ries (2011). Lean's basic principle is to make a minimum viable product (MVP) as quickly and cheaply as possible, release it into the market and gather live feedback from actual consumers in the market. Then take the information derived from the feedback and change the product or service in a manner where the market will find it more appealing. The idea is to keep each stage as cost-effective and small as possible so many cycles can be done. Once the entrepreneur starts to see exponential growth, the entrepreneur then looks to develop a more sustainable version of his or her MVP. This requires raising capital from professional investors, and where the
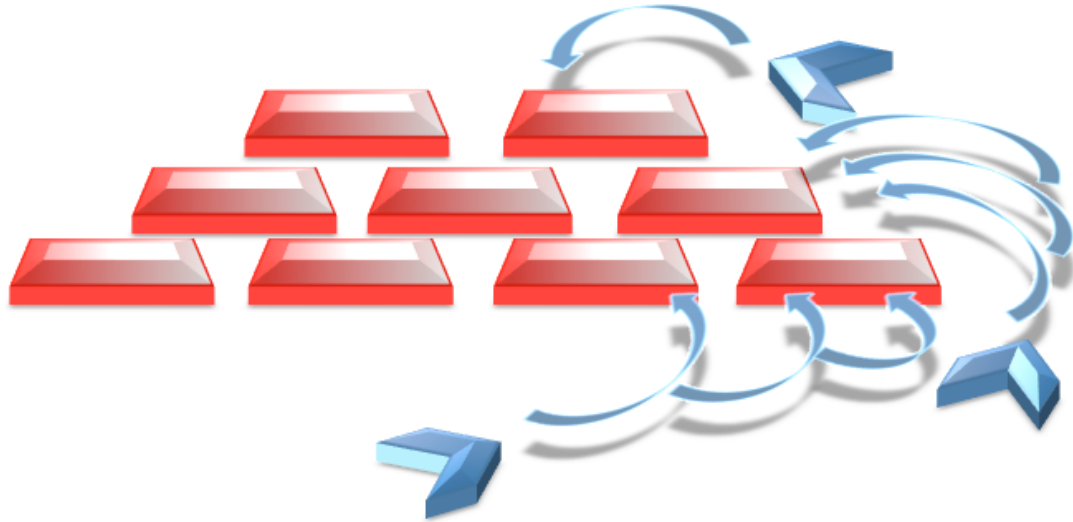
Figure 2: General Khalid ibn al-Walid hit and run formation

value of information gathered and the exponential growth achieved is of value. This is the equivalent to executing an option contract.

In terms of businesses focusing on consumers, exponential growth is the critical point. It serves as the optimal execution point for the entrepreneur to raise capital from professional investors and build the current venture to scale. Many of the larger and established Venture Capitalists stress to their portfolio companies that capturing a large number of users is more important then generating revenue. The idea being that building a large consumer base is much more difficult then establishing ways to profit from said consumer base. In addition, growth is a way to shield against new competition Deeb (2014). This strategy has lead to the founding of a sizeable number of start up companies, called unicorns; private companies with a valuation over one billion dollars, even if their net profits are significantly less. As long as they have exponential or reinforced growth, then investors will continue to valuate them at high levels. With large consumer bases only a tiny of fraction of monetization is needed to reap asymmetrically large profits. Yet, many companies with exponential growth and a large initial demand, have failed when demand, for various reasons, reverse course
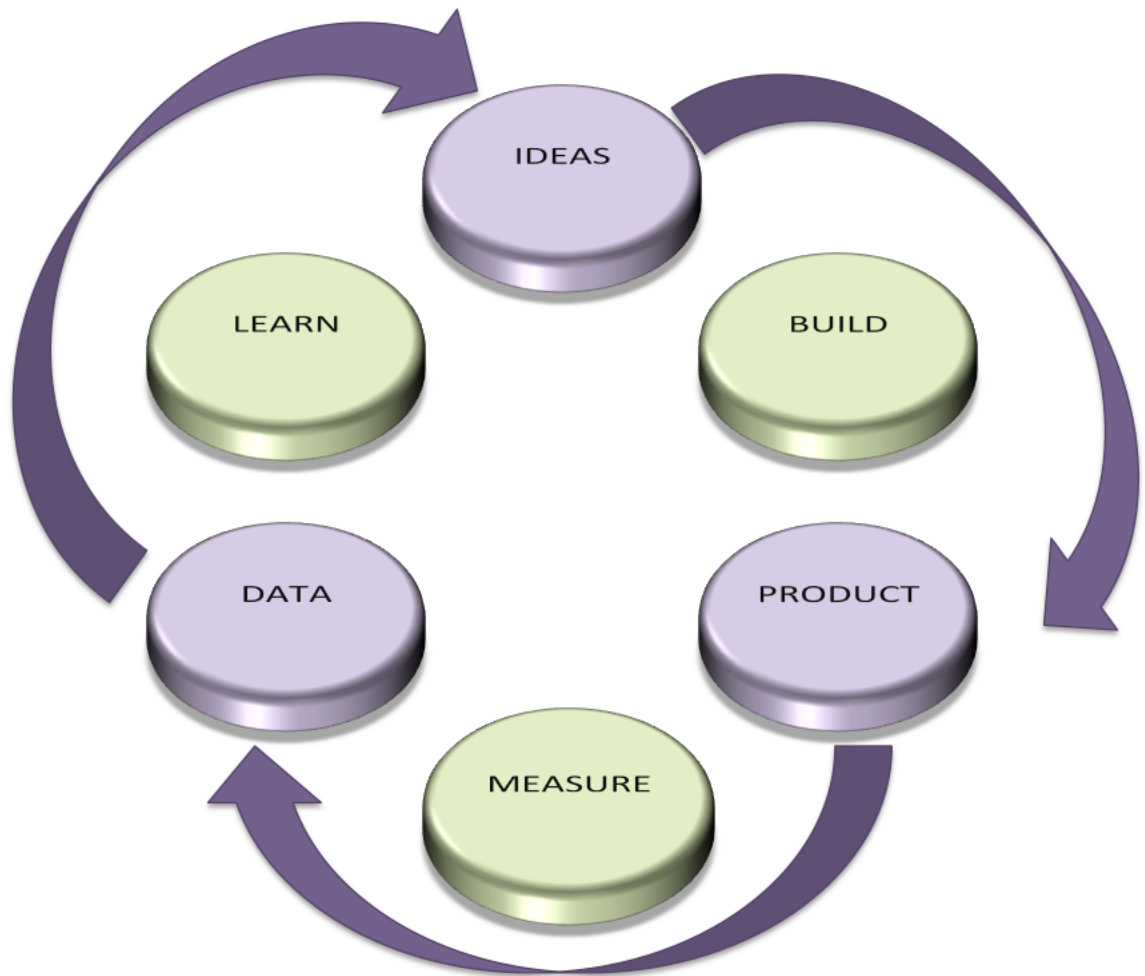
Figure 3: Lean Start-up Methodology

CBInsights (2015).

The focus of this paper is on the optimal decision to be made at each stage of a multi-period investment, which may have more then one stage, when faced with uncertainty and the lack of a statistical base line. We define a multi-period investment as any investment where there is more then one preiod of investing and where one option has the right to purchase another option, or where one decision can lead to another decision. Technology startups fulfill this definition because they raise capital through different phases of their growth. Infrastructure investments can also satisfy this definition if prior to starting the project, live experiments are done to measure the efficacy of the project, i.e. measuring the flow and speed of traffic at different times of the day within a certain location. This can be considered a two-stage investment problem, with many period, where the live measuring of traffic and collection of data is the first stage of investment and the decision to pursue the infrastructure project is the second and last stage of investing.

0.1   Barriers To Real Option Application

The exisiting literature in Real Options is filled with assumptions that are difficult to make for startups or unique types of businesses. In the case of Herath and Park (2002), they define three different volatilities, σ_1, σ_2, and σ_3, for three unique investments that are made in sequential stages and are interrelated to each other. The volatilities σ_1, σ_2, and σ_3 are independent of one another and have unique distributions. This is an assumption our model cannot make because our focus is on multi-stage investments where we have no direct historical information on the volatility σ distribution of the underlying revenues for each incremental investment. To address the lack of historical information, we can construct a customized benchmark from the distributions of

similar features of other investments where historical information exists. In Xu et al. (2014), a feature-based regression algorithm is used for a recommendation system when a new user with no prior interaction with the system attempts to use it. The user's features are compared to users of similare features and their probability models are applied to the new user. Another way to address the uncertainty of our model's volatilities σ for the underlying investments, is to apply a uniform distribution on our parameters within rational bounds. In Barmish and Lagoa (1997), the case is made to use the uniform distribution for uncertain parameters given certain assumptions about the underlying density functions. When the assumptions are weaker they state truncated uniform distributions give similar results.

In both cases whether we use the feature equivalent constructed probabilities or the uniform distribution, our model will be sensitive to the assummptions used in the construction of those probablity models. The constructed probability model's critical assumption is that the correlation of the features' probabilities have a negligble effect on the whole benchmarked probability. The critical assumption made on the uniform distributions are the boundaries. In some cases, the bounds (or at least one bound) are clear, but in some cases the theoretical boundary, or boundaries, may not exist. For example, stocks have a clear lower bound, they can never go below zero, but in the case of the upper bound there is no clear boundary because stock prices, in theory, can go up to infinity. Our model is a multi-stage decision which allows us the opportunity to process new signals about the distributions of our volatility σ and modify them accordingly. The feature based benchmark distribution or the uniform distribution both can serve as an initial best guess or prior distribution, for our model's uncertain distributions. In Herath and Park (2001), using bayesian analysis they perform a pre-posterior analysis to calculate the expected value of sampling information (EVSI) which is then compared to their calculated quasi-real option value. If the EVSI reduces a significant amount of uncertainty, then the sampling is done and the value

of the real option is re-assessed with better information, leading to another decision of whether to keep holding the real option. Their examples assume that the samples and bayesian processes follow normal distributions, but our model not having any direct historical data does not make any assumptions about the distributions.

Attempting to learn the probability distribution of an underlying value of a multi-stage decision or option contract complicates our model with three underying assumptions. The first assumption is that our prior distribution is an adequate representation of the underlying distribution to price the primary real option at stage one of the investment. This assumption affects the decision whether or not to buy the first option or make the first investment decision. At this point, all that is known is the prior distribution, and in contrast to other literature, we have do not assume a transition probability. Second, the optimal number of stages to derive a probability distribution that is within an acceptable rate of error from the true underlying volatility σ of the investment can be quantified for our model. Given the uncertain nature of the underlying distributions it is difficult to find a scenario where the data informed distribution can be deemed "complete" without an exogenous signal. This assumption relates to knowing when to stop exploring and to start exploiting an opportunity, or when to start different stages of partial exploration/exploitantion. Third, the true underlying volatility σ of our multistage investments is not associated with fat tails and non-charateristic scale, i.e. power law, which if it were the case would expose our model to tail risk as discussed in Taleb and Bar-Yam Taleb et al. (2014). The possible exposure to tail risk in probabilistic investing can cause us to see a temporal opportunity when the actual volatility and expected profit might be far less or uncertain as discussed in Makridakis and Taleb (2009). These properties highlight the difficulty in using option valuaition techniques in illiquid and incomplete markets and show why managers and consultants prefer simpler modeling techniques over option valuation techniques as cited in Rigby Rigby (2001). They found real option analysis

had low satisfaction rates and high defection rates. In summary, our model needs to answer; if the first or outer layer option should be purchased, when to exploit or move to another stage of exploration, and if the distribution of the derived value is stable.

To construct our multi-stage option model we look to the frameworks used by Venture Capital (VC) firms that invest in disruptive and scalable technologies. VC firms look to invest in ventures that look to challenge established businesses and markets and by their very nature are initiatives that attempt to be unique in products and/or services. The innovative and unique nature of these businesses carries with them extreme risk and uncertainty, both about the company and the market. To attempt to manage the uncertainty VC firms fund these startrups with several rounds of funding with customized contracts sometimes linked to performance and milestones. In Bienz and Hirsch (2011), they show how the nature and magnitude of uncertainty and information asymmetry drives VC firms to use and customize multi-staged financing contracts allowing VC firms to manage uncertainty in a more efficient manner. The miletstones and performances help guide both the start-up companies and VC firms in determing when or if to move to another round of funding, or when to execute another subset of the options. In Choi et al. (2008), they discusses the ignorance reduction process from the start-up companies point of view in determining how to define a time threshold to go from exploration to exploitation. They derive propositions focusing on the defensibility of the business from competitors and the rate of cost of exploratory phase of the business. In this paper, we structure our framework around the growing philosophy of the VC world, that looks to invest in "networks", Weissman (2012). Investing in networks, essentially means investing in the product or service that is a by product of a semi-auntonomous micro market. Facebook, twitter, Uber are companies that can all be considered networks, to some degree. The network, is analogous to a freemium product or service that has a base model that is free and

then can be upgraded or can offer additional services for a payment of some kind. The concept of network is superior to a traditional freemium in the sense that the network is semi-autonomous and is part of the solution process to solve whatever the objective of that network is trying to solve, or provide services and products to reinforce the purpose of that network existing in the first place. A better freemium product or service analogy would be, a product or service where the consumer has some control to enhance the product or service to some degree, or at the very least contribute ideas towards a better product. For the purposes of the is paper we will use networks and freemium interchangeably. The benefits for investors and entrepreneurs in investing in a network are two-fold. First, with any new product or service, there is the question of whether there is a significant and sustainable need for said product or service. Many initial startups have exponential growth and solid marketshare, only to have that marketshare dwindle after a period of time. A common example is MySpace whose size became dwarfed by its simpler more viral competitor Facebook Giliette (2011). A growing and vibrant community, or network, where growth is largely driven by the network itself, is the quantitative evidence that there is a set of needs that has a market looking for solutions. This segways into the second benefit, because any product or solution that is tailored and/or inspired by the network, can be seemlessly integrated into the network and has a built-in marketshare, and which allows it to be defensible against copycat competitors. If the product or service fails to convert network members into paying customers, then the product or solution failed solve the need and want of the network, but the failure will be faster via the network, saving potential money and time. A great example of a network is the arts and crafts market place Etsy. The founders initially had worked on a community forum dedicated to crafters. The common theme of the discussions on the board was a way to see their crafts that was cheaper and easier then Ebay. The founders built Etsy, per the requirements and needs they could gather from the forums and then invited

the forum to join. From there it grew exponentially, as more crafting communities were invited and the sellers themselves marketed the site, since it directly benefited them if more users were coming to the site to buy products.

We want to combine the concept of networks to the real option approach to make a Network Option Decision Process (NODP) that can apply option-based techniques in a way that is applicable to the industry. In contrast to the NODP Model we want to compare the application of the Partially Observable Markov Decision Process (POMDP) to the same type of problem. The POMDP Model assumes that the probability of investments are unobservable during the investment process unless the full investment is exploited. In addition, there is an observation process that is probabilistically related to the unobservable core process, so that a probabilistic guess can be made about to the true state of the investment in any give period. This converts the POMDP model into a MDP model with a very larget state space. These large intractable decison problems require the development of heuristics and faster methodologies to solve. In this paper we propose optimal stopping heuristic for the POMDP model that is based on the Dowry Problem Gilbert and Mosteller (2006) and marginal profits. There has been much work on addressing large dimensionality problems and faser alorithms. In Katehakis et al. (2015), they introduce a faster methodology for solving large scale systems of equations that arise in markov decision problems, with the use of a matrix based algorithm, and in Ertiningsih et al. (2015) they extend the concepts to quasi-skipfree processes, where as in Cowan and Katehakis (2015a) the exploitation vs exploration problem is explored in a stochastic multi-armed bandit. In Cowan and Katehakis (2015b), they address the problem of making a choice in a multi-armed bandit setting where the means of the options available can be infinite, through the use of use bounded distributions, where as in Cowan and Katehakis (2015c) they provide an optimal policy of maximizing the expected sum out of N random populations. In, Katehakis and Smit (2012), they introduce a successive

lumping procedure to solve a class of markov chains and in KATEHAKIS et al., they compare successive lumping methodology to lattice path counting, finding that SL based algorithms outperform the lattice path and the former includes a method to calculate steady state distributions. In Fleischhacker et al. (2015), they optimize the discounted revenue of a firm with one pool of inventory selling into two different markets, using efficient dynamic pricing policies. In Shi et al. (2012), the paper address the optimal management of the two basic cash flows of that businesses have to manage, the purchase of inventory and the selling of inventory. Typically in the industry there is a lag and can be financially dentrimental to a business. They establish a two-threshold policy for when to order based on asset levels. In Burnetas et al. (2015), they address the bandit problem with known dependent costs and construct a class of policies that have an asymptotic lower bound. In Katehakis and Puranam (2012), an opitmal value function is established throught the use of monotonic properties. These references go into detail of various optimization and efficiency techniques used in markov related problems but the list is not nearly exhaustive.

In contrast the MDP and POMDP modelds, the NODP Model makes no assumptions about the underlying core process or the observational process, and instead relies on two sets of observations. One we define as the new paying user (NPU) and one we define as new freemium user (NFU). The NPU is observational process from the POMDP model but the NFU is just part of the NODP model. This reduces expectaction complications that the above MDP algorithms attempt to solve but at the expense of a transition matrix. The objective is to quantify the likelihood of the permanent paying user base (PPU), that will in turn drive the valuation of the investment. The PPU cannot be known unless the investment is fully vested and exploited. In chapter 1, we start with a two-state POMDP model and explore the compclications that quickly arise in solving it, and provide a heuristic for an optimal stopping policy. In chapter 2, we define the two-state NODP model and develop a

best-case stopping policy that we extend to a N-state NODP model as an average best-case stopping policy. In chapter 3, we attempt to aggregate the states of the N-State NODP model to apply a dynamic program, in spite of not assuming transition probabilities.

