Creative Destruction in Multi-Source Marketplaces: Exploring Factors Influencing Success or Failure in Multi-Sided Marketplaces

by

James A Barkley
Bachelor of Science Computer Science
University of Colorado, Boulder

Submitted to the Faculty in partial fulfillment of the requirements for the degree of Master of Science in Engineering and Management at Massachusetts Institute of Technology January, 2015

© 2015 James A Barkley. All Rights Reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

All others may use this work under the terms and conditions of the Creative Commons BY 4.0 License.

Author: ....................................................

Signature redacted

James A Barkley
System Design and Management Program

Certified by: ................................................

Signature redacted

Dr. David Robert Wallace
Thesis Supervisor
Professor of Mechanical Engineering, MacVicar Faculty Fellow

Accepted by: ................................................

Signature redacted

Patrick Hale
Chairman & Senior Lecturer, Engineering Systems Division
Director, System Design and Management Program
Creative Destruction in Multi-Source Marketplaces:
Exploring Factors Influencing Success or Failures in Multi-Sided Marketplaces

By
James A Barkley

Submitted to the System Design and Management Program on January, 2016 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Engineering and Management.

Abstract
This work explores key factors that influence the patterns of creative destruction in Multi-Sided Platforms (MSPs) with the intent to determine which controls are likely to lead to particular patterns of success or failure of such a platform in the marketplace. This work builds on previous research in crowd-sourcing and multi-sided marketplaces by examining previously discovered factors in the marketplace to understand their impacts, especially as they act as determinants to an MSP’s success or failure. Eleven key factors of business strategy, technology strategy, and awareness-building were identified through an extensive literature review; dynamic simulations and uncertainty modeling were used to assess the level of influence of these factors. Simulation experiments for Facebook and Twitter were conducted and compared to historical adoption and financial data of both platforms, along with a hypothetical case study and a sensitivity analysis of all variables.

Implications for future research are that more study of user motivations for value creation and their impact is needed. Furthermore, the macro-economic dynamics add complexity, but are critical to understanding creative destruction. Implications for business leaders are that special attention should be given to anything that can enhance or inhibit Contact Rate and Adoption Fraction, and that when assessing tradeoffs, entrepreneurs should remember this fundamental tenet: enhance adoption incentives and limit adoption inhibitors.

The results of this research suggest that indirect network effects behave as amplifying or inhibiting forces acting on direct network-effect forces, and can be controlled through policy, or in some cases managed as constraints. This study finds that the key elements to focus on for understanding, forecasting, or optimizing a Multi-Sided Platform in the wild are: Coopetition, Content Per User, Awareness Effectivity, Revenue Per User, Cost Per User, Market Competitiveness, and Content Quality.

Thesis Supervisor: Dr. David Robert Wallace
Title: Professor of Mechanical Engineering, MacVicar Faculty Fellow
Acknowledgments

Thank you to both The MITRE Corporation and UI LABS for sponsoring my work at the Massachusetts Institute of Technology. Their flexibility and commitment to investing in an educated workforce should be a standard for all companies.

Thank you to my advisor, Dr. David Wallace, whose commitment to his students is legendary, and who exhibited great patience and wisdom in the face of every roadblock, paralysis, and soul-searching moment.

Thank you to my family: Mom, Dad, and Kristen, for constantly pushing me to dream the impossible dream and to strive for achievement in life, and without whose sacrifices my betterment would not have been possible. Especially, thanks to my loving wife Laura, an intellectual titan as well as a steadfast and immutable force in my tumultuous existence.

Thank you also to my extended family: The System Design & Management cast of characters including professors, colleagues, classmates and playmates. You have been my inspiration. Especially, thanks to the SDM administrators Pat Hale and Joan Rubin, who gave me limitless opportunity, and wise counsel.

Thank you to Joe Zajic, Harry Sleeper, and Bill King. I’ve been blessed with amazing role models in my career who have been both the cause and the antidote of my greatest challenges! I owe them a debt of gratitude!