CHAPTER 1

Partially Observable Markov Decision Process

1.1   Two State Model

To develop a network option decision process (NODP) model we will compare and contrast different valuation scenarios against an equivacol partially observable markov decision process (POMDP) model. We assume the valuation V for our multi-stage investment is dependent on the number of permanent paid users (PPU), U_T^p, at execution period T and can only take the values H or L, where

$$U_t^p = \{H, L\} \text{ in period } t$$

$$H > L$$

The concept of permanent paid users or clients, U^p is an ephemeral concept in the tech and business world but investors and entreprenuers focus on the average number of users for a fiscal quarter to represent the permanent paid user level. For the purposes of our paper the permanent paid users U_T^p at period T will be a sufficient representative. In our model the number of permanent paid users is not directly visible by either the POMPDP or the NODP model, but we assume that U_t^p follows a core

process that is represented by the following conditional transitional probabilities

$$P(U_t^p = H | U_{t-1}^p = H) = \alpha \tag{1.1}$$

$$P(U_t^p = L | U_{t-1}^p = H) = 1 - \alpha \tag{1.2}$$

$$P(U_t^p = L | U_{t-1}^p = L) = \beta \tag{1.3}$$

$$P(U_t^p = H | U_{t-1}^p = L) = 1 - \beta \tag{1.4}$$

This is representative of the probabilistic nature of uncertain investments. At no given time do investors know what the true demand for a product or service will be with certainty unless they execute all stages of the investment. Until then, throught the multi-stage investment, management looks to data or signals that are indicative of the probability of the unobservable demand being at a certain level. The conditional transitional probabilies for the PPU U^p represent the strength of the product or service's ability to create a positive network effect. If the α is 100% that implies that the product or service has a strong network affect and the multi-stage investment will be very profitable with little risk. If the α is less then 50%, that implies that the product or service has a negative network effect and the innovative solution was rejected by the particular segment of the market. Our model assumes that the observations of the data or signals that are correlated to the true PPU U_t^p, in every period t, for new paying clients or users is available to both models. In our model the number of new paid users (NPU) per period are the obsevations that are correlated to the true state of the PPU U_t^p at period t. The variable u_t^p, represents the number of new paid users (NPU) in period t, and u_t^p can only take on one of two values h or l where

$$u_t^p = \{h, l\} \text{ in period } t$$

$$h > l$$

The probability of observing the NPU u_t^p at period t is conditional on the true state of PPU U_t^p at period t and is denoted by the following observation probabilities

$$P(u_t^p = h | U_t^p = H) = \delta \tag{1.5}$$

$$P(u_t^p = l | U_t^p = H) = 1 - \delta \tag{1.6}$$

$$P(u_t^p = l | U_t^p = L) = \theta \tag{1.7}$$

$$P(u_t^p = h | U_t^p = L) = 1 - \theta \tag{1.8}$$

Just like the conditional transitional probabilities, only the POMDP model assumes to know the observation probabilities. The NODP model does not assume to know the conditional transition probabilities for U_t^p or the observation probabilities for u_t^p and can only observe its values for any given period t, as can the POMDP model. It is important to note that if the δ, for example, were to be 50%, then when the PPU U^p is H that would imply that the NPU u^p observations are uncorrelated with the PPU U^p and therefore is an incorrect measurement to track. If the δ were less then 50%, that could imply that h is much greater then l to account for the high PPU U^p result.

The NODP model makes one additional assumption that the POMPD model does not make. The variable u_t^f represents the new free users (NFU) at period t, for all the new users of the freemium product or service. The NFU u_t^f can only be one of the two values g or k where

$$u_t^f = \{g, k\} \text{ in period } t$$

$$g > k$$

Model Assumptions		
Assumption	POMDP	NODP
U^p transition probabilities	yes	no
U^p and u^p observation probabilities	yes	no
u^p observations	yes	yes
u^f observations	no	yes

Table 1.1: List of Assumptions for each Model

Table 1.1 summarizes the assuptions our POMDP and NODP models make. In multi-stage investments for unique and innovative ventures it is highly unlikely that the investor or entrepreneur will be able to assume the transition and observation probabilities of the POMDP model. This is why we chose the POMDP model as a good comparison against the NODP model, which is more typical of the type and depth of information entrepreneurs and investors have access to. The POMDP model does not assume the NFU u^f, this is because the transition and observation probabilities reveal more information about the PPU's value. Our model assumes the cost C_t to the be the cost of continuing the investment at every period t. The problem is defined as a compound option, where the entrepreneur can decide to 1) continue exploring, 2) abandon the investment opportunity altogether, or 3) swich to exploitation. At every new stage for the POMDP model the probabilistic value for U^p is re-evaluated, where as for the NODP model the objective is to see either exponential growth in u^p relative to u^f or to a lesser degree linear growth of u^p conditional on the linear growth on u^f. In terms of our current two-state model, this means consistently seeing $u^p = h$ when we see $u^f = g$, where as in the technology industry VC firms focus on exponential growth in u^f followed by exponential growth in u^p. Our model assumes the profit π received from this multi-stage investment will be valuation V^i net of the costs C_t for all stages executed

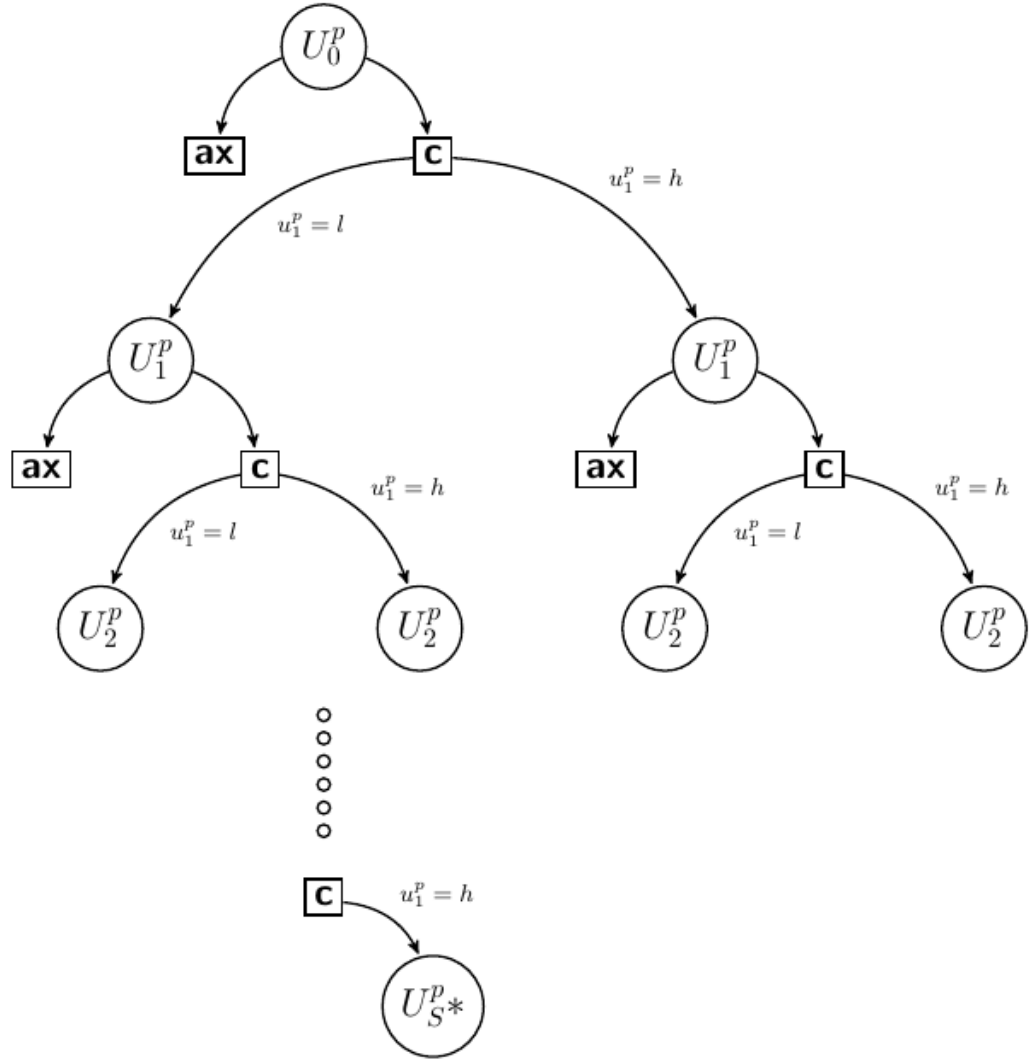Figure 1.1: Two-State POMDP Transtional Diagram

$$\pi = V^i - \sum_{t=1}^{T^*} C_t \tag{1.9}$$

where T^* is the stage at which the entrepreneur or investor chooses to abandon or exploit the investment and becomes the final stage.

Figure 1.1 is the transitional diagram for our two-state POMDP model. At perod 0, the probability of PPU $P(U_0^p = H)$ is represented by the variable ϵ^H and defined

by the bounded uniform distribution or the feature-based model, and serve as "best-guesses" for the investment at period 0. Since the state of our model can only have two option, $P(U_0^p = L) = \epsilon^L = 1 - \epsilon^H$. At this point the investor or entrepreneur will have to make the decision whether to purchase the initial option at period 0 given the initial best guess ϵ distribution and the underlying transitional probabilities or abandon or execute the deal. At point c in the transition the entrepreneur has continued on the option and has moved to the next period 1 where NPU u_1^p is now observed. We now calculate the new probability for PPU $P(U_1^p)$ in period 1 of the POMPDP model using the result of the NPU u_1^p in period 1. If the NPU $u_1^p = h$ was observed, then the probability for the PPU $P(U_1^p = H | u_1^p = h)$ at period 1 will be

$$P(U_1^p = H | u_1^p = h) = \frac{P(u_1^p = h | U_1^p = H)P(U_1^p = H)}{P(u_1^p = h | U_1^p = H)P(U_1^p = H) + P(u_1^p = h | U_1^p = L)P(U_1^p = L)}$$

(1.10)

where using our transitional probabilities probability $P(U_1^p = H)$ is

$$P(U_1^p = H) = P(U_1^p = H | U_0^p = H)P(U_0^p = H) + P(U_1^p = H | U_0^p = L)P(U_0^p = L)$$

$$\Rightarrow \alpha\epsilon^H + (1 - \beta)\epsilon^L$$

(1.11)

and probability $P(U_1^p = L)$ is

$$P(U_1^p = L) = P(U_1^p = L | U_0^p = H)P(U_0^p = H) + P(U_1^p = L | U_0^p = L)P(U_0^p = L)$$

$$\Rightarrow (1-\alpha)\epsilon^H + \beta\epsilon^L \tag{1.12}$$

and therefore equation (1.10) is

$$\Rightarrow \frac{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \tag{1.13}$$

If the NPU $u_1^p = l$ was observed, then the probability for the PPU $P(U_1^p = H | u_1^p = l)$ at period 1 will be

$$P(U_1^p = H | u_1^p = l) = \frac{P(u_1^p = l | U_1^p = H)P(U_1^p = H)}{P(u_1^p = l | U_1^p = H)P(U_1^p = H) + P(u_1^p = l | U_1^p = L)P(U_1^p = L)} \tag{1.14}$$

$$\Rightarrow \frac{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \tag{1.15}$$

At the end of period 1, the entrepreneur or investor has spent C_1 to observe the NPU u_1^p during this period and now must make the decision again to; 1) abandon, 2) continue, or 3) to exploit. Given that a NPU $u_1^p = h$ was observed in period 1 and assuming that the investor is risk neutral, if the investor decided to exploit in period 1 then expected profit function is

$$\pi_1^h = V^H \left[\frac{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] +$$

$$V^L \left[\frac{(1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] - C_1 \qquad (1.16)$$

where as if the NPU $u_1^p = l$ was oberved and the investors chose to exploit, the expected profit function is

$$\pi_1^h = V^H \left[\frac{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] +$$

$$V^L \left[\frac{\theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] - C_1 \qquad (1.17)$$

If the investor decides to abandon the investment after period 1, then only the cost C_1 would be incurred. If the investor chose to continue the multi-stage investment then the transitional probability of $P(U_2^p = H)$ would be

$$P(U_2^p = H) = P(U_2^p = H|U_1^p = H)P(U_1^p = H) + P(U_2^p = H|U_1^p = L)P(U_1^p = L)$$

where if in period 1 NPU $u_1^p = h$ was observed then

$$P(U_2^p = H | u_1^p = h) = \alpha \left[\frac{\delta(\alpha \epsilon^H + (1-\beta)\epsilon^L)}{\delta(\alpha \epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta \epsilon^L)} \right] +$$

$$(1-\beta) \left[\frac{(1-\theta)((1-\alpha)\epsilon^H + \beta \epsilon^L)}{\delta(\alpha \epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta \epsilon^L)} \right] \qquad (1.18)$$

and

$$P(U_2^p = L | u_1^p = h) = (1-\alpha) \left[\frac{\delta(\alpha \epsilon^H + (1-\beta)\epsilon^L)}{\delta(\alpha \epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta \epsilon^L)} \right] +$$

$$\beta \left[\frac{(1-\theta)((1-\alpha)\epsilon^H + \beta \epsilon^L)}{\delta(\alpha \epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta \epsilon^L)} \right] \qquad (1.19)$$

else if in period 1 NPU $u_1^p = l$ was observed then

$$P(U_2^p = H | u_1^p = l) = (1-\alpha) \left[\frac{(1-\delta)(\alpha \epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha \epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta \epsilon^L)} \right] +$$

$$\beta \left[\frac{\theta((1-\alpha)\epsilon^H + \beta \epsilon^L)}{(1-\delta)(\alpha \epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta \epsilon^L)} \right] \qquad (1.20)$$

and

$$P(U_2^p = L|u_1^p = l) = \alpha \left[\frac{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] +$$

$$(1-\beta) \left[\frac{\theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] \qquad (1.21)$$

Then the investor would pay C_2 to observe the NPU for period 2, u_2^p. If $u_1^p = h$ is observed in period 1 and $u_2^p = h$ is observed in period 2 then the probability of the PPU $U_2^p = H$ will be

$$P(U_2^p = H|u_2^p = h, u_1^p = h) =$$

$$\Rightarrow \frac{P(u_2^p = h|U_2^p = H)P(U_2^p = H|u_1^p = h)}{P(u_2^p = h|U_2^p = H)P(U_2^p = H|u_1^p = h) + P(u_2^p = h|U_2^p = L)P(U_2^p = L|u_1^p = h)}$$

$$\Rightarrow \frac{\delta \left[\alpha \left[\frac{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] + (1-\beta) \left[\frac{(1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] \right]}{\left(\delta \left[\alpha \left[\frac{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] + (1-\beta) \left[\frac{(1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] \right] \right) +}$$

$$(1-\theta)\left[(1-\alpha)\left[\frac{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)+(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]+\right.$$

$$\left.\beta\left[\frac{(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)+(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]\right]\right) \qquad (1.22)$$

If $u_1^p = h$ is observed in period 1 and $u_2^p = l$ is observed in period 2 then the probability of the PPU $U_2^p = H$ will be

$$P(U_2^p = H|u_2^p = l, u_1^p = h) =$$

$$\Rightarrow \frac{P(u_2^p = l|U_2^p = H)P(U_2^p = H|u_1^p = h)}{P(u_2^p = l|U_2^p = H)P(U_2^p = H|u_1^p = h) + P(u_2^p = l|U_2^p = L)P(U_2^p = L|u_1^p = h)}$$

$$\Rightarrow \frac{(1-\delta)\left[\alpha\left[\frac{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)+(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]+(1-\beta)\left[\frac{(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)+(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]\right]}{\left((1-\delta)\left[\alpha\left[\frac{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)+(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]+(1-\beta)\left[\frac{(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)+(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]\right]+}$$

$$\theta\left[(1-\alpha)\left[\frac{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)}{\delta(\alpha\epsilon^H+(1-\beta)\epsilon^L)+(1-\theta)((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]+\right.$$

$$\beta \left[\frac{(1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)}{\delta(\alpha\epsilon^H + (1-\beta)\epsilon^L) + (1-\theta)((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] \right] \right) \tag{1.23}$$

If $u_1^p = l$ is observed in period 1 and $u_2^p = h$ is observed in period 2 then the probability of the PPU $U_2^p = H$ will be

$$P(U_2^p = H | u_2^p = h, u_1^p = l) =$$

$$\Rightarrow \frac{P(u_2^p = h | U_2^p = H)P(U_2^p = H | u_1^p = l)}{P(u_2^p = h | U_2^p = H)P(U_2^p = H | u_1^p = l) + P(u_2^p = h | U_2^p = L)P(U_2^p = L | u_1^p = l)}$$

$$\Rightarrow \frac{\delta \left[(1-\alpha) \left[\frac{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] + \beta \left[\frac{\theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] \right]}{\left(\delta \left[(1-\alpha) \left[\frac{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] + \beta \left[\frac{\theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] \right] \right) +}$$

$$(1-\theta) \left[\alpha \left[\frac{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] + \right.$$

$$(1-\beta) \left[\frac{\theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)} \right] \right] \right) \tag{1.24}$$

If $u_1^p = l$ is observed in period 1 and $u_2^p = l$ is observed in period 2 then the probability of the PPU $U_2^p = H$ will be

$$P(U_2^p = H | u_2^p = l, u_1^p = l) =$$

$$\Rightarrow \frac{P(u_2^p = l | U_2^p = H) P(U_2^p = H | u_1^p = l)}{P(u_2^p = l | U_2^p = H) P(U_2^p = H | u_1^p = l) + P(u_2^p = l | U_2^p = L) P(U_2^p = L | u_1^p = l)}$$

$$\Rightarrow \frac{(1-\delta)\left[(1-\alpha)\left[\frac{(1-\delta)(\alpha\epsilon^H+(1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H+(1-\beta)\epsilon^L)+\theta((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right] + \beta\left[\frac{\theta((1-\alpha)\epsilon^H+\beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H+(1-\beta)\epsilon^L)+\theta((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]\right]}{\left((1-\delta)\left[(1-\alpha)\left[\frac{(1-\delta)(\alpha\epsilon^H+(1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H+(1-\beta)\epsilon^L)+\theta((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right] + \beta\left[\frac{\theta((1-\alpha)\epsilon^H+\beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H+(1-\beta)\epsilon^L)+\theta((1-\alpha)\epsilon^H+\beta\epsilon^L)}\right]\right] +}$$

$$\theta\left[\alpha\left[\frac{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}\right] + \right.$$

$$\left. (1-\beta)\left[\frac{\theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}{(1-\delta)(\alpha\epsilon^H + (1-\beta)\epsilon^L) + \theta((1-\alpha)\epsilon^H + \beta\epsilon^L)}\right]\right]\right) \tag{1.25}$$

The equations (1.22), (1.23), (1.24), and (1.25), show that the POMDP model is path dependent and to properly value the option to continue into the next period (i.e, from period 1 to period 2), requires that all the observations and subsequent transitions be calculated (i.e., from period 2 onward).