To laugh about the mornin’,
Keeping midnight’s way,
Erstwhile I adorning,
Crown of Delphic, nebula
# Contents

Chapter 1: Introduction ........................................................................................................... 6
  Section 1.1: Background & Motivation .................................................................................. 6
  Section 1.2: Summary of Key Findings .................................................................................. 7
  Section 1.3: Research Scope ................................................................................................ 10
  Section 1.4: Research Objectives .......................................................................................... 11
  Section 1.5: Methodology & Approach .................................................................................. 12
    Section 1.5.1: Methodology & Approach Overview .......................................................... 12
    Section 1.5.2: Tools ............................................................................................................ 13
    Section 1.5.3: Core Framework .......................................................................................... 14
    Section 1.5.4: Literature Review ......................................................................................... 15
    Section 1.5.5: Modeling Approach ...................................................................................... 15
    Section 1.5.6: Logistic Growth Models ............................................................................... 17
    Chapter 1.5.7: Bass Diffusion Model .................................................................................. 19
    Chapter 1.5.8: Data and Model Sources ............................................................................. 20
  Chapter 2: Literature Review & Previous Research ............................................................... 22
    Section 2.1: Crowdsourcing & Innovation Research ......................................................... 24
    Section 2.2: Multi-Sided Platforms & Markets Research ............................................... 26
    Section 2.3: Psychology of User Motives Research ......................................................... 28
    Section 2.4: Open Source and Multi-Source Research .................................................... 31
    Section 2.5: Key Findings of Previous Research ............................................................... 33
  Chapter 3: System Models .................................................................................................... 36
    Section 3.1: Core Model Variables ..................................................................................... 37
    Section 3.2: Core Model Development .............................................................................. 39
    Section 3.3: Key Influencing Variables .............................................................................. 43
      Section 3.3.1: Directly Represented Factors .................................................................. 44
      Section 3.3.2: Complex Factors for Representation ....................................................... 45
    Section 3.4: Extended Model Development .................................................................... 49
  Chapter 4: Simulation Experiments ....................................................................................... 57
    Section 4.1: Using the Model ............................................................................................. 57
      Section 4.1.1: Model Interface ......................................................................................... 57
      Section 4.1.2: Advanced User Options ........................................................................... 59
      Section 4.1.3: External Data and Model Time Steps ....................................................... 60
Chapter 1: Introduction

Section 1.1: Background & Motivation

The term Creative Destruction is taken from the work of Austrian-American economist Joseph Schumpeter. Sometimes called “Schumpeter’s gale”, creative destruction is a theory of economic innovation and the business cycle (Schumpeter, 1950) that describes how new economic orders are continually created and prior economic orders are continually destroyed. Much study has been given to understanding the phenomenon of creative destruction since Schumpeter’s original work was published in 1942 (third and final edition in 1950). However, with the rise of the Internet and web-based services, the economic forces and prevailing market winners have become even more disruptive and uncertain (Foster & Kaplan, 2001), (Hagiu & Wright, Multi-Sided Platforms, 2015).

In their 2001 book by the same title (Creative Destruction), Sarah Kaplan and Richard Foster draw from a McKinsey database of over 1,000 companies across 15 industries over a 38-year period and propose that the average lifetime of a company in the S&P is decreasing and will be ten years by 2020 (Foster & Kaplan, 2001). Kaplan & Foster propose levers and strategies for making businesses longer-lasting and higher-performing. However, the majority of organizations studied in this work are traditional businesses, such as chip manufacturing, pharmaceuticals, and cookware. Companies in service industries are excluded. This study is representative of the majority of work up into the early 2000’s.

Online marketplaces and widespread connectivity have led to the rise of numerous platforms, typically using a business model commonly referred to as a Multi-Sided Platform, or MSP. MSPs – such as Uber, Facebook, and Groupon – bring together two or more groups of
customers that have interdependencies (Ankaraju, 2010). This is very different paradigm than traditional operating modes for a company (e.g., vertically integrated firms, resellers, or input suppliers), with one of the biggest differences being indirect, or side-channel, network effects (Hagiu & Wright, Multi-Sided Platforms, 2015).

There are many business strategy, technology strategy, and awareness management choices that influence the behavior of markets and market actors in MSPs:

I hypothesize that a small, identifiable set of strategy, policy, and technology choices are key drivers that can be used to accurately forecast and control user growth, cost, and revenue of an MSP.

I also hypothesize that users are tied directly to revenue and valuation and that cost/user and revenue/user are key metrics for assessing the viability of a business venture.

Section 1.2: Summary of Key Findings

In this research, a conceptual framework was developed for understanding multi-sided platforms and the dynamics of both direct and indirect network effects, including what constraints and strategic decisions influence these dynamics. Eleven key factors of business strategy, technology strategy, and awareness-building were identified through an extensive literature review, and dynamic simulations and uncertainty modeling were used to assess the order of magnitude of influence of these factors. Simulation experiments for Facebook and Twitter were conducted and compared to historical adoption and financial data of both platforms to validate the model, along with a hypothetical case study and a sensitivity analysis of all variables.
The core framework, key factors culled from the literature review, and system dynamics model developed are valuable artifacts which can serve to explore market factors and key decisions for an MSP; the historical case studies with Facebook and Twitter data can be forecasted with surprising accuracy. The model has been published with a special interface that allows easy experimentation by providing all inputs with sliders which show changes in real-time to the output variables and custom graphs.