In figure 1.2, a transition diagram for a three-period two-state POMDP model is shown. In order for an investor to make the decision whether to continue from state
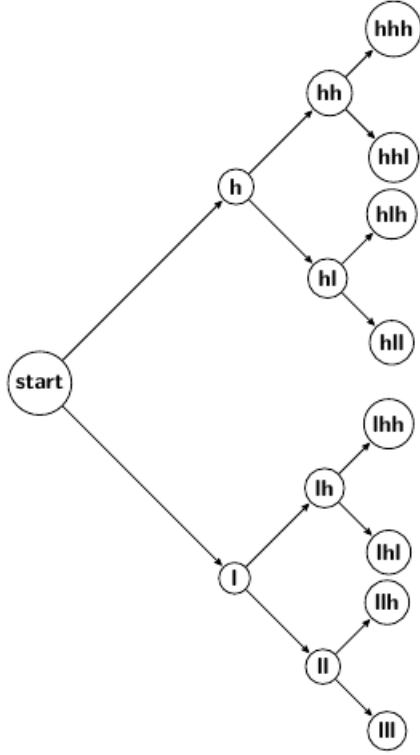
Figure 1.2: Two-State Three-Period POMDP Transtional Diagram

h in period 1, the values of state hh and hl must be known. Likewise to value states hh and hl, it is necessary to have the values of states hhh, hhl, hlh, and hll. Given the computational complexity in the POMDP model our paper uses an optimal stopping point based on marginal profit (MP). The MP stopping point calculates the probability of the states $U^p = H$ and $U^p = L$ at which point there no marginal gain in continuing the multi-stage investment and abandonment or execution must be chosen. This is possible because the two-state POMDP allows us to quantify the two values where in an n-state POMDP model, where $n > 2$ the values would have to be approximated. The optimal prability values for the two states are represented by the variable mp^H for the optimal probability $P(U^p = H)$ and mp^L for the optimal probability $P(U^p = L)$. This is done by focusing on the marginal revenue of the

transitional probabilities where the states change, as in our model

$$P(U_t^p = L|U_{t-1}^p = H) = 1 - \alpha$$

$$P(U_t^p = H|U_{t-1}^p = L) = 1 - \beta$$

We want to calculate the probabilities, mp^H and mp^L by calculating the next periods marginal revenue as being equal to marginal cost

$$MR = V^H P(U_{t+1}^p = H|U_t^p = L)mp^L + V^L P(U_{t+1}^p = L|U_t^p = H)mp^H$$

$$mp^H + mp^L = 1$$

$$MC = C_{t+1}$$

$$\Rightarrow V^H P(U_{t+1}^p = H|U_t^p = L)mp^L + V^L P(U_{t+1}^p = L|U_t^p = H)mp^H = C_{t+1}$$

$$\Rightarrow V^H P(U_{t+1}^p = H|U_t^p = L)(1 - mp^H) + V^L P(U_{t+1}^p = L|U_t^p = H)mp^H = C_{t+1}$$

$$\Rightarrow V^H P(U_{t+1}^p = H|U_t^p = L) - mp^H V^H P(U_{t+1}^p = H|U_t^p = L) + V^L P(U_{t+1}^p = L|U_t^p = H)mp^H$$

$$= C_{t+1}$$

$$\Rightarrow -V^H P(U_{t+1}^p = H|U_t^p = L)mp^H + V^L P(U_{t+1}^p = L|U_t^p = H)mp^H$$

$$= C_{t+1} - V^H P(U_{t+1}^p = H|U_t^p = L)$$

$$\Rightarrow mp^H(V^L P(U_{t+1}^p = L|U_t^p = H) - V^H P(U_{t+1}^p = H|U_t^p = L)) = C_{t+1} - V^H P(U_{t+1}^p = H|U_t^p = L)$$

$$\Rightarrow mp^H = \frac{C_{t+1} - V^H P(U_{t+1}^p = H|U_t^p = L)}{V^L P(U_{t+1}^p = L|U_t^p = H) - V^H P(U_{t+1}^p = H|U_t^p = L)} \qquad (1.26)$$

and

$$\Rightarrow (1 - mp^L) = \frac{C_{t+1} - V^H P(U_{t+1}^p = H | U_t^p = L)}{V^L P(U_{t+1}^p = L | U_t^p = H) - V^H P(U_{t+1}^p = H | U_t^p = L)}$$

$$\Rightarrow -mp^L = \frac{C_{t+1} - V^H P(U_{t+1}^p = H | U_t^p = L)}{V^L P(U_{t+1}^p = L | U_t^p = H) - V^H P(U_{t+1}^p = H | U_t^p = L)} -$$

$$\frac{V^L P(U_{t+1}^p = L | U_t^p = H) - V^H P(U_{t+1}^p = H | U_t^p = L)}{V^L P(U_{t+1}^p = L | U_t^p = H) - V^H P(U_{t+1}^p = H | U_t^p = L)}$$

$$\Rightarrow mp^L = \frac{V^L P(U_{t+1}^p = L | U_t^p = H) - C_{t+1}}{V^L P(U_{t+1}^p = L | U_t^p = H) - V^H P(U_{t+1}^p = H | U_t^p = L)} \qquad (1.27)$$

The solutions for mp^H in equation (1.26) and mp^L in equation (1.27), are conditional on four assumptions. The first assumption is that

$$V^H > C_t \geq V^L \qquad (1.28)$$

V^H being greater then C_t and V^L is self explanatory but if C_t is greater then V^L, then it implies that the marginal revenue of continuing will always be greater then marginal cost and the multi-stage investments could theoretically never end. In the industry, this is not a possible scenario. If there is a theoretical investment in the industry where

$$V^H > 0$$

$$C_t = V^T \geq 0$$

Then this implies there will be a positive gain with certainty, and the investment should immediately be exploited in period 0. If

$$V^H > 0$$

$$C_t = V^T \leq 0$$

Then this implies that the business is being paid for each period of the multi-stage investment, and no exploitation is needed. Again, this is not a possible scenario in the industry. Construction, programming, Research & Development, labor, commodity inputs, etc., are not free let alone revenue-positive inputs for any business or industry (the Investment Banking industry is a special exception).

The second condition we stated before is

$$P(U_{t+1}^p = H | U_t^p = L) > P(U_{t+1}^p = L | U_t^p = H) \tag{1.29}$$

If condition two is violated then this implies that the venture probabilistically will have negative growth or no network effect, and therefore the investment would never have been chosen in the industry. The third condition is

$$\text{for } i = 1, 2, \ldots, T \tag{1.30}$$

$$(V_i^H - V_{i-1}^H) \leq 0 \text{ and} \tag{1.31}$$

$$(V_i^L - V_{i-1}^L) \leq 0 \tag{1.32}$$

This implies that our model is assuming a constant or decreasing V^H and V^L, where if V^H and V^L were increasing over time then the MP stopping point will give suboptimal results because it is not taking into account future possibilities. This assumption is very possible in the industry. As data from multi-stage investments are processed the initial assumptions about V^H and V^L can be revised and in some cases are functions

of data being collected, $V^L(u^p)$ and $V^H(u^p)$. The MP optimal stopping point can serve as a guide in these situations, serving as a signal to entrepreneurs to justify the future stages of investments, as a lowerbound.

Since the MP optimal stopping poing, mp^H is decreasing in C_t and increaseing in V^L, there is a possibility that for investments where there is little downside, the mp^H will be too high to be used as a practical signal. The greater $P(U_{t+1}^P = L | U_t^p = H)$ is the more likely it becomes. Therefore, we define the variable, mp^{H^*} as the highest value of $P(U^p = H)$ observed during the first m periods of the multi-period investment. If after the first m periods, the $P(U^p = H)$ was never greater then or equal to mp^H, then the entrepreneur continues the multi-period investment until $P(U^p = H) \geq mp^{H^*}$. To define m, we first calculate the longest period, t^*, where the total cost of executing the investment, with $P(U^p = H) = mp^H$ equals the total cost of all the t^* periods of the investment. In other words, the point at which exploiting the investment at the optimal stopping probability, mp^h, no longer becomes profitable.

$$\lfloor t^* \rfloor = \frac{V^H mp^H + V^L P(1 - mp^H) - C^F}{C} \tag{1.33}$$

Where C^F is the cost of exploiting the multi-stage investment. It could be added into V^H and V^L, but for application purposes we separated it into its own variable. Now m equals

$$m = \frac{t^*}{e} \tag{1.34}$$

where e is the natural exponent. Therefore, mp^{H^*} will be

$$mp^{H^*} \in \{P(U^p = H)_1, \ldots, P(U^p = H)_m\} \tag{1.35}$$

$$\text{such that } P(U^p = H)_i \leq mp^{H^*}, \tag{1.36}$$

$$\forall P(U^p = H)_i \in \{P(U^p = H)_1, \ldots, P(U^p = H)_m\} \tag{1.37}$$

If all the $P(U^p = H)$'s in a multi-period investment were unique, then this stragey will identify the highest value (less then mp^h), 37% of the time Gilbert and Mosteller (2006). Since $\alpha > (1 - \alpha)$, this implies that the POMDP model will have a high concentration of $P(U^p = H)$'s, in a range of larger values, and so it's probability of finding the max value should be greater then 37%. Therefore, the entrepreneur's strategy is to continue the investment until $P(U^p = H) \geq mp^H$ in the first m months. If that condition is met, exploit the investment. If after m periods, the condition $P(U^p = H) \geq mp^H$ has not been met, then continue the investment until $P(U^p = H) \geq mp^{H^*}$ (the highest value in the first m periods) and then stop the investment. If the expected profit is desirable then exploit, if not abandon. If $P(U^p = H) \geq mp^{H^*}$ is not satisfied, then abandon the investment after period t^*.

1.2   Example

Consider the case of a two-state POMDP problem, where an entrepreneur creates a prototype of a widget that users can pay to use. The widget costs \$100 a month to run and manage, but a commercial version, that is stable and can be scaled, needs \$10,000 additional invested. The monthly payments are sunk costs once they are made and once the entrepreneur invests the \$10,000 that cannot be recuperated. If the monthly paid user base (MPU) can get to 7500, then the entrepreneur will make \$20,000, on the commercial version of the new widget, but if the number of paid

users does not make it to 7500, the value of the investment will be a loss of $9000. Time is not a limiting factor, after the completion of the commercial grade widget, the demand will be realized instantaneously. The unobservable core process has two states, 7500 and 0 (any value less then 0 is considered 0 and any value greater then 7500 can be considered 7500).

$$U^p = \{7500, 0\}$$

$$V^H = 20,000$$

$$V^L = -9,000$$

The entrepreneur knows from similar products, that during the prototype stage the number of users that will become MPU's tend to follow these conditional probabilities.

$$P(U_t^p = 7500 | U_{t-1}^p = 7500) = 90\%$$

$$P(U_t^p = 0 | U_{t-1}^p = 7500) = 10\%$$

$$P(U_t^p = 0 | U_{t-1}^p = 0) = 40\%$$

$$P(U_t^p = 7500 | U_{t-1}^p = 0) = 60\%$$

This implies that if users decide to become MPUs, most likely they won't change their decision, where users who don't want to use the widget or who haven't made up their minds have more uncertainty. The entrepreneur can observe how many new paid users join the prototype, and can only have two values, 80 or 10, where 80 is

considered a positive sign of user adoption and 10 is considered a negative sign.

$$u^p = \{80, 10\}$$

$$P(u_t^p = 80|U_t^p = 7500) = 75\%$$

$$P(u_t^p = 10|U_t^p = 7500) = 25\%$$

$$P(u_t^p = 10|U_t^p = 0) = 55\%$$

$$P(u_t^p = 80|U_t^p = 0) = 45\%$$

The entrepreneur has information on what the users initial reaction to the widget will be, therefore the initial values of U_0^p are

$$P(U_0^p = 7500) = 50\%$$

$$P(U_0^p = 0) = 50\%$$

We generate observations for three scenarios. The first scenario generates observations for just the $U^p = 7500$ unobservable state, the second scenario generates observations for just the $U^p = 0$ unobservable state, and the third scenario generates observations for the unobservable process following the probability transitions. We generate 30

observations for each scenario.

Scenario 1: $10, 80, 10, 10, 80, 80, 10, 80, 10, 80, 80, 80, 80, 80, 80,$

$10, 10, 80, 10, 80, 80, 80, 10, 80, 10, 80, 80, 80, 80, 80$

Scenario 2: $10, 80, 80, 80, 10, 10, 80, 80, 80, 80, 10, 80, 80, 10, 80,$

$10, 10, 80, 10, 80, 10, 10, 10, 10, 80, 10, 80, 10, 80, 10$

Scenario 3: $80, 80, 80, 10, 80, 80, 10, 80, 80, 10, 80, 80, 80, 80, 80,$

$80, 10, 80, 80, 80, 80, 10, 80, 10, 10, 80, 10, 10, 80, 80$

core: $7500, 7500, 0, 7500, 7500, 7500, 0, 7500, 7500, 7500,$

$7500, 7500, 7500, 7500, 7500, 0, 7500, 7500, 7500, 7500,$

$7500, 7500, 7500, 7500, 0, 7500, 7500, 7500, 7500, 7500$

Using the equations from 1.26 and 1.27 we calculate the value of the probabilities of the two states of the MPU, at which point no more value is being gained by the multi-period investment.

$$P(U^p = 7500) = 92.2\%$$
$$P(U^p = 0) = 7.8\%$$

Running through each scenario we stop the investment when the probability of the MPU being 7500 is 92.2%, and exploit the commercial grade option.

Scenario 1: $57.7, 85, 72.8, 67.2, 87.1, 91.2, 75.8, 88.9, 74.7, 88.6, 91.5, 92.1, 92.2, 92.2, 92.2,$
$76.4, 68.8, 87.4, 74, 88.5, 91.5, 92.1, 76.3, 89, 74.8, 88.7, 91.5, 92.1, 92.2, 92.2$

Scenario 2: $57.7, 85, 90.8, 91.9, 76.2, 68.7, 87.4, 91.2, 92, 92.2, 76.3, 89, 91.6, 76, 88.9,$
$74.7, 68.1, 87.3, 73.9, 88.5, 74.5, 68, 65.1, 63.8, 86.4, 73.5, 88.4, 74.5, 88.6, 74.6$

Scenario 3: $83.3, 90.4, 91.9, 76.2, 89, 91.6, 76, 88.9, 91.6, 76, 88.9, 91.6, 92.1, 92.2, 92.2,$
$92.2, 76.4, 89, 91.6, 92.1, 92.2, 76.3, 89, 74.8, 68.1, 87.3, 73.9, 67.7, 87.2, 91.2$

In Scenario 1, the entrepreneur would have stopped the investment in month 13 and because the expected profit is \$6,438 the option is exploited and the entrepreneur makes \$8,700. In Scenario 2, the entrepreneur stops the investment in month 10, with an expected profit of \$6,738. The entrpreneur executes the deal and loses \$11,000. In Scenario 3, the investment is stopped in 14 months and with an expected profit of \$6,338, the investment is exploited and the entrepreneur makes \$8,600. The last period t^*, of this multi-period investment would be 77, and the number of period m to search for an adjusted mp^{H^*} would be 28. In the 3 scenarios we had no need for the mp^{H^*} because the mph was reached, before m, for each scenario.

1.3 Simulation

Two simulations are run, in the first simulation 1.2, the transitional and observational probabilities are identical to the example. In the second simulation 1.3, the transition

Simulation 1				
V^L	Mean	STD	Max	Min
0	-816.77	7462.75	6300	-19800
500	-616.04	7502.71	6400	-20200
1000	-654.98	7545.28	6400	-20600
1500	-689.44	7586.33	6400	-21000
2000	-514.17	7642.76	6500	-21300
2500	-543.3	7686.43	6500	-21800
3000	-568.52	7725.58	6500	-22200
3500	-437.95	7829.86	6600	-22600
4000	-470.49	7879.97	6600	-23000
4500	-279.53	7951.52	6700	-23400
5000	-315.05	8007.05	6700	-23800
5500	-120.06	8052.53	6800	-24100
6000	-147.48	8102.44	6800	-24500
6500	-181.87	8158.35	6800	-24900
7000	3.53	8216.13	6900	-25200
7500	-27.72	8261.59	6900	-25600
8000	186.51	8357.67	7000	-26000
8500	162.76	8379.43	7000	-26300
9000	320.49	8484.95	7100	-26700
9500	6810.38	8172.58	9600	-22500

Table 1.2: Simulation Results where $P(U_t^p = 7500|U_{t-1}^p = 7500) = 90\%$

probabilities to:

$$P(U_t^p = 7500|U_{t-1}^p = 7500) = 100\%$$

$$P(U_t^p = 0|U_{t-1}^p = 7500) = 0\%$$

Both simulations are run with V^L 20 different values ranging from 0 to 9500, in 500 increments. At each V^L price point, 10,000 iterations are run and averaged togther.

The MP optimal stopping point, for the first set of assumptions, 1.2 on average yields negative results. It's interesting to note, that at as V^L decreased, the performance got worse. This implies that our heuristic, borrowing from Gilbert and Mosteller (2006), was not effective in stopping the POMDP model optimally, as the mp^H and mp^{H^*}

Simulation 2				
V^L	Mean	STD	Max	Min
0	7659.86	5858.62	9700	-11100
500	7553.8	6095.98	9700	-11500
1000	7434.75	6365.58	9700	-12000
1500	7246.25	6714.1	9700	-12500
2000	7073.51	7052.32	9700	-13000
2500	6797.74	7489.76	9700	-13600
3000	6770.07	7661.78	9800	-14000
3500	6698.01	7838.99	9800	-14400
4000	6624.94	8016.75	9800	-14900
4500	6557.65	8188.71	9800	-15200
5000	6454.73	8401.2	9800	-15700
5500	6373.26	8591.39	9800	-16200
6000	6272.61	8800.35	9800	-16700
6500	6137.98	9047.05	9800	-17200
7000	6070.68	9217.69	9800	-17700
7500	5852.45	9551.99	9800	-18200
8000	5696.41	9823.27	9800	-18700
8500	5624.61	9998.61	9800	-19200
9000	5200.47	10529.91	9800	-19700
9500	5122.42	10711.39	9800	-20200

Table 1.3: Simulation Results where $P(U_t^p = 7500 | U_{t-1}^p = 7500) = 100\%$

were too high. This can be seen by the difference in the max and mins, which are indicative of the lopsided loses taken on by the entrepreneur. The results in 1.3 are in line with the assumptions made. With high probable chance that $P(U_t^p = 7500$, the odds of profit were for the entrepreneur.

CHAPTER 2

Network Option Decision Process

2.1   Two State NODP Model

The Network Option Decision Process (NODP) Model does not assume any known or prior probabilities, instead it observes two different performance indicators, new paid users (NPU) u_t^p and new free users (NFU) u_t^f, during each period t. The NFU u_t^f indicates the size and depth of the market that the multi-stage investment attempts to capture, and the NPU u_t^p indicates the success rate of the investment in capturing said market, in terms of permanent paid users (PPU) U^p. As previously stated for the two-state model, the NFU can only take on the two values $u_t^f = \{g, k\}$ at period t where $g > k$, and the NPU can only take on the two values $u_t^p = \{h, l\}$ at period t where $h > l$. Figure 2.1, shows the transitional diagram for the two-state NODP model.

The POMDP model has three choices like the NODP model, abandon, exploit or continue to learn. The difference here is that the NODP model has two separate exploitation stages x^1 and x^2. The first stage of the process focuses on just the NFU u_t^f and whether to continue seeing the results for it or abandoning the multi-stage investment or to exploit the next stage or level (stage 2) of the multi-stage investment, x^1. Once the next the stage, stage 2, is exploited then it becomes similar to the POMDP model where the entrepreneur must decide to continue observing the NPU u_t^p, aban-

doning the multi-stage investment or exploiting the final opportunity. The difference here is that the NODP does not assume the observational or transitional probabilities in the POMDP models, but it continues to observe the NFU u_t^f along with the NPU u_t^p simultaneously. The logic behind the NODP model is that steady or exponential growth between the NPU u_t^p and NFU u_t^f, implies a PPU of high valuation, $U^p = H$, whereas negative growth between the NPU u_t^p and NFU u_t^f observations implies a PPU of $U^p = L$. The probability of the PPU $P(U^p)$ is dependent on the combination of observations between the NPU u_t^p and NFU u_t^f from periods 1 to t.

$$P(U_{t+1}^p = H | (u_t^p, u_t^f), (u_{t-1}^p, u_{t-1}^f), \ldots, (u_1^p, u_1^f)) = \gamma \tag{2.1}$$

$$P(U_{t+1}^p = L | (u_t^p, u_t^f), (u_{t-1}^p, u_{t-1}^f), \ldots, (u_1^p, u_1^f)) = 1 - \gamma \tag{2.2}$$

Since our two state model has only two states for both u_t^p and u_t^f, then in period t any four combinations are possible

$$(u_t^p, u_t^f) = \{(h, g), (h, k), (l, k), (l, g)\}$$

$$\text{where } h > l, g > k$$

To translate these four combinations into a probability the NODP model uses an algebraic variation of the logistic function, which is from the family of sigmoid functions used to model the non-linearity components of complex systems. The logistic function will allow us to approximate the probability after every observation of the NPU u^p and NFU u^f. The NODP model is based on the principle that despite not having a historical probability distribution, good investments are made based on the

growth rate of the adoption of the different stages of the investment. The two-state NODP model has two stages, and the relationship, (h, g), describes an exponential growth in both stages implying the market and the adoption of the market to the entrepreneur's product or service are both growing. The relationship (l, k), implies that there is a very small market for the problem the entrepreneur is trying to service and little adoption of the product or service that the entrepreneur is providing. It is important to note that the second stage of the NODP model is exploited after a sufficient series of NFU's $u_t^f = g$ are observed based on the cost of each period of observing the NFU's u_t^f versus the potential reward of each observation of $u_t^f = g$. We approach this problem by starting in stage two of the model, which implies that a sufficient number of $u_t^f = g$ were observed in stage one. The combination of (l, g) implies that there is growth in the target market, but the product or service is not being adopted by the target market. The combination of (h, k) is more complicated to translate because it is conditional on the stage and size of the investment and the nature of the observations of u_t^f and u_t^p themselves. Combination (h, k), implies that there is a growth in adoption of the product or service but no growth in the market itself. First, if the multi-stage investment is in the initial periods then this implies that the market being serviced is a small subset of the target market and the product or service is sucessfully being adopted by this niche market. We define a period as an initial-period when the total free users (TFU) is less then the target paid-users (TPU). The TPU is a guideline number that investors define as a necessary condition for exploiting the second-stage of the multi-stage investment, and the rate at which the NPU u_t^p can grow to the TPU indicates the certainty at which unobservable PPU U^p will reach state H and therefore profitability. The TPU is less then the PPU because the current infrastructure of the product or service cannot accomodate the PPU at its desired state and significant investments must be made, hence the need to know when or when not to exploit the second-stage of the multi-period investment.

Investors quantify the TPU based on the combination of exploitation costs versus the cost of continuing the exploration phase of the investment and quarterly revenue that the TPU generates. The TPU revenue does need to be profitable it just needs to be large enough to justify further investment.

The combination of (h, k) in the post-initial periods (periods where TFU $\geq$ TPU) of the multi-period investment implies that there is growth and adoption of the entrepreneur's product or service, but despite the NFU u_t^f having positive growth rates in the initial-periods, the lack of subsequent growth past the TFU level indicates lack of support for addtional growth of the NPU u_t^p will continue beyond exploitation. In addtion, this could be a sign that the NPU u_t^p may or not may be correlated with the NFU u_t^f, which makes the observations obsolete when trying to predicte the probability of the PPU U^p. In practice, if the growth is coming from users that upgraded from the freemium service to the paid service, and not from new users that never were part of the freemium service, then the market is considered to be smaller then initially assumed and the product or service is proven to be succesful at servicing the niche market. If the latter is the case, then this could be indicative of the product or service being extremely successful at capturing the market because the new users are going straight to purchasing the paid feature, an extremely rare scenario in the industry. For the purposes of our two-state NODP model we assume the combination (h, k) as a neutral combination that shows no correlation between growths and therefore does not affect the probability of U^p. Table 2.1 summarizes the combinations and the marginal gains/losses in probability implied by each combination.

We define our algebraic logistic-based probability function as a percentage of two variables, Υ^H and Υ^L each one representing a possible state of the PPU U^p. The probability of the two different possible states of the PPU, in period t, is defined

Two-state Combinations		
Combination	Result	Incremental Probability (Increase/Decrease)
(h, g)	strong growth in freemium and paid service	increase success probability
(h, k)	not correlated, random growth	no affect on probability
(l, g)	freemium adoption but not paid service	medium decrease in success probability
(l, k)	small growth in freemium and paid service	strong decrease in success probability

Table 2.1: List of Combinations

as

$$P(U_t^p = H) = \frac{\Upsilon_t^H}{\Upsilon_t^H + \Upsilon_t^L} \tag{2.3}$$

$$P(U_t^p = L) = \frac{\Upsilon_t^L}{\Upsilon_t^H + \Upsilon_t^L} \tag{2.4}$$

The initial values of Υ_0^H and Υ_0^L at period $t = 0$ of the multi-period investment are equal and greater then one

$$\Upsilon_0^H = \Upsilon_0^L$$

$$\text{where } \Upsilon_0^H \geq 1 \text{ and } \Upsilon_0^L > 1$$

Investors typically target a minimum rate of return (MRR), given that the TPU for the NODP model is assumed, we can calculate the longest period of time, t^*, that a benchmark combination of the NPU $u_t^p = b$ and the NFU $u_t^f = g$, consectively observed (best case scenario), takes for the multi-period investment to reach the target TPU, where $b < h$. If h were to be less then b, then that would imply that under the best case scenario, the TPU would not be reachable under circumstances

that would satisfy the MRR and the investment would not be prudent. Also, the higher the h the faster the rate at attaining the TPU and the greater the probability implied by our model. If the NPU observation is a constant random whole number then t^* can be calculated by simply dividing the V^H by MRR times C, where V^H is the revenue generated if the PPU $U^p = H$ and the C is the cost of each period of investment.

$$\lfloor t^* \rfloor = \frac{V^H}{(1 + MRR)C} \tag{2.5}$$

Using t^*, we can then calculate b, by dividing the TPU by t^*

$$b = \frac{TPU}{\lfloor t^* \rfloor} \tag{2.6}$$

If h represent a growth ratio of paid users NPU, then define TFU* as the amount of new free users acquired before the investor exploited the second stage of the multi-period investment. The maximum number of period t^* would still be calculated using equation 2.5 and then b would be calculated using

$$\text{TFU}^*(1 + b)(1 + b) \dots (1 + b) = TPU$$

$$\text{TFU}^*(1 + b)^{t^*} = TPU$$

$$(1 + b)^{t^*} = \frac{TPU}{\text{TFU}^*}$$

$$(1 + b) = \left(\frac{TPU}{\text{TFU}^*} \right)^{\frac{1}{t^*}}$$

with the final equation

$$b = \left(\frac{TPU}{\text{TFU}^*} \right)^{\frac{1}{t^*}} - 1 \tag{2.7}$$

Function Values				
Combinations	O^H	O^L	D^H	D^L
(h, g)	$\frac{h}{g}$	-	g	-
(h, k)	-	-	-	-
(l, g)	-	$\frac{l}{g}$	-	g
(l, k)	-	$\frac{\frac{h}{g}}{g}$	-	g

Table 2.2: Two-state model values for functions $D()$ and $O()$

We define the formula that incrementally changes the values of Υ_t^H and Υ_t^L with each combination of observations u_t^f and u_t^p as

$$\Upsilon_t^i = \Upsilon_0^i + \frac{\omega^{n_t^c O^c(u_t^p, u_t^f)}}{n_t^c D^c(u_t^p, u_t^f)} \tag{2.8}$$

Where i are the possible states for U^p and c are the possible combination of NPU and NFU observations. In our two-state model i can be H and L, and our combinations are listed in table 2.1. The variable n_t^c is the number of times the unique combinations of the NPU and NFU (u_t^p, u_t^f) have been observed up to period t. The functions $O^i(u_t^p, u_t^f)$ and $D^i(u_t^p, u_t^f)$ map the observed combinations to numeric values to approximate desired sigmoid shapes. For the purpose of our two-stage model the values of h, g, k, and l are sufficient for our sigmoid probability distribution, table 2.2 summarizes the functions values for each observed combination.

The values in table 2.2 directly relate to the desired probability increases and decreases from table 2.1, by changing the values of Υ^H and Υ^L with every combination observed. For our two-state model, the observation of (h, g) increases the value of Υ^H thereby increasing the probability of $U^P = H$. The combination of (h, k) is defined to have no effect, and the combinations of (l, g) and (l, k) increase the value of Υ^L by different rates, one implying a faster increase then the other. The structure of equation 2.8 shows that the values of Υ^H and Υ^L will grow exponentially which gives the probability the sigmoid shape the NODP model assumes.

The value of ω determines the speed, and subsequently the slope at the x origin, at which the logistic curve accelerates from the lower bound, which at period 0 is 50% as seen in the example logistic curve in figure 2.2. Using the value of t^*, we can optimize the value of ω to approach the upper asymptotic boundary at the approximate rate that corresponds to the best-case scenario of our two-state model. If the best-case combination is observed consecutively in the first t^* periods, then the TPU will be reached and the entrepreneur will exploit the opportunity. It is important to note that for our two-state model we have simplified the TPU to a constant number, but in the industry the TPU tends to be a rate of targeted paying-users per a defined period of time or customer-acquisition cost. This can also be accomodated in our model as long as the TPU, the NPU u^p and the NFU u^f have equivalent units. Additionally, the TPU by itself is not a sufficient methodology for signaling an optimal exploitation point. The path to the TPU is more important, whether the performance indicators, NPU u^p and NFU u^f, are growing towards the TPU or are their trajectories more indicative of negative growth. For this reason, we calculate the longest number of periods t^* required to reach the TPU from 0 (period 0), which is defined as consecutively observing the unique-case scenarios of the benchmark b combination (b, g). This gives the lowest rate b possible for the longest period of time t^* for our two-state model to have the highest probability of indicating the PPU $U^p = H$. Therefore, by dividing equation 2.8, by the summation of itself plus Υ^L and setting that equal to a very high probability of certainty (for example 99%), defined here as φ, and setting n_t^c equal to t^*, $O(u_t^p, u_t^f)$ equal to $\frac{b}{g}$, and $D^c(u_t^p, u_t^f)$ equal to g we get

$$\frac{\Upsilon_0^H + \dfrac{\omega^{t^*(\frac{b}{g})}}{t^* g}}{\Upsilon_0^L + \left(\Upsilon_0^H + \dfrac{\omega^{t^*(\frac{b}{g})}}{t^* g} \right)} = \varphi$$

At period 0, Υ^H equals Υ^L, so we can replace Υ^L with Υ^H and reduce the amount of

variables.

$$\Rightarrow \frac{\Upsilon_0^H + \dfrac{\omega^{t^*\left(\frac{b}{g}\right)}}{t^*g}}{2\Upsilon_0^H + \dfrac{\omega^{t^*\left(\frac{b}{g}\right)}}{t^*g}} = \varphi$$

$$\Rightarrow \frac{\dfrac{\Upsilon_0^H t^*g + \omega^{t^*\left(\frac{b}{g}\right)}}{t^*g}}{\dfrac{2\Upsilon_0^H t^*g + \omega^{t^*\left(\frac{b}{g}\right)}}{t^*g}} = \varphi$$

$$\Rightarrow \frac{\Upsilon_0^H t^*g + \omega^{t^*\left(\frac{b}{g}\right)}}{2\Upsilon_0^H t^*g + \omega^{t^*\left(\frac{b}{g}\right)}} = \varphi$$

$$\Rightarrow \Upsilon_0^H t^*g + \omega^{t^*\left(\frac{b}{g}\right)} = 2\Upsilon_0^H t^*g\varphi + \omega^{t^*\left(\frac{b}{g}\right)}\varphi$$

$$\Rightarrow \omega^{t^*\left(\frac{b}{g}\right)}(1-\varphi) = 2\Upsilon_0^H t^*g\varphi - \Upsilon_0^H t^*g$$

Which can be reduced to

$$\omega^* = \left(\frac{2\Upsilon_0^H t^*g\varphi - \Upsilon_0^H t^*g}{1-\varphi}\right)^{\frac{g}{bt^*}} \tag{2.9}$$

We can now use ω^* in equation 2.8 for every combination that is observed. The magnitude of Υ_0^H and Υ_0^L at period 0 controls the length of the lower bound asymptote of the sigmoid model before exponential growth is observed. This is to account for the assumption that the first consecutive observations of any type of combination account for very little in terms of evidence of growth. Only after a longer pattern of consecutive observations are observed does the sigmoid curve exponentially grow and then starts to decline towards the upper bound. The larger the value chosen for Υ_0^H and Υ_0^L at period 0, the longer the lower bound asymptote and the more consecutive observations needed to be indicative for a sign of growth. In Figure 2.3, the red graph has a larger Υ_0^H and Υ_0^L in period 0.