Table 1.1 – Top eleven key determinants in MSP success
The findings of this research support the hypothesis that these factors are worth careful consideration, and that they can be predicted and, to some extent, controlled for improving outcomes. This research suggests that the key elements to focus on for understanding, forecasting, or optimizing an MSP in the wild are those which are sensitive but controllable: Coopetition; Content Per User; Awareness Effectivity; Revenue Per User; and Cost Per User; Market Competitiveness; and Base Creation Quality.

Larger context of global macroeconomics are an important part of the story that were beyond the scope of this research, but are implied as a direction for future research. In addition, motivational forces for content creation and user communities are ripe for additional exploration.

For entrepreneurs and business-builders, strategies for maximizing the adoption factors such as ambassador programs, referral schemes, and advertising offer a huge return on investment, especially if done early. Decisions about pricing policies, community selection, and barriers to entry will limit as much as possible any inhibiting effects on adoption. For example, the users on the side of the marketplace that stand to benefit most from interactions/transactions with other sides of the marketplace are also the most likely to tolerate fees.

1 References to direct model variables are given in All Caps Notation.
Section 1.3: Research Scope

This research explores the hypotheses that a small set of key decisions are the primary drivers in multi-sided platforms and can be used to accurately forecast growth, cost, and revenue of an MSP. This core set of “levers”, it is hypothesized, are the primary influencers of:

- user adoption,
- user retention,
- support costs for the MSP owner/operator,
- revenue generation for the owner/operator, and
- value creation within the marketplace

A core framework has been developed for exploring these hypotheses, along with a set of system dynamics models, monte-carlo simulations, uncertainty and sensitivity analyses, and financial metrics over time.

In modeling and simulation, defining the system boundary and selecting the right variables at the right level of fidelity is always a critical challenge with no hard and fast rules. In addition to the core model elements based on existing mathematical models, the key influencing factors for modeling have been derived from a literature review and assessment (Chapter 2). This work has been scoped to explore a dozen or less key influencing factors as well as a set of 20 or less core model variables. This strikes the right balance between achievability and serious research.

Several revenue streams and models are available for MSPs, such as charging for the service outright, charging for only premium features (so-called “freemium” model), or through offering complementary services, and understanding these is a core objective of this work; however, the scope of this effort only demonstrates how this framework and model set can be applied to assess the revenue streams in a real-world scenario.

The models are tested using public-sourced data from existing MSPs. This includes information such as public financial statements and user growth numbers, but excludes working closely under NDA with any particular company in order to incorporate proprietary data such as...
operating costs. I have also tested and tuned the model with data from my current work at UI LABS, but none of that data will be made available under this work.

Section 1.4: Research Objectives

MSPs are still a fairly new concept, and are an emergent phenomenon in many sectors of the economy. My objective is to build a framework and a set of tools to validate the hypotheses – a small set of strategy, policy, and technology choices are the key drivers which can be used to accurately forecast and control user growth, cost, and revenue of an MSP; users are tied directly to revenue and valuation and that cost/user and revenue/user are key metrics for assessing the viability of a business venture) – in the hopes that it is a tool which can be used by the author and others who are introducing an MSP into the marketplace. I’m particularly interested in exploring the potential for MSPs in the manufacturing sector, which is undergoing a disruptive digital revolution right now.

The outcomes and deliverables of this thesis, also shown in figure 2, are:

- a literature review and case study comparison; Chapter 2
- a proposed set of core factors for consideration; Section 2.5, Section 3.2
- a set of executable models with simulation results and interpretation; Section 3.3
- a thesis document entailing the background, methodology, results, and implications for future research, among other sections;
- and a web site for providing access to all research artifacts under an open source license. http://cdimm.net/
Section 1.5: Methodology & Approach

Section 1.5.1: Methodology & Approach Overview

The methodology and approach started with defining the core framework for investigation, which was done iteratively in conjunction with a thorough literature review. The literature review was performed based on a perspective decomposition which formed the categories of literature surveyed. These categories were likewise informed in conjunction with the development of the core framework.

During the literature review, key influencing factors were tagged, extracted, and placed into a scoring matrix across the literature categories and grouped by the core framework pillars using a systematic method of concept classification in text analysis known as “coding.” The key
influencing factors have been prioritized by the number of mentions across various literature categories. In order to appropriately scope the effort, I've only chosen to model the 11 highest-scoring factors. This is a little below the obvious "cut-line" for the top group of scores, but I felt it important to incorporate the next two highest.