The expected profit function in period 0, for a risk-neutral investor, is defined as

$$\pi_0 = V^H \left[\frac{\Upsilon_0^H}{\Upsilon_0^H + \Upsilon_0^L} \right] + V^L \left[\frac{\Upsilon_0^L}{\Upsilon_0^H + \Upsilon_0^L} \right] \tag{2.10}$$

where both initial ratios at period 0 are 50%. Just like in the POMDP Model, the investor spends C_1 to observe the NPU u_1^p in period 1, however unlike the POMDP Model, the investor also observes the NFU u_1^f. For comparison sake, we don't separate out the cost for the NFU u_1^f. This isn't a large assumption, since most freemiums in the industry by definition are cheap proof-of-concepts. In addition, the POMDP Model does have costs associate with quantifying the transitional and observation probabilities that are also not separated out. The investor now has to choose whether to abandon, continue or exploit the multi-period investment. If combination $(u_1^p = h, u_1^f = g)$ is observed, then first the Υ's are updated

$$\Upsilon_1^H = \Upsilon_0^H + \frac{\omega^{*\frac{h}{g}}}{g} \tag{2.11}$$

$$\Upsilon_1^L = \Upsilon_0^L \tag{2.12}$$

If the investor chooses to exploit the investment then the expected profit function for period 1 is

$$\pi_1^{(h,g)} = V^H \left[\frac{\Upsilon_1^H}{\Upsilon_1^H + \Upsilon_1^L} \right] + V^L \left[\frac{\Upsilon_1^L}{\Upsilon_1^H + \Upsilon_1^L} \right] - C_1 \tag{2.13}$$

If combination $(u_1^p = l, u_1^f = k)$ is observed then the Υ's are updated as such

$$\Upsilon_1^H = \Upsilon_0^H \tag{2.14}$$

$$\Upsilon_1^L = \Upsilon_0^L + \frac{\omega^{*\frac{h}{g}}}{g} \tag{2.15}$$

Where as if combination $(u_1^p = l, u_1^f = g)$ is observed then

$$\Upsilon_1^H = \Upsilon_0^H \tag{2.16}$$

$$\Upsilon_1^L = \Upsilon_0^L + \frac{\omega^{*\frac{l}{g}}}{g} \tag{2.17}$$

Combination $(u_1^p = h, u_1^f = k)$ is considered an outlier and as such no update is made to the Υ's. The expected profit function, equation 2.13, for period 1 remain the same as long as the proper algebraic updates are made to the Υ's. If the investor decides to continue the multi-period investment to period 2, then the expected profit function, after observing combination c^2 (assume that in period 1 the combination $(h, g)_1$, equation 2.11, is observed) will generalize to

$$\pi_2^{(h,g)_1, c^2} = V^H \left[\frac{\Upsilon_2^H}{\Upsilon_2^H + \Upsilon_2^L} \right] + V^L \left[\frac{\Upsilon_2^L}{\Upsilon_2^H + \Upsilon_2^L} \right] - C_2 \tag{2.18}$$

The Υ's, in period 2 will update accordingly, if combination $(u_2^p = h, u_2^f = g)$ is observed then

$$\Upsilon_2^H = \Upsilon_0^H + \frac{\omega^{*\frac{2h}{g}}}{2g} \tag{2.19}$$

$$\Upsilon_2^L = \Upsilon_0^L \tag{2.20}$$

If combination $(u_2^p = l, u_2^f = k)$ is observed then

$$\Upsilon_2^H = \Upsilon_0^H + \frac{\omega^{*\frac{h}{g}}}{g} \tag{2.21}$$

$$\Upsilon_2^L = \Upsilon_0^L + \frac{\omega^{*\frac{h}{g}}}{g} \tag{2.22}$$

If combination $(u_2^p = l, u_2^f = g)$ is observed then

$$\Upsilon_2^H = \Upsilon_0^H + \frac{\omega^{*\frac{h}{g}}}{g} \tag{2.23}$$

$$\Upsilon_2^L = \Upsilon_0^L + \frac{\omega^{*\frac{l}{g}}}{g} \tag{2.24}$$

and if combination $(u_2^p = h, u_2^f = k)$ is observed then no change is made. Unlike the POMDP Model we do not have transition probabilities to update, but the dilemma facing the NODP Model is that since there are no observations and transitional probabilities, we cannot calculate a MP stopping point to use as a heuristic to signal when to abandon, exploit or continue the multi-period investment. Therefore, to optimize decision making in the NODP Model we apply a best-case (BC) stopping limit heuristic. Given the profit function at the end of period t

$$\pi_t^c = V^H \left[\frac{\Upsilon_t^H}{\Upsilon_t^H + \Upsilon_t^L} \right] + V^L \left[\frac{\Upsilon_t^L}{\Upsilon_t^H + \Upsilon_t^L} \right] - C_t$$

we find if there exists a local maximum profit function, $\pi_t^o m$, that is greater then π_t^c, if only the most optimal combination (h, g) is consecutively observed (best case scenario) from periods $t + 1$ to tm, where tm is the period of the local maximum profit function under the best case scenario. The variable o is defined as $tm - t$ and represents the number of consecutive combinations (h, g) that are needed for $\pi_t^o m > \pi_t^c$, if $\pi_t^o m$ exists. Therefore the profit function for the local maximum is

$$\pi_t^o m = V^H \left[\frac{\Upsilon_t^H m}{\Upsilon_t^H m + \Upsilon_t^L m} \right] + V^L \left[\frac{\Upsilon_t^L m}{\Upsilon_t^H + \Upsilon_t^L m} \right] - C_t m$$

where the Υ's are defined as

$$\Upsilon_{tm}^H = \Upsilon_0^H + \frac{\omega^{*\frac{tmh}{g}}}{tmg} \tag{2.25}$$

$$\Upsilon_{tm}^L = \Upsilon_t^L \tag{2.26}$$

The logic here is that if under the best-case scenario going forward there will not be
an expected value in the future greater then present expected value, then the multi-
period investment should be stopped. At this stopping point, if the expected-profit is
greater then the MRR, then the multi-periond investment should be exploited. If not,
then the investment should be abandoned. The solve for the local maxima at time tm
would require the derivative of equation 2.25 with respect to tm, which would require
the use of the Lambert function. We propose instead to use an algorithm take an
estimated range of tm. First, for the general profit function

$$\pi = V^H \varphi + V^L (1 - \varphi) - C(tm) \tag{2.27}$$

where φ is the probability at which point the profit function is optimal given the
linear cost of the multi-period options C, here defined as a fixed cost. To solve for
equation 2.27, we have to define tm in terms of φ, which requires solving

$$\frac{\Upsilon_0^H + \frac{\omega^{*\frac{th}{g}}}{gt}}{\Upsilon_t^L + \Upsilon_0^H + \frac{\omega^{*\frac{th}{g}}}{gt}} = \varphi \tag{2.28}$$

in term of t. Given the ratio $\frac{\omega^{*\frac{th}{g}}}{gt}$, to solve this would again require the use of the
Lambert function. To avoid this, we define an upper bound t^Q and lower bound t^q
for t, by replacing the gt in equation 2.28, with TPU/b, defined as g^*, and g. The
TPU divided by b gives a longer time then is necessary for the optimal combination

to achieve the same probability, where as g by itself will give a value smaller then our target tm.

$$\Rightarrow \frac{\Upsilon_0^H + \frac{\omega^{*\frac{t^Q h}{g}}}{g^*}}{\Upsilon_t^L + \Upsilon_0^H + \frac{\omega^{*\frac{t^Q h}{g}}}{g^*}} = \varphi$$

$$\Rightarrow \Upsilon_0^H + \frac{\omega^{*\frac{t^Q h}{g}}}{g^*} = \varphi\left[\Upsilon_t^L + \Upsilon_0^H + \frac{\omega^{*\frac{t^Q h}{g}}}{g^*}\right]$$

$$\Rightarrow \frac{\omega^{*\frac{t^Q h}{g}}}{g^*} = \varphi\Upsilon_t^L + \varphi\Upsilon_0^H + \varphi\frac{\omega^{*\frac{t^Q h}{g}}}{g^*} - \Upsilon_0^H$$

$$\Rightarrow \frac{\omega^{*\frac{t^Q h}{g}}}{g^*} = \varphi\Upsilon_t^L + \varphi\Upsilon_0^H + \varphi\frac{\omega^{*\frac{t^Q h}{g}}}{g^*} - \Upsilon_0^H$$

$$\Rightarrow \frac{\omega^{*\frac{t^Q h}{g}}}{g^*}(1 - \varphi) = \varphi\Upsilon_t^L + \varphi\Upsilon_0^H - \Upsilon_0^H$$

$$\Rightarrow \omega^{*\frac{t^Q h}{g}} = \frac{g^*(\varphi\Upsilon_t^L + \varphi\Upsilon_0^H - \Upsilon_0^H)}{(1 - \varphi)}$$

$$\Rightarrow \frac{t^Q h}{g} = \log_\omega\left(\frac{g^*(\varphi\Upsilon_t^L + \varphi\Upsilon_0^H - \Upsilon_0^H)}{(1 - \varphi)}\right)$$

$$t^Q = \frac{g}{h}\log_\omega\left(\frac{g^*(\varphi\Upsilon_t^L + \varphi\Upsilon_0^H - \Upsilon_0^H)}{(1 - \varphi)}\right) \tag{2.29}$$

Equivalently, t^q can be defined by equation 2.29, with g in place of g^*

$$t^q = \frac{g}{h}\log_\omega\left(\frac{g(\varphi\Upsilon_t^L + \varphi\Upsilon_0^H - \Upsilon_0^H)}{(1 - \varphi)}\right) \tag{2.30}$$

It is evident to see that g and g^* have a direct relationship with t and since $g < g^*$, t^q is the lowerbound for tm and t^Q is the upperbound. The upperbound and lowerbound equations are substituted for tm in equation 2.27, which allows the general profit

function to be solved for the local maxima, with respect to φ.

$$\underset{\varphi}{\text{maximize }} V^H \varphi + V^L(1-\varphi) - C\frac{g}{h}\log_\omega\left(\frac{g(\varphi\Upsilon_t^L + \varphi\Upsilon_0^H - \Upsilon_0^H)}{(1-\varphi)}\right) \qquad (2.31)$$

The derivative of the first two parts of equation 2.31 are trivial, but to solve the the third part requires that the equation be written in terms of natural log and the use of the quotient rule.

$$\Rightarrow C\frac{g}{h\log(\omega)}\log\left(\frac{(g\varphi\Upsilon_t^L + g\varphi\Upsilon_0^H - g\Upsilon_0^H)}{(1-\varphi)}\right)'$$

$$\Rightarrow C\frac{g}{h\log(\omega)}\left[\frac{(1-\varphi)}{(g\varphi\Upsilon_t^L + g\varphi\Upsilon_0^H - g\Upsilon_0^H)}\right]\left[\frac{(1-\varphi)(g\Upsilon_t^L + g\Upsilon_0^H) + (g\varphi\Upsilon_t^L + g\varphi\Upsilon_0^H - g\Upsilon_0^H)}{(1-\varphi)^2}\right]$$

$$\Rightarrow C\frac{g}{h\log(\omega)}\left[\frac{1}{(g\varphi\Upsilon_t^L + g\varphi\Upsilon_0^H - g\Upsilon_0^H)}\right]\left[\frac{g\Upsilon_t^L + g\Upsilon_0^H - g\Upsilon_t^L\varphi - g\Upsilon_0^H\varphi + g\Upsilon_t^L\varphi + g\Upsilon_0^H\varphi - g\Upsilon_0^H}{(1-\varphi)}\right]$$

$$\Rightarrow C\frac{g}{h\log(\omega)}\left[\frac{1}{(g\varphi\Upsilon_t^L + g\varphi\Upsilon_0^H - g\Upsilon_0^H)}\right]\left[\frac{g\Upsilon_t^L + g\Upsilon_0^H - g\Upsilon_t^L\varphi - g\Upsilon_0^H\varphi + g\Upsilon_t^L\varphi + g\Upsilon_0^H\varphi - g\Upsilon_0^H}{(1-\varphi)}\right]$$

$$\Rightarrow C\frac{g}{h\log(\omega)}\left[\frac{g\Upsilon_t^L}{g\Upsilon_t^L\varphi + g\Upsilon_0^H\varphi - g\Upsilon_0^H - g\Upsilon_t^L\varphi^2 - g\Upsilon_0^H\varphi^2 + g\Upsilon_0^H\varphi}\right]$$

We can now solve equation 2.31 for the maximum φ.

$$\Rightarrow V^H - V^L = C\frac{g}{h\log(\omega)}\left[\frac{\Upsilon_t^L}{\Upsilon_t^L\varphi + \Upsilon_0^H\varphi - \Upsilon_0^H - \Upsilon_t^L\varphi^2 - \Upsilon_0^H\varphi^2 + \Upsilon_0^H\varphi}\right]$$

$$\Rightarrow \Upsilon_t^L\varphi + \Upsilon_0^H\varphi - \Upsilon_0^H - \Upsilon_t^L\varphi^2 - \Upsilon_0^H\varphi^2 + \Upsilon_0^H\varphi = C\frac{g\Upsilon_t^L}{h\log(\omega)(V^H - V^L)}$$

$$\Rightarrow -(\Upsilon_t^L + \Upsilon_0^H)\varphi^2 + (\Upsilon_t^L + 2\Upsilon_0^H)\varphi = C\frac{g\Upsilon_t^L}{h\log(\omega)(V^H - V^L)} + \Upsilon_0^H$$

Using the quadratic formula we solve for φ^q and φ^Q, and since the g inside the brackets corresponds to g and g^* of our upper and lower bounds cancel out we are left with

just φ^q and φ^Q for both t^q and t^Q.

$$\varphi^Q = \frac{-(\Upsilon_t^L + 2\Upsilon_0^H) + \sqrt{(\Upsilon_t^L + 2\Upsilon_0^H)^2 - 4(\Upsilon_t^L + \Upsilon_0^H)\left[C\dfrac{g\Upsilon_t^L}{h\log(\omega)(V^H - V^L)} + \Upsilon_0^H\right]}}{-2(\Upsilon_t^L + \Upsilon_0^H)}$$

(2.32)

$$\varphi^q = \frac{-(\Upsilon_t^L + 2\Upsilon_0^H) - \sqrt{(\Upsilon_t^L + 2\Upsilon_0^H)^2 - 4(\Upsilon_t^L + \Upsilon_0^H)\left[C\dfrac{g\Upsilon_t^L}{h\log(\omega)(V^H - V^L)} + \Upsilon_0^H\right]}}{-2(\Upsilon_t^L + \Upsilon_0^H)}$$

(2.33)

Then substituting φ^Q and φ^q for φ in the profit function

$$\pi(\varphi) = V^H\varphi + V^L(1 - \varphi) - C\frac{g}{h}\log_\omega\left(\frac{g(\varphi\Upsilon_t^L + \varphi\Upsilon_0^H - \Upsilon_0^H)}{(1 - \varphi)}\right)$$

we define φ^* as

$$\varphi^* = max[\pi(\varphi^Q), \pi(\varphi^q)]$$

With the optimal φ^*, we define the future optimal profit function as $\pi^*(\varphi^*)$ and the optimal upper and lower bounds, t^{q*} and t^{Q*} as

$$\lceil t^{Q*} \rceil = \frac{g}{h}\log_\omega\left(\frac{g^*(\varphi^*\Upsilon_t^L + \varphi^*\Upsilon_0^H - \Upsilon_0^H)}{(1 - \varphi^*)}\right) - Ct$$

(2.34)

$$\lfloor t^{q*} \rfloor = \frac{g}{h}\log_\omega\left(\frac{g(\varphi^*\Upsilon_t^L + \varphi^*\Upsilon_0^H - \Upsilon_0^H)}{(1 - \varphi^*)}\right) - Ct$$

(2.35)

The time of the global maximum profit under the best-case scenario, tm, from period

t will be in the range between t^{q*} and t^{Q*}. Depending on the difference between g and $g*$ the possible range of the optimal upper and lower bounds can be large or small, but the entrepreneur does know that the optimal profit function at period tm will be $\pi^*(\varphi^*)$. Therefore, if the current profit function at the end of time t, π^c_t, is greater or equal then $\pi^*(\varphi^*)$, then the multi-period investment should be stopped. Then if the profit function π^c_t is greater then or equal to the MPP, the investment should be exploited, else the multi-period investment should be abandoned. If the future optimal profit function $\pi^*(\varphi^*)$ is greater, then the multi-period investment should continue because conditional on observing the best-case combinations (h, g), the entrepreneur has the possibility of gaining more value through the gathering of more data. Therefore, at time 0, the best-case future optimal profit function is $\pi^{*,0}$ and the multi-period investment should continue as long as the current profit function, π^c_t, is less then that, but when a non best-case combination is observed, then the best-case future optimal profit function becomes $\pi^{*,1}$, and the multi-period investment should continue until it is equal to it. The 0 and 1 from the future optimal profit functions correspond to the number of non best-case combinations observed, and because of the sigmoid shape of the probability and the cost, C_t, of each period of the investment, the future optimal profit function has an inverse relationship to every non best-case combination observed so that

$$\pi^{*,n^c} > \pi^{*,n^{c+1}} \tag{2.36}$$

where n^c is the number of non best-case combinations observed.

2.2 N-State NODP Model

The NODP N-State model is defined as the PPU, U^p, having N possible states.

$$U^p = \{H^1, \ldots, H^N\}$$

$$H^i < H^{i+1}$$

The NFU can take on R values, $u_t^f = \{g^1 \ldots g^R\}$, in period t where $g^r < g^{r+1}$. The NPU has S possible states, $u_t^p = \{h^1 \ldots h^S\}$, in period t where $h^s < h^{s+1}$. The number of both NPU and NFU observations are both greater then the number of PPU states.

$$R > N$$

$$S > N$$

The model is now reflective of all the possible range of observations an entrepreneur can observe, althought R and S are from the same state space, they are not necessarily equivalent. This causes the N-state model to possibly have an extremely large set of combinations that can be observed. If the multi-period investment is exploited, the values of the possible states are V^{H^1} to V^{H^N}, where V^{H^1} is less then V^{H^N}. The probability of the PPU, U_t^p, being H^i at period t is

$$P(U_t^p = H^i) = \frac{\Upsilon_t^{H^i}}{\sum_{n=1}^{N} \Upsilon_t^{H^n}} \tag{2.37}$$

where

$$\Upsilon_0^{H^1} = \Upsilon_0^{H^2} = \ldots = \Upsilon_0^{H^{N-1}} = \Upsilon_0^{H^N}$$

$$\text{where } \Upsilon_0^{H^i} \geq 1, \text{ for } i \text{ in } 1 \ldots N$$

In the N-state model there is still one minimum rate of return (MRR) and one TPU, because it is associated with the minimum number of paid users needed to justify an expansion or a limit to the capacity of the temporary infrastructure in stage two of the multi-period investment. Here we hightlight a that in the industry, VC's equate rapid and exponential sales with a large opportunity and high valuation, therefore the reality of multiple demand levels being observed should signal different levels of PPU and subsequently different V^{H^i}'s. For industry application purposes we present a simple and transparent way to structure the observations, for valuation purposes. There are now N periods of time t^{*i}'s, where i is 1 to N, that are the longest possible periods of time that the entrepreneur can continue the multi-period investment while still meeting the MRR. Each corresponding to every possible state of the PPU. Using equation 2.5,

$$\lfloor t^{*i} \rfloor = \frac{V^{H^i}}{(1 + MRR)C}$$

$$\text{where } i = 1, \ldots, N$$

There are m^+ t^{*i}'s that are positive, and m^- t^{*i}'s that are negative or 0 because the V^{H^i}'s are less then 0 or the V^{H^i}'s are too small given the cost of each period. The total sum of m^-, and m^+ equals N. Since the PPU U^p are ordered according to their

valuations, V^{HN} to V^{HN^*} have t^{*i}'s in m^+, and V^{H1} to V^{HN^-} have t^{*i}'s in m^-.

$$\overbrace{V^{H1},\ldots,V^{HN^-}}^{m^-},\overbrace{V^{HN^*},\ldots,V^{HN}}^{m^+}$$
$$\overbrace{t^1,\ldots,t^{N^-}}^{m^-},\overbrace{t^{N^*},\ldots,t^N}^{m^+}$$

We calculate the adjusted NPU lowerbound b for the t^{*i}'s in set m^+.

$$b^i = \frac{TPU}{\lfloor t^{*i} \rfloor} \tag{2.38}$$

$$\text{where } i \in \{m^+\} \tag{2.39}$$

Or if u^p is a growth ratio

$$b^i = \left(\frac{TPU}{\text{TFU}^*} \right)^{\frac{1}{t^{*i}}} - 1 \tag{2.40}$$

$$\text{where } i \in \{m^+\} \tag{2.41}$$

The objective of the entrepreneur is to continue the multi-period investment with the assumption that at least the MRR is attainable, so we identify the largest b^i that is less then the largest NPU u^p observation possible, h^S. The larger the b^i, the smaller the value of V^{Hi}, but we want to choose a b^i that is possible. If there is a b^i that is larger then h^S then that implies under the best case scenario, the targer MRR would not be attainable and therefore that should not be our target b^m.

$$b^m \in m^+ \text{ such that } b^m \geq b^i \wedge b^m \leq h^S, \forall b^i \in m^+ \tag{2.42}$$

Now we create adjusted b_a^i's for every V^{Hi}. Let V^{Hm} denote the V^{Hi} that is associated with b^m. Using V^{Hm}, V^{HN}, the highest valuation possible, and h^S we calculate the

dollar rate increase from the minimum observation rate b, to the highest NPU u^p observation possible.

$$\Delta^Q = \frac{h^S - b^m}{V^{H^N} - V^{H^m} + 1}$$ (2.43)

where Δ^Q is the dollar rate for b_a^i's greater then b^m and

$$\Delta^q = \frac{b^m - h^1}{V^{H^m} - V^{H^1}}$$ (2.44)

where Δ^q is the dollar rate for b_a^i's less then b^m. The difference in equation 2.43 and 2.44 is that h^S is the upper bound for Υ^{H^N}, where for Υ^{H^1}, h^1 is the lower bound. With Δ^Q we define b_a^i, for values greater then m as

$$b_a^i = (V^{H^i} - V^{H^m}) * \Delta^Q + b^m$$ (2.45)

$$\text{where } i > m$$ (2.46)

and for values less then m, we use Δ^q

$$b_a^i = b^m - (V^{H^m} - V^{H^i}) * \Delta^q$$ (2.47)

$$\text{where } i < m$$ (2.48)

Now for each b_a^i we create N ω factors of the algebraic sigmoid incremental function using equation 2.9, g^R, and t^m, the length of periods associated with b^m

$$\omega^{*i} = \left(\frac{N \Upsilon_0^{H^i} t^m |g^R| \varphi - \Upsilon_0^{H^i} t^m |g^R|}{1 - \varphi} \right)^{\left| \frac{g^R}{b_a^i t^m} \right|}$$ (2.49)

$$\text{for } \{i, \dots, N\}$$ (2.50)

Just like in the 2-Stage model we choose a very high φ less then one, for example 99%. We choose the highest NFU, g^R, because ω is increasing in g and we want to establish the highest ω for each b_a^i, so that high ratios of NPU to NFU will increase the Υ values faster.

We now define our N-Stage incremental algebraic sigmoid formula from equation 2.8 as

$$\Upsilon_t^i = \Upsilon_0^i + \frac{\omega^{O_t^i}}{D_t^i} \tag{2.51}$$

The variable O_t^i is now the product of all the ratios observed from period $1, \ldots, t$ in the range of b_a^i, and D_t^i is the product of all the observed NFU, g^r, from period $1, \ldots, t$ where

$$O_t^i = \prod_{n=1}^{t} \left| \frac{h_n^s}{g_n^r} \right| \tag{2.52}$$

$$D_t^i = \prod_{n=1}^{t} |g_n^r| \tag{2.53}$$

$$\text{where } b_a^i \leq h_n^s < b_a^{i+1} \tag{2.54}$$

Equations 2.52 and 2.53 replace the need for table 2.2. For each b_a^i we define h_a^i as the largest NPU observation possible within each categorey of b_a^i.

$$h_a^i \in \{h^1, \ldots, h^S\} \text{ such that } h^s \leq h_a^i < b_a^{i+1}, \forall h^s \in \{h^1, \ldots, h_a^i\} \tag{2.55}$$

Since, we now have N φ's, we calculate N best-case scenarios, one for each b_a^i category, but this makes the maximization problem in equation 2.31 difficult to solve, because of the additional variables. Since, the φ is a fraction of the Υ's, the only component that the model needs to maximize for in the range b_a^i, is the Υ^i's. We re-define the

upper and lower bounds t^q and t^Q in terms of Υ^i.

$$t^{Q^i} = \frac{g^R}{h_a^i} \log_\omega^i [g^*(\Upsilon_t^{H^i} - \Upsilon_0^{H^i})] \tag{2.56}$$

$$t^{q^i} = \frac{g^R}{h_a^i} \log_\omega^i [g(\Upsilon_t^{H^i} - \Upsilon_0^{H^i})] \tag{2.57}$$

$$\text{For } i \in \{1, \ldots, N\} \tag{2.58}$$

Where

$$\varphi^i = \frac{\Upsilon^{H^i}}{\displaystyle\sum_{n=1}^{N} \Upsilon^{H^n}} \tag{2.59}$$

The N-State maximization problem (with t^Q substituted in) is now

$$\underset{\Upsilon_t^{H^i}}{\text{maximize}} \; \frac{\displaystyle\sum_{n=1}^{N} V^{H^n} \Upsilon_t^{H^n}}{\displaystyle\sum_{n=1}^{N} \Upsilon_t^{H^n}} - C \frac{g^R}{h_a^i} \log_\omega^i [g^*(\Upsilon_t^{H^i} - \Upsilon_0^{H^i})] \tag{2.60}$$

$$\Rightarrow \frac{\displaystyle\sum_{n=1}^{N} V^{H^i} \Upsilon_t^{H^n} - \sum_{n=1}^{N} V^{H^n} \Upsilon_t^{H^n}}{\left(\displaystyle\sum_{n=1}^{N} \Upsilon_t^{H^n}\right)^2} = C \frac{g^R}{h_a^i \log(\omega^i)} \left[\frac{1}{\Upsilon_t^{H^i} - \Upsilon_0^{H^i}} \right]$$

$$\Rightarrow \frac{\displaystyle\sum_{n \neq i}^{N} \Upsilon_t^{H^n} (V^{H^i} - V^{H^n})}{\left(\displaystyle\sum_{n=1}^{N} \Upsilon_t^{H^n}\right)^2} = C \frac{g^R}{h_a^i \log(\omega^i)} \left[\frac{1}{\Upsilon_t^{H^i} - \Upsilon_0^{H^i}} \right]$$

$$\Rightarrow \frac{\Upsilon_t^{H^i} - \Upsilon_0^{H^i}}{\left(\displaystyle\sum_{n=1}^{N} \Upsilon_t^{H^n}\right)^2} = C \frac{g^R}{h_a^i \log(\omega^i)} \left[\frac{1}{\displaystyle\sum_{n \neq i}^{N} \Upsilon_t^{H^n} (V^{H^i} - V^{H^n})} \right]$$

Since the right side of the equation is a constant, for simpler notation we define the variable Z as

$$Z^i = C \frac{g^R}{h_a^i \, log(\omega^i)} \left[\frac{1}{\sum\limits_{n \neq i}^{N} \Upsilon_t^{H^n}(V^{H^i} - V^{H^n})} \right] \tag{2.61}$$

The maximization problem becomes

$$\Rightarrow \frac{\Upsilon_t^{H^i} - \Upsilon_0^{H^i}}{\left(\sum\limits_{n=1}^{N} \Upsilon_t^{H^n} \right)^2} = Z^i$$

$$\Rightarrow \Upsilon_t^{H^i} - \Upsilon_0^{H^i} = Z^i((\Upsilon_t^{H^i})^2 + \sum_{n \neq i}^{N} 2\Upsilon_t^{H^n}\Upsilon_t^{H^i}) + Z^i(\sum_{n \neq i}^{N} \Upsilon_t^{H^n}(\sum_{p \neq i}^{N} \Upsilon_t^{H^p}))$$

$$\Rightarrow -Z^i(\Upsilon_t^{H^i})^2 + \Upsilon_t^{H^i} - Z^i\Upsilon_t^{H^i}\sum_{n \neq i}^{N} 2\Upsilon_t^{H^n} - \Upsilon_0^{H^i} = Z^i(\sum_{n \neq i}^{N} \Upsilon_t^{H^n}(\sum_{p \neq i}^{N} \Upsilon_t^{H^p}))$$

$$\Rightarrow -Z^i(\Upsilon_t^{H^i})^2 + \Upsilon_t^{H^i}(1 - Z^i\Upsilon_t^{H^i}\sum_{n \neq i}^{N} 2\Upsilon_t^{H^n}) - \left[\Upsilon_0^{H^i} + Z^i(\sum_{n \neq i}^{N} \Upsilon_t^{H^n}(\sum_{p \neq i}^{N} \Upsilon_t^{H^p})) \right]$$

Using the quadratic formula, solve for $\Upsilon_t^{H^{iQ}}$ and $\Upsilon_t^{H^{iq}}$

$$\Upsilon_t^{H^{iQ}} = \frac{-(1 - Z^i\Upsilon_t^{H^i}\sum\limits_{n \neq i}^{N} 2\Upsilon_t^{H^n}) + \sqrt{(1 - Z^i\Upsilon_t^{H^i}\sum\limits_{n \neq i}^{N} 2\Upsilon_t^{H^n})^2 - 4Z^i\left[\Upsilon_0^{H^i} + Z^i(\sum\limits_{n \neq i}^{N} \Upsilon_t^{H^n}(\sum\limits_{p \neq i}^{N} \Upsilon_t^{H^p})) \right]}}{-2Z^i}$$

$$\tag{2.62}$$

$$\Upsilon_t^{H^{iq}} = \frac{-(1 - Z^i\Upsilon_t^{H^i}\sum\limits_{n \neq i}^{N} 2\Upsilon_t^{H^n}) - \sqrt{(1 - Z^i\Upsilon_t^{H^i}\sum\limits_{n \neq i}^{N} 2\Upsilon_t^{H^n})^2 - 4Z^i\left[\Upsilon_0^{H^i} + Z^i(\sum\limits_{n \neq i}^{N} \Upsilon_t^{H^n}(\sum\limits_{p \neq i}^{N} \Upsilon_t^{H^p})) \right]}}{-2Z^i}$$

$$\tag{2.63}$$

Substitute $\Upsilon_t^{H^{iQ}}$ and $\Upsilon_t^{H^{iq}}$ into the profit function and then take the maximum

$$\frac{\sum_{n=1}^{N} V^{H^n} \Upsilon_t^{H^n}}{\sum_{n=1}^{N} \Upsilon_t^{H^n}} - C\frac{g^R}{h_a^i} \log_\omega^i [g^*(\Upsilon_t^{H^i} - \Upsilon_0^{H^i})]$$

The future maximum value profit function for the best-case scenrio for each range b_a^i at period t is $\pi_t^{i*} = \max\{\pi^i(\Upsilon_t^{H^{iQ}})_t, \pi^i(\Upsilon_t^{H^{iq}})_t\}$. At period t, after observing h^s and g^r, the current profit function is π_t. If there are no future maximum profit functions π_t^{i*} that are greater then π_t, then the multi-period investment should be stopped, and if π_t is greater or equal to MRR, the multi-period investment should be exploited otherwise it should be abandoned. If all π_t^{i*} are greater then π_t, then the investment should continue. For the majority of the scenarios, π_t will have several π_t^{i*}'s greater then it. Here we provide a heuristic based on b^m. Let $\pi_t^{\bar{m}}$ equal the average of future profit functions for ranges greater or equal to b^m, at period t

$$\pi_t^{\bar{m}} = \frac{\sum_{i=m}^{N} \Upsilon^i \times \pi_t^{i*}}{\sum_{i=m}^{N} \Upsilon^i} \tag{2.64}$$

If π_t is less then $\pi_t^{\bar{m}}$, then continue the multi-period investment. Once π_t is greater then or equal to $\pi_t^{\bar{m}}$, then stop the multi-period investment, and then exploit or abandon based on the value to the MPP.

2.3   Example

Consider the case of a two-state NODP problem, where an entrepreneur creates a community around the subject of a widget (NFU). After observing growth of activity and membership in the community the entrepreneur releases a prototype of a widget, tailored to the community that users can pay to use. The widget costs $100 a month

to run and manage, but a commercial version, that is stable and can be scaled, needs $10,000 additional invested. The monthly payments are sunk costs once they are made and once the entrepreneur invests the $10,000 that cannot be recuperated. If the monthly paid user base (PPU) can get to 5000, then the entrepreneur will make $20,000, on the commercial version of the new widget, but if the number of paid users does not make it to 5000, the value of the investment will be a loss of $9000. Time is not a limiting factor, after the completion of the commercial grade widget, the demand will be realized instantaneously. The unobservable core process has two states, 5000 and 0 (any value less then 0 is considered 0 and any value greater then 5000 can be considered 5000).

$$U^p = \{5000, 0\}$$

$$V^H = 20,000$$

$$V^L = -9,000$$

The entrepreneur does not know what the transtional probabilities are, but knows the monthly addition of new members in the widget community, u^f, and how many people from the widget community transition to paying for the widget prototype, u^p. The number of new widget community members can only have two values, 120 and 20, where 120 is a positive sign of problem validation and 20 is a negative sign. The number of new paid users joining the prototype, can also only have two values, 80 or 10, where 80 is considered a positive sign of user adoption and 10 is considered a

negative sign.

$$u^f = \{120, 20\}$$

$$u^p = \{80, 10\}$$

$$P(u_t^p = 80, u_t^f = 120 | U_t^p = 7500) = 70\%$$

$$P(u_t^p = 80, u_t^f = 20 | U_t^p = 7500) = 5\%$$

$$P(u_t^p = 10, u_t^f = 120 | U_t^p = 7500) = 5\%$$

$$P(u_t^p = 10, u_t^f = 20 | U_t^p = 7500) = 20\%$$

$$P(u_t^p = 10, u_t^f = 120 | U_t^p = 0) = 15\%$$

$$P(u_t^p = 10, u_t^f = 20 | U_t^p = 0) = 40\%$$

$$P(u_t^p = 80, u_t^f = 120 | U_t^p = 0) = 15\%$$

$$P(u_t^p = 80, u_t^f = 20 | U_t^p = 0) = 30\%$$

The entrepreneur requires at least a 20% return (MRR) on the multi-period invest-
ment.

We generate observations for three scenarios. The first scenario generates observations
for just the $U^p = 5000$ unobservable state, the second scenario generates observations
for just the $U^p = 0$ unobservable state, and the third scenario generates observations
for the unobservable process following the probability transitions. We generate 30

observations for each scenario.

Scenario 1: $1, 1, 1, 4, 4, 1, 1, 1, 1, 1, 4, 1, 1, 3, 1,$

$4, 1, 2, 1, 1, 4, 1, 4, 1, 1, 1, 2, 1, 4, 2$

Scenario 2: $2, 4, 4, 1, 3, 3, 4, 2, 4, 4, 4, 3, 3, 4, 1,$

$4, 4, 2, 2, 1, 4, 4, 2, 2, 4, 1, 4, 4, 2, 3$

Scenario 3: $4, 1, 1, 3, 1, 1, 1, 3, 4, 4, 4, 1, 1, 3, 1,$

$1, 3, 1, 1, 4, 4, 1, 3, 3, 1, 1, 3, 1, 1, 1$

core: $0, 5000, 5000, 0, 5000, 0, 0, 0, 0, 0,$

$0, 5000, 5000, 5000, 5000, 5000, 5000, 5000, 5000, 0,$

$5000, 5000, 5000, 5000, 5000, 5000, 5000, 5000, 5000, 5000$

Index: $1 - (80, 120)$

Index: $2 - (80, 20)$

Index: $3 - (10, 120)$

Index: $4 - (10, 20)$

Even though the entrepreneur has no notion of the transitional probabilities, the numbers in scenario 3 are being generated by them.

$$P(U_t^p = 5000 | U_{t-1}^p = 5000) = 90\%$$

$$P(U_t^p = 0 | U_{t-1}^p = 5000) = 10\%$$

$$P(U_t^p = 0 | U_{t-1}^p = 0) = 40\%$$

$$P(U_t^p = 5000 | U_{t-1}^p = 0) = 60\%$$

We calculate the values needed for our best-case heuristic.