These key influencing factors then became the variables for modeling and simulation. Two initial model types (Logistic Growth and Bass-Diffusion) were explored but only one was selected for final modeling and experimentation in order to keep the effort in line with the scope and resources.

Finally, several validation and sensitivity experiments were explored with the final model, followed by interpretation of the results and conclusions. A graphical view of the methodology and approach can be seen in Figure 1.3 below.

![Figure 1.3 - Overview of methodology and approach](image)

**Section 1.5.2: Tools**
In addition to this paper-based framework, the following tools were used in various ways to build out mathematical and executable models for exploration of this basic framework:

- **System Dynamics** – System dynamics is a science that emerged in the 1950's to aid understanding of complex socio-technical systems with a high degree of interdependence and circular causality (System Dynamics Society, 2015). I used this for both modeling and simulation using Vensim DSS from Ventana Systems.

- **Uncertainty Analysis** – Uncertainty Analysis is the science of identifying and modeling key sources of uncertainty in a system to understand the distribution of possible outcomes and the sensitivity of the system overall to changes in certain parameters (de Neufville & Scholtes, 2011). VenSim and The DecisionTools Suite from Palisade Corporation are examples of tools available that use assist in use of these techniques.

- **Financial Modeling** – Financial modeling is a way of mathematically representing a particular financial situation, usually for purposes of forecasting financial performance or valuation over time. I have incorporated financial forecasting elements into the Vensim models and have done some auxiliary work using Excel spreadsheets for analysis of financial measures (e.g., NPV, payback period, etc.) as well.

**Section 1.5.3: Core Framework**

Our core framework for exploring MSPs begins with the premise that traditional network effects are at work in MSPs, but so are indirect or “side-channel” network effects. The traditional network effects function as in a traditional marketplace or product adoption scenario, but the indirect network effects occur when different sides of the marketplace interact with each other. This view is supported strongly by the literature review in Section 2.

In addition to these core elements, this framework needs to include “levers" that the platform builder/operator has influence over. In this framework, these have been categorized into three core pillars of the model: Business strategy, technology strategy, and “net positive awareness.” Some examples of each of these can be seen in Figure 1.4 below.
Section 1.5.4: Literature Review

The key variables for modeling have been culled from a comprehensive literature review, in which I have extracted and prioritized the top variables that show up routinely in publications and research. These were placed into a scoring matrix across the literature categories and grouped by the core framework pillars. The key influencing factors have been prioritized by the number of mentions across various literature categories. Sections 2.5 and 3.2 of this document contain the analysis and resulting key influencing factors for modeling.

Section 1.5.5: Modeling Approach
The general approach to modeling here is focused around users of the platform. User population is a key factor in revenue, support costs, carrying capacity, user growth, and value generation, and almost all variables in which people are concerned with the outcome; therefore, user population adoption and retention must be central to the model. In addition, user population must be linked to almost all relevant factors.

The variables are broken down into two groups:

- **Core model variables** – These variables represent the basic features of the foundational model like "users" as well as most off the output variables which are desirable to predict, such as "revenue."

- **Key influencing variables** – These variables are mostly input and intermediate variables that represent the levers in our core framework that have influence over the output variables which are desirable to predict or control.

Two foundational models were initially explored: the logistic growth model and the Bass model of technology diffusion, each of which is discussed in more detail below. These two models were chosen as the foundation for this work because of their classic role in product adoption and population growth, and because they align nicely with the core framework. These models serve as a good foundation, but are fairly basic and contain a number of assumptions. In this research, an initial investigation of both models was performed, and ultimately the Bass-diffusion model was selected as the final model for development. Using this model, the basic approach to population modeling has been built on, including incorporated feedback elements that have strong acausal relationships with the key variables and other previously unconsidered factors. In other words, the foundational model was extended by layering on the key variables culled from the literature review in Section 2.5 of this work. For example, user populations are linked to many other factors through intermediate variables such as revenue per user, support costs per user, deterrents to adoption, number of sides of the MSP, etc.

The models have been validated by being fit to real-world, publicly available data sets, such as Facebook reported user growth and published revenue numbers. "Fit" in this case means the models have been simulated with initial conditions matching that of the projects we are simulating (e.g., Facebook). The output variables are then compared with the actual historical results, and the models are either sufficiently matched to the real data, or need to be refined

---

2 Not governed or operating by the laws of cause and effect (Oxford University Press, n.d.).
and tweaked. Multiple MSPs were fit to the models in order to guard against over-fit. Where the exact data was not available, such as internal support costs, values that yield accurate results from the models were selected and then compared those values to expected