$$\Upsilon^H = 100$$

$$\Upsilon^L = 100$$

$$t^* = 83$$

$$b = 60.24$$

$$\omega = 1.34$$

Combo 1 will contribute to Υ^H, where as combo 3 and 4 will contribute to Υ^L. Combo 2 will not contribute to anything and be treated as noise, but in scenario 2, we'll contribute it to Υ^H. In all 3 scenarios, the multi-period investment was stopped after 18 months, for a loss of $1800. Where for scneario 2, that was a great result for scenario 1 and 3, the entrepreneur passed on an investment that otherwise would have been profitable. The reason is because the NODP model needs several periods of consistent data that signals growth, to increase the probability beyond the range of 50%. Given, that this example has the not practical expected return of a negative profit at 50%, the NODP model ends the investment early. We state that the example is not practical because if investors and entrepreneurs would not invest in a scenario, where in period 0, there is an equal amount of loss and profit possible. If we run all 3 scenarios again excluding the $10,000 sunk cost required to build the commercial grade widget, in sceario 1, the entrepreneur will make $13,000, in sceario 2, the entrepreneur will lose $12,000, and in sceario 3, the entrepreneur will make $12,900. The loss in scenario 2, is just from the multi-period cost, which shows that the NODP is more expensive when the unobservable state and its subsequent observations are not profitable. In that scenario, the entrepreneur lost more then if the commercial grade investment was made initially and failed. The upside in this case drove the investment past the $9,000 mark.
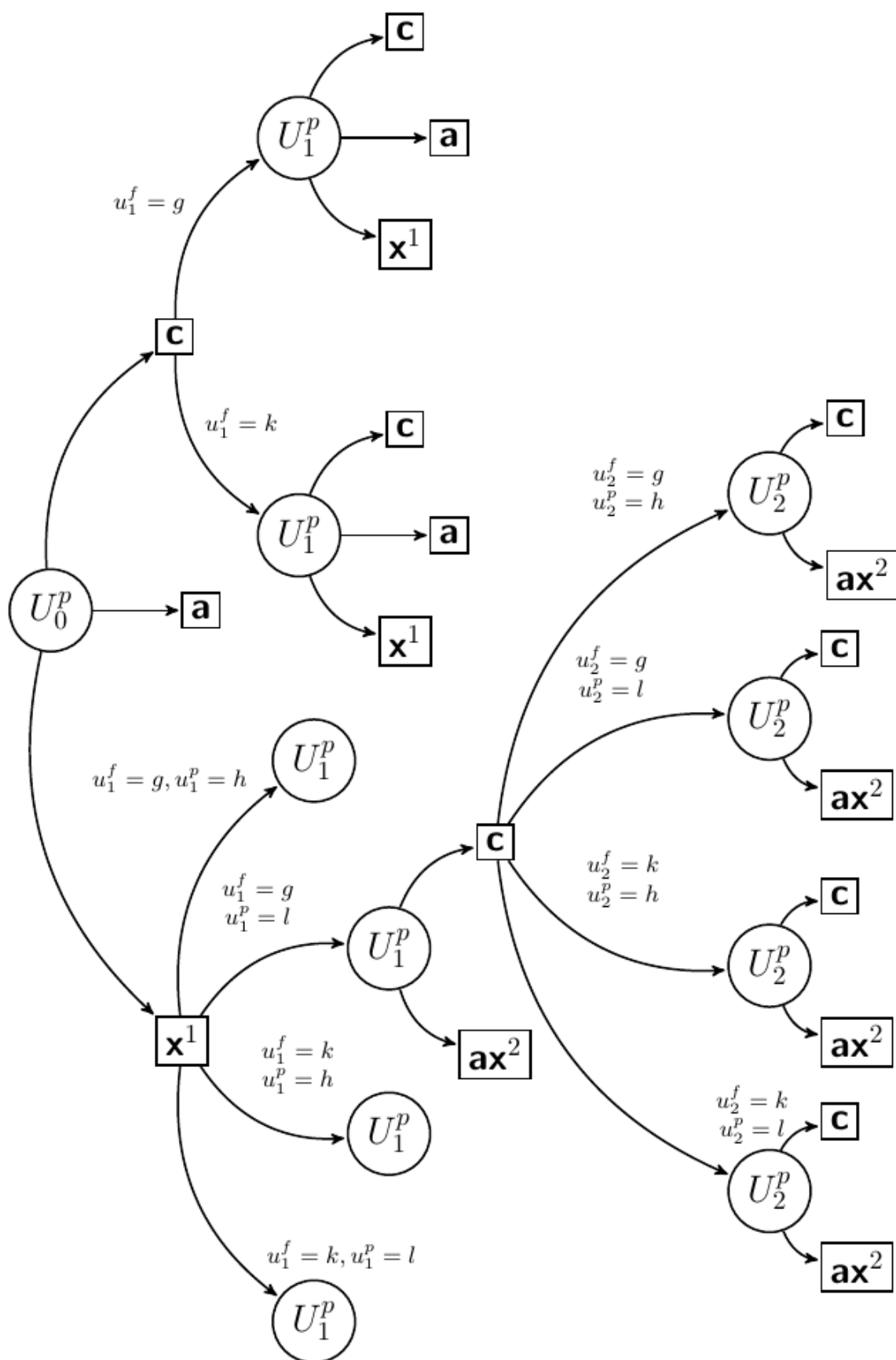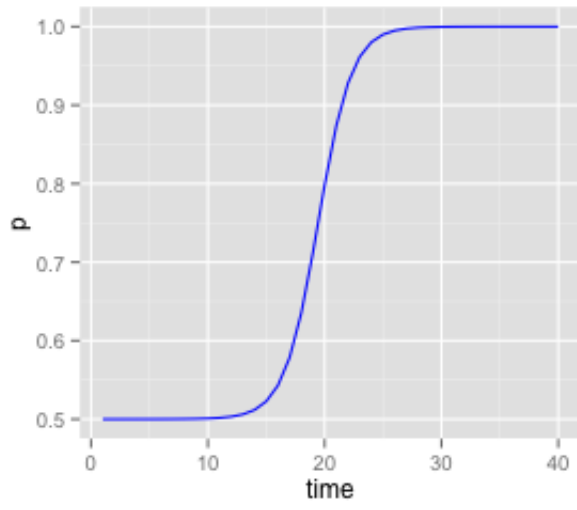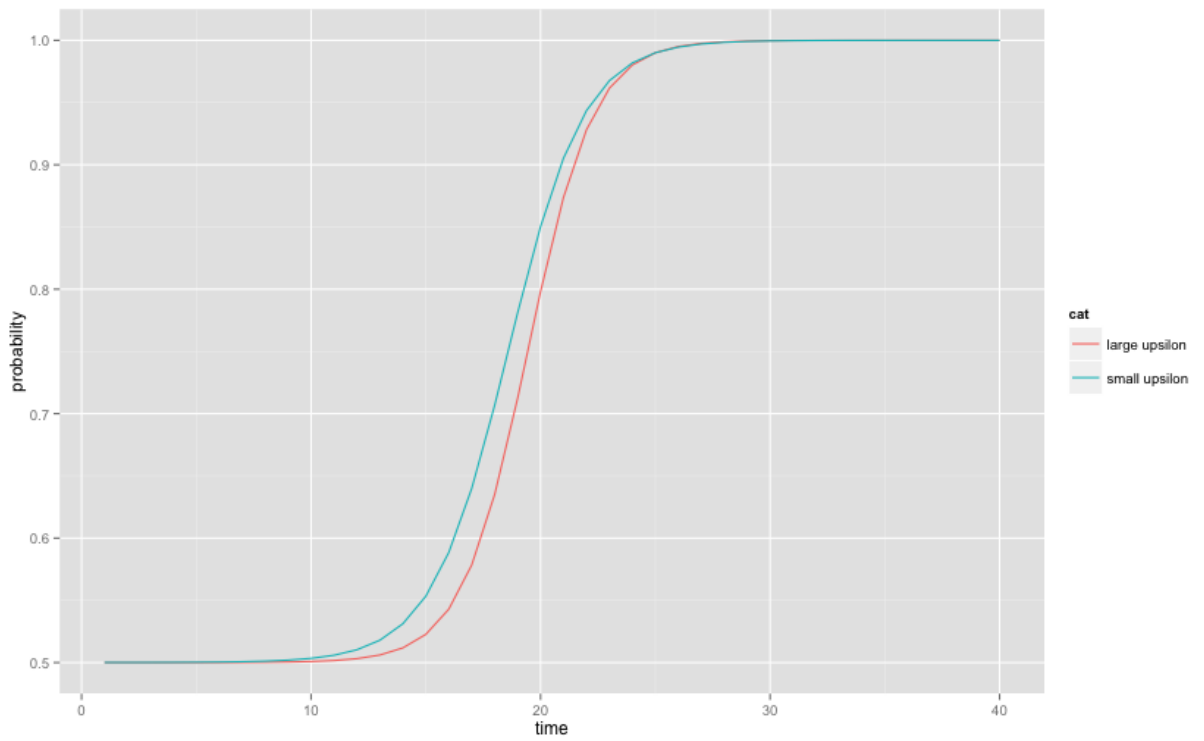
Figure 2.1: Two-State NODP Transitional Diagram

Figure 2.2: Example of Logistic curve

Figure 2.3: Examples of Logistic curves with different Υ's

CHAPTER 3

Policy Function Approximation

3.1 NODP Policy Function

In this section we model the N-Stage Network Option Decision Process as an approximate dynamic program. We now assume that the PPU U^p now as N possible states corresponding to N possible valuations, so that N is greater then $S \times R$, where S is the number of possible states of the NPU and R is the number of possible states of the NFU.

$$U^p = H^1, \ldots, H^N \text{ where } H^i < H^{i+1}$$

$$u^f = g^1, \ldots, g^R \text{ where } g^r < g^{r+1}$$

$$u^p = h^1, \ldots, h^S \text{ where } h^s < h^{s+1}$$

$$\text{so that } N > S \times R$$

The size of the NPU is now indicative of conditional nature of the valuations on the speed and value of the aggregate observations (h^s, g^r). The N-State heuristic could still be used to find the optimal stopping point as long the variance for the expected future profit function, $\bar{\pi}_t^m$, isn't too large. For small PPU's, NPU's and NFU's, we could use regular dynamic programming, using pre-calculated lookup tables, but here our assumption is that N,S, and R are very large, reflecting the continuous nature of ratios entrepreneurs observe. On the other hand, VC's tend to aggregate valuations

according to even lot ranges, \$1,000,000, \$5,000,000, \$10,000,000, etc., in spite of the continuous nature of the data they observe. We define the state variable as Φ_t at period t, as a vector of all the possible Υ's.

$$\Phi_t = \{\Upsilon_t^1, \ldots, \Upsilon_t^N\} \tag{3.1}$$

$$\text{where } \Upsilon_0^1 = \Upsilon_0^2 = \ldots = \Upsilon_0^{N-1} = \Upsilon_0^N \text{ and } \Upsilon^n \geq 1 \tag{3.2}$$

The actual state should be the PPU U^p, but because it is not observable during the multi-period investment, the vector of Υ's is used as function to represent it. We use the variable a_t to represent the three discrete action spaces in our model.

$$A_t = \{c, x, a\} \tag{3.3}$$

We let W_t represent the combinations of NFU and NPU observed at and up to period t.

$$W_t = \{(h^s, g^r)_1, \ldots, (h^{s'}, g^{r'})_t\} \tag{3.4}$$

The variable, Φ^M, represents the transition function from state Φ_t to Φ_{t+1}.

$$\Phi_{t+1} = \Phi^M(\Phi_t, A_t, W_{t+1}) \tag{3.5}$$

The objective function is the expected net profit, and is defined as

$$B(\Phi_t, A_t) = \frac{\sum_{n=1}^{N} V^{H^n} \Upsilon_t^n}{\sum_{n=1}^{N} \Upsilon_t^n} - (C \times t) \tag{3.6}$$

where C is the cost of continuing the multi-period investment each period. We want to maximize the objective function, as a expected net profit. Our objective is to find the best policy Π that will give net returns greater then or equal to zero.

$$\max_{\Pi} \mathbb{E}^{\Pi} \sum_{t=1}^{T} B(\Phi_t, A^{\Pi}(\Phi_t)) \tag{3.7}$$

where function A^{Π} determines the action choice according to policy Π. Typically, we can compute an optimal policy using Bellman's optimality equation which would be

$$\Gamma_t(\Phi_t) = \max_{A_t} \left(B(\Phi_t, A_t) + \sum_{s^*} p(s^*|\Phi_t, A_t) \Gamma_{t+1}(s^*) \right) \tag{3.8}$$

where $\Gamma_t(\phi)$ equals the value of being in state $\Phi_t = \phi$ at period t to the end of the mult-period investment following the optimal policy Π. The function $p(s^*|\Phi_t, A_t)$ is the probability of transitioning to state s^*, given that the current state is Φ_t and action A_t is chosen. The dilemma our model faces is that we have no way to calculate the probability transition from one state to another. We are wholly reliant on the combinations we observe that are generate according to a unobservable random distribution. We developed a heuristic for the N-Stage model, where we created ranges to aggregate NPU observations associated with the possible valuation, V^{H^n}, levels and then developed probabilities using our algebraic sigmoid function that are dependent on the combinations observed to calculate expected value. We did not create transition probabilities, instead we created future best-case expected values for each aggregated range and then took an average of them. The strategy being, if there is an expected return greater then or equal to the MRR possible, then we should continue the multi-period investment because the primary objective was gaining value through information. For our approximate dynamic program we now need to develop a probability approximation for the transition function and policy function, which will map

states to actions. Typically, we would create a lookup table, but given the size of the observations now, an approximation is needed. We can write the optimality equation in terms of our transition function Φ^M as

$$\Gamma_t(\Phi_t) = \max_{A_t} \left(B(\Phi_t, A_t) + \mathbb{E}\{\Gamma_{t+1}(\Phi^M(\Phi_t, A_t, W_{t+1}))|\Phi_t\} \right) \tag{3.9}$$

The first step is to aggregate the N possible states of the PPU into M states. In forecasting a multi-period investment in the industry, there is little difference between \$1,000,000, \$1,039,052.35, or \$1,109,802. In addition, there are fundamental limits to how much money can be lost (initial costs plus option based costs) and how much can be profited (limited size of the network).

$$\overbrace{H^1, \ldots, H^n}^{H^{*1}}, \ldots, \overbrace{H^{N-o}, \ldots, H^N}^{H^{*M}} \tag{3.10}$$

$$\overbrace{V^{H^1}, \ldots, V^{H^n}}^{V^{H^{*1}}}, \ldots, \overbrace{V^{H^{N-o}}, \ldots, V^{H^N}}^{V^{H^{*M}}} \tag{3.11}$$

The state variable Φ^* reflects this in its vector of Υ's size M. The next step is to calculate the adjusted lower bound for the NPU, b_a^i, and the maximum periods the entrepreneur can continue without violating the MRR, just as we did with the best-case heuristic.

$$\lfloor t^m \rfloor = \frac{V^{H^{*m}}}{(1 + MRR)C}$$

$$\text{where } m = 1, \ldots, M$$

$$b^m = \frac{TPU}{lfloor t^m rfloor}$$

$$\text{where } t^m > 0$$

Or if u^p is a growth ratio

$$b^m = \left(\frac{TPU}{TFU^*}^{\frac{1}{t^m}} \right) - 1$$

where $t^m > 0$

We then identify, b^e, such that it is the largest b^m less then or equal to the largest possible value for the NPU, h^S, and then calculate the positive and negative dollar rates, Δ^Q and Δ^q.

$$\Delta^Q = \frac{h^S - b^e}{V^{HM} - V^{He} + 1} \tag{3.12}$$

$$\Delta^q = \frac{b^e - h^1}{V^{He} - V^{H1}} \tag{3.13}$$

$$b_a^m = (V^{H^m} - V^{H^e}) \times \Delta^Q + b^e \tag{3.14}$$

where $m \geq e$ \hfill (3.15)

$$b_a^m = b^e - (V^{H^e} - V^{H^e}) \times \Delta^q \tag{3.16}$$

where $m < e$ \hfill (3.17)

We now define the ω^m and the algebraic sigmoid function from equation 2.9 and equation 2.51.

$$\omega^{*m} = \left(\frac{N \Upsilon_0^{H^m} t^e |g^R| \varphi - \Upsilon_0^{H^m} t^e |g^R|}{1 - \varphi} \right)^{\left| \frac{g^R}{b_a^m t^e} \right|} \tag{3.18}$$

for $\{m, \ldots, M\}$ \hfill (3.19)

$$\Upsilon_t^m = \Upsilon_0^m + \frac{\omega^{O_t^m}}{D_t^m} \tag{3.20}$$

Where O_t^m and D_t^m are defined as

$$O_t^m = \prod_{i=1}^{t} \left| \frac{h_i^s}{g_i^r} \right|$$

$$D_t^m = \prod_{i=1}^{t} |g_i^r|$$

where $b_a^m \leq h_i^s < b_a^{m+1}$

We now define the adjusted upper bound of the NPU.

$$h_a^m \in \{h^1, \ldots, h^S\} \text{ such that } h^s \leq h_a^m < b_a^{m+1}, \forall h^s \in \{h^1, \ldots, h_a^m\}$$

We can now define our policy approximation Π, which is a rule based function. Let P_t be the set of all Υ's greater then 0 at period t.

$$\{P_t\} \in \{\Upsilon_t^1, \ldots, \Upsilon_t^M\} \text{ such that } \Upsilon_t^m > 0 \tag{3.21}$$

For all the Υ's in set P_t, we now calculate $|P|$ scenarios of only the the NPU ranges in set P. Let Λ_{t+1}^p equal the expected profit if in the next period, the corresponding range's h_a^p was observed as the NPU and the average $\bar{D}_t^m$ was observed.

$$\Upsilon_{t+1}^p = \Upsilon_0^p + \frac{\omega^{O_t^m \times \frac{h_a^p}{D_t^m}}}{D_t^m \times \bar{D}_t^m} \tag{3.22}$$

$$\Lambda_{t+1}^p = \frac{\sum_{m=1}^{M} V^{H^m} \Upsilon_{t+1}^M}{\sum_{m=1}^{M} \Upsilon_{t+1}^m} - (C \times (t+1)) \tag{3.23}$$

where $p \in \{1, \ldots, P\}$ \hfill (3.24)

The expected value of the continuing therefore is

$$\mathbb{E}[\Lambda] = \frac{1}{|P|} \sum_{p=1}^{P} \Lambda_{t+1}^{p} \tag{3.25}$$

So the policy is therefore, that after t^e periods (least amount of time to meet MRR limit), if the current $B(\Phi_t, A_t)$ is less then $\mathbb{E}[\Lambda]$, then continue on with the multi-period investment. If it is greater then or equal to, then stop the multi-period investment and exploit if the expected return is greater then or equal to the MRR and abandon otherwise. It is important to note that in our policy function we delay making a decision till t^e. The reason is that enough observations have not been sampled yet, and therefore the model waits the minimum time that the probability can convincingly signal a PPU state.

References

Agha Ibrahim Akram. *The Sword of Allah: Khalid bin al-Waleed âĂŞ His Life and Campaigns.* Oxford University Press, NY, New York, 2004.

BR Barmish and Constantino M Lagoa. The uniform distribution: A rigorous justification for its use in robustness analysis. *Mathematics of Control, Signals and Systems*, 10(3):203–222, 1997.

Carsten Bienz and Julia Hirsch. The dynamics of venture capital contracts. *Review of Finance*, page rfr003, 2011.

Apostolos N Burnetas, Odysseas Kanavetas, and Michael N Katehakis. Asymptotically optimal multi-armed bandit policies under a cost constraint. *arXiv preprint arXiv:1509.02857*, 2015.

CBInsights. 146 startup failure post-mortems, December 2015. URL https://www.cbinsights.com/blog/startup-failure-post-mortem/.

Young Rok Choi, Moren Lévesque, and Dean A Shepherd. When should entrepreneurs expedite or delay opportunity exploitation? *Journal of Business Venturing*, 23(3):333–355, 2008.

Wesley Cowan and Michael N Katehakis. Asymptotic behavior of minimal-exploration allocation policies: Almost sure, arbitrarily slow growing regret. *arXiv preprint arXiv:1505.02865*, 2015a.

Wesley Cowan and Michael N Katehakis. Asymptotically optimal sequential experimentation under generalized ranking. *arXiv preprint arXiv:1510.02041*, 2015b.

Wesley Cowan and Michael N Katehakis. Normal bandits of unknown means and variances: Asymptotic optimality, finite horizon regret bounds, and a solution to an open problem. *arXiv preprint arXiv:1504.05823*, 2015c.

George Deeb. Growth vs. profit: What should rising startups focus on first?, March 2014. URL http://thenextweb.com/entrepreneur/2014/03/16/growth-vs-profit-startups-focus-first/#gref.

Dwi Ertiningsih, Michael N Katehakis, Laurens C Smit, and Floske M Spieksma. Level product form qsf processes and an analysis of queues with coxian inter-arrival distribution. *arXiv preprint arXiv:1507.05298*, 2015.

Adam Fleischhacker, Michael N Katehakis, et al. Dynamic pricing in a dual market environment. *arXiv preprint arXiv:1509.07338*, 2015.

John P Gilbert and Frederick Mosteller. Recognizing the maximum of a sequence. In *Selected Papers of Frederick Mosteller*, pages 355–398. Springer, 2006.

Felix Giliette. The rise and inglorious fall of myspace. 2011.

Hemantha SB Herath and Chan S Park. Real options valuation and its relationship to bayesian decision-making methods. *The Engineering Economist*, 46(1):1–32, 2001.

Hemantha SB Herath and Chan S Park. Multi-stage capital investment opportunities as compound real options. *The Engineering Economist*, 47(1):1–27, 2002.

Michael N Katehakis and Kartikeya S Puranam. On optimal bidding in sequential procurement auctions. *Operations Research Letters*, 40(4):244–249, 2012.

Michael N Katehakis and Laurens C Smit. A successive lumping procedure for a class of markov chains. *Probability in the Engineering and Informational Sciences*, 26 (04):483–508, 2012.

MICHAEL N KATEHAKIS, LAURENS C SMIT, and FLOSKE M SPIEKSMA. A comparative analysis of the successive lumping and the lattice path counting algorithms.

Michael N Katehakis, Laurens C Smit, and Floske M Spieksma. On the solution to a countable system of equations arising in stochastic processes. *arXiv preprint arXiv:1510.05859*, 2015.

Spyros Makridakis and Nassim Taleb. Decision making and planning under low levels of predictability. *International Journal of Forecasting*, 25(4):716–733, 2009.

S.K. Malik. *Khalid bin Walid: the general of Islam: a study in Khalid's generalship*. Ferozsons publishers, Lahore, 1968.

Eric Ries. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Random House LLC, 2011.

Darrell Rigby. Management tools and techniques: A survey. *California Management Review*, 43(2):139, 2001.

Eduardo S Schwartz and Lenos Trigeorgis. *Real options and investment under uncertainty: classical readings and recent contributions*. MIT press, 2004.

J Shi, Michael N Katehakis, and Benjamin Melamed. Cash-flow based dynamic inventory management. Technical report, Working paper, Rutgers Business School. Newark, NJ 07102, 2012.

Nassim Nicholas Taleb, Yaneer Bar-Yam, Raphael Douady, Joseph Norman, and Rupert Read. The precautionary principle: fragility and black swans from policy actions. *Extreme Risk InitiativeâĂŤNYU School of Engineering Working Paper Series*, 2014.

Andrew Weissman. Investment thesis @USV, May 2012. URL https://www.usv.com/blog/investment-thesis-usv.

Xiujuan Xu, Zhu Lizhong, Zhao Xiaowei, Xu Zhenzhen, and Liu Yu. A feature-

based regression algorithm for cold-start recommendation. *Journal of Industrial and Production Engineering*, 31(1):17–26, 2014.

APPENDIX A

Examples and Simulations - R Code

A.1 Example 1

```
#POMDP − Scenario 1
scen1=sample(c(80,10),30,T,c(.75,.25))
#POMDP − Scenario 2
scen2=sample(c(80,10),30,T,c(.45,.55))
#POMDP − Scenario 3
scen3=rep(0,30)
core=rep(0,30)
core[1] = sample(c(7500,0),1,T,c(.5,.5))


for(i in 1:NROW(core))
{
        if(core[i]==7500)
        {
                scen3[i]=sample(c(80,10),1,T,c(.75,.25))
                if(i<NROW(core))
                {
                        core[i+1]=sample(c(7500,0),1,T,c(.9,.1))
                }
        }
        else
        {
                scen3[i]=sample(c(80,10),1,T,c(.45,.55))
                if(i<NROW(core))
                {
                        core[i+1]=sample(c(7500,0),1,T,c(.6,.4))
                }
        }
}

#optimal probabilities where no more value can gained
mph=(100−20000*.6)/(−9000*.1−20000*.6)
mpl=(−9000*.1−100)/(−9000*.1−20000*.6)
```

```
calcScenario=function(scen,actual=7500,p_7500=.5,p_0 = .5,
prob_80_7500=.75,prob_10_7500=.25,prob_80_0=.45,prob_10_0=.55,
probHH=.9,probLH=.1,probHL=.6,probLL=.4)
{
        #The core state probabilities at time 0
        prob_7500=p_7500
        prob_0 = .5

        p7500res=rep(0,NROW(scen))
        p0res=rep(0,NROW(scen))

        for(i in 1:NROW(scen))
        {
                prob_7500t = probHH*prob_7500 + probHL*prob_0
                prob_0t = probLH*prob_7500 + probLL*prob_0
                if(scen[i]==10)
                {
                        prob_7500=(prob_10_7500*prob_7500t)/
                        (prob_10_7500*prob_7500t+prob_10_0*prob_0t)
                        prob_0 = (prob_10_0*prob_0t)/
                        (prob_10_7500*prob_7500t+prob_10_0*prob_0t)
                }
                else
                {
                        prob_7500=(prob_80_7500*prob_7500t)/
                        (prob_80_7500*prob_7500t+prob_80_0*prob_0t)
                        prob_0 = (prob_80_0*prob_0t)/
                        (prob_80_7500*prob_7500t+prob_80_0*prob_0t)
                }
                p7500res[i]=prob_7500
                p0res[i]=prob_0
        }
        l=list(p7500res,p0res)
        return(l)
}

res=calcScenario(scen1)
res7500=res[[1]]
res0=res[[2]]

res=calcScenario(scen2)
res7500=res[[1]]
res0=res[[2]]
```

```
res=calcScenario(scen3)
res7500=res[[1]]
res0=res[[2]]

#POMDP Simulation
scensim = matrix(0,10000,100)
coresim = matrix(0,10000,100)

makeScen=function()
{
        scenf=rep(0,100)
        core=rep(0,100)
        core[1] = sample(c(7500,0),1,T,c(.5,.5))

        for(i in 1:NROW(core))
        {
                if(core[i]==7500)
                {
                        scenf[i]=sample(c(80,10),1,T,c(.75,.25))
                        if(i<NROW(core))
                        {
                                core[i+1]=sample(c(7500,0),1,T,c(.9,.1))
                        }
                }
                else
                {
                        scenf[i]=sample(c(80,10),1,T,c(.45,.55))
                        if(i<NROW(core))
                        {
                                core[i+1]=sample(c(7500,0),1,T,c(.6,.4))
                        }
                }
        }
        l=list(core,scenf)
        return(l)
}

for(i in 1:10000)
{
        ll=makeScen()
        coresim[i,]=ll[[1]]
        scensim[i,]=ll[[2]]
}

ret = rep(0,10000)
```

```
vres = matrix(0,20,10000)
for(v in 1:20)
{
val = (v-1)*500
mph=(100-20000*.6)/(-val*.1-20000*.6)
mpl=(-val*.1-100)/(-val*.1-20000*.6)
t_end = (20000*mph + -val*(1-mph) -10000)/100
m_end = floor(t_end/exp(1))
t_end = floor(t_end)

for(i in 1:10000)
{
ismax = 0
ll=calcScenario(scensim[i,], actual=7500,p_7500=.5,
p_0=.5,prob_80_7500=.75,prob_10_7500=.25,prob_80_0=.45,
prob_10_0=.55,probHH=1,probLH=0,probHL=.6,probLL=.4)
mphend=max(ll[[1]][1:m_end])
for(h in 1:t_end)
{
if(h<=m_end)
{
if(ll[[1]][h]>=mph)
{
        if(coresim[i,h]==7500)
        {
                ret[i]=20000 - 10000 - h*100
        }
        else
        {
                ret[i]=-val-10000-h*100
        }
        ismax = 1
        break
}
}
else
{
if(ll[[1]][h]>=mphend)
{
        if(coresim[i,h]==7500)
        {
                ret[i]=20000 - 10000 - h*100
        }
        else
        {
```

```
                        ret[i]=−val−10000−h∗100
            }
            ismax = 1
            break
}
}
}
if(ismax==0)
{
            ret[i]= −h∗100
}
}
vres[v,]=ret
}

finres = matrix(0,20,5)
for(i in 1:20)
{
            finres[i,1]=(i−1)∗500
            finres[i,2]=mean(vres[i,])
            finres[i,3]=sd(vres[i,])
            finres[i,4]=max(vres[i,])
            finres[i,5]=min(vres[i,])
}

finres2 = matrix(0,20,1)
for(i in 1:20)
{
            finres2[i,1]=paste(round(finres[i,],2), collapse="␣\&␣")
}
```

A.2   Example 2

```
#NODP − Scenario 1
#1 (80,120), 2 (80,20), 3 (10,120), 4 (10,20)
scen1=sample(c(1,2,3,4),200,T,c(.7,.05,.05,.2))
#NODP − Scenario 2
scen2=sample(c(1,2,3,4),200,T,c(.15,.3,.15,.4))
#NODP − Scenario 3
scen3=rep(0,100)
core=rep(0,100)
core[1] = sample(c(7500,0),1,T,c(.5,.5))

for(i in 1:NROW(core))
```

```
{
if ( core [ i]==7500)
{
scen3 [ i]=sample ( c ( 1 ,2 ,3 ,4 ) ,1 ,T, c ( .7 ,.05 ,.05 ,.2 ))
if ( i<NROW( core ))
{
core [ i+1]=sample ( c (7500 ,0 ) ,1 ,T, c ( .9 ,.1 ))
}
}
else
{
scen3 [ i]=sample ( c ( 1 ,2 ,3 ,4 ) ,1 ,T, c ( .15 ,.3 ,.15 ,.4 ))
if ( i<NROW( core ))
{
core [ i+1]=sample ( c (7500 ,0 ) ,1 ,T, c ( .6 ,.4 ))
}
}
}


bigQ=function ( upsl , upsh ,C, g , h ,w, vha , vhl )
{
a=−1∗( upsl+upsh )
b=( upsl+2∗upsh )
c=−1∗(C∗( g∗upsl /( h∗log (w)∗( vha − vhl )))+upsh )

if (( b^2 − 4∗a∗c)>0)
{
bq =  (−1∗b+sqrt ( b^2 − 4∗a∗c ))/(2∗a )
}
else {bq=0}
return (bq)
}


lilQ=function ( upsl , upsh ,C, g , h ,w, vha , vhl )
{
a=−1∗( upsl+upsh )
b=( upsl+2∗upsh )
c=−1∗(C∗( g∗upsl /( h∗log (w)∗( vha − vhl )))+upsh )

if (( b^2 − 4∗a∗c)>0)
{
bq =  (−1∗b−sqrt ( b^2 − 4∗a∗c ))/(2∗a )
}
else {bq=0}
return (bq)
```

```
}

w = ((2*100*83*120*.99 − 100*83*120)/(.01))^(120/(60.24*83))
calcSigmoid = function(scen,vh,vl,C)
{

uph = 100
upl = 100
cnt=rep(0,4)

cur_profit=rep(0,NROW(scen))
fin_profit=0
mrr=0
mp = rep(0,NROW(scen))

for(i in 1:NROW(scen))
{
mq=lilQ(upl,100,100,120,80,w,10000,−19000)
mQ=bigQ(upl,100,100,120,80,w,10000,−19000)
inlogq=(120*(mq*100+mq*100−100)/(1−mq))
inlogQ=(120*(mQ*100+mQ*100−100)/(1−mQ))

if(inlogq>0)
{
pq = vh*mq + vl*(1−mq) −C*(120/80)*log(inlogq,w) −10000
−C*(i−1)
qrate=(vh*mq + vl*(1−mq)−C*(120/80)*log(inlogq,w) −10000
−C*(i−1))/(C*(120/80)*log(inlogq,w) +10000 +C*(i−1))
}
else
{
pq=−10000
}

if(inlogQ>0)
{
pQ = vh*mQ + vl*(1−mQ) −C*(120/80)*log(inlogQ,w) −10000
− C*(i−1)
Qrate=(vh*mQ + vl*(1−mQ) −C*(120/80)*log(inlogQ,w) −10000
− C*(i−1))/(C*(120/80)*log(inlogQ,w) +10000 + C*(i−1))
}
else
{
pQ=−10000
}
```

```
max_profit = max(pq,pQ)
m_rate   = max(Qrate,qrate)
mp[i]=max_profit
if(scen[i]==1)
{
cnt[1]=cnt[1]+1
uph = 100 + (w^(cnt[1]*(80/120)))/(cnt[1]*120)
}
else  if(scen[i]==2)
{
cnt[2]=cnt[2]+1
}
else  if(scen[i]==3)
{
cnt[3]=cnt[3]+1
upl = 100 + (w^(cnt[3]*(10/120)))/(cnt[3]*120)
}
else
{
cnt[4]=cnt[4]+1
upl = 100 + (w^(cnt[4]*(80/120)))/(cnt[4]*120)
}
cur_profit[i]=vh*(uph/(uph+upl)) + vl*(upl/(uph+upl))
 - C*i - 10000
mrr=cur_profit[i]/(C*i+10000)
if(cur_profit[i]>=max_profit)
{
break
}
else  if(m_rate<.2)
{
break
}

}
if(mrr>=.2)
{
fin_profit=20000-10000-C*i
}
else
{
fin_profit=-C*i
}
l = list(fin_profit,mrr,cur_profit,mp,cnt)
```

```
return(l)
}

calcSigmoid2 = function(scen,vh,vl,C)
{

uph = 100
upl = 100
cnt=rep(0,4)

cur_profit=rep(0,NROW(scen))
fin_profit=0
mrr=0
mp = rep(0,NROW(scen))

w = ((2*100*166*120*.99 - 100*166*120)/(.01))^(120/(30.12*166))

for(i in 1:NROW(scen))
{
mq=lilQ(upl,100,100,120,80,w,20000,-9000)
mQ=bigQ(upl,100,100,120,80,w,20000,-9000)
inlogq=(120*(mq*100+mq*100-100)/(1-mq))
inlogQ=(120*(mQ*100+mQ*100-100)/(1-mQ))

if(inlogq>0)
{
pq = vh*mq + vl*(1-mq) -C*(120/80)*log(inlogq,w) -C*(i-1)
qrate=(vh*mq + vl*(1-mq)-C*(120/80)*log(inlogq,w)
-C*(i-1))/(C*(120/80)*log(inlogq,w) +C*(i-1))
}
else
{
pq=-10000
}

if(inlogQ>0)
{
pQ = vh*mQ + vl*(1-mQ) -C*(120/80)*log(inlogQ,w) - C*(i-1)
Qrate=(vh*mQ + vl*(1-mQ) -C*(120/80)*log(inlogQ,w)
- C*(i-1))/(C*(120/80)*log(inlogQ,w) +10000 + C*(i-1))
}
else
{
pQ=-10000
}
```

```
max_profit = max(pq,pQ)
m_rate   = max(Qrate,qrate)
mp[i]=max_profit
if(scen[i]==1)
{
cnt[1]=cnt[1]+1
uph = 100 + (w^(cnt[1]*(80/120)))/(cnt[1]*120)
}
else if(scen[i]==2)
{
cnt[2]=cnt[2]+1
}
else if(scen[i]==3)
{
cnt[3]=cnt[3]+1
upl = 100 + (w^(cnt[3]*(10/120)))/(cnt[3]*120)
}
else
{
cnt[4]=cnt[4]+1
upl = 100 + (w^(cnt[4]*(80/120)))/(cnt[4]*120)
}
cur_profit[i]=vh*(uph/(uph+upl)) + vl*(upl/(uph+upl)) - C*i
mrr=cur_profit[i]/(C*i)
if(cur_profit[i]>=max_profit)
{
break
}
else if(m_rate<.2)
{
break
}

}
if(mrr>=.2)
{
fin_profit=20000-C*i
}
else
{
fin_profit=-C*i
}
l = list(fin_profit,mrr,cur_profit,mp,cnt)
return(l)
```

```
}

rs1=calcSigmoid2(scen1,20000,-9000,100)
fp1  = rs1[[1]]
mrr= rs1[[2]]
cp  = rs1[[3]]
mp1 = rs1[[4]]
cnt1 = rs1[[5]]

rs2=calcSigmoid2(scen2,20000,-9000,100)
fp2  = rs2[[1]]
mrr2= rs2[[2]]
cp2  = rs2[[3]]
mp2  = rs2[[4]]
cnt2 = rs2[[5]]

rs3=calcSigmoid2(scen3,20000,-9000,100)
fp3  = rs3[[1]]
mrr3= rs3[[2]]
cp3  = rs3[[3]]
mp3  = rs3[[4]]
cnt3 = rs3[[5]]
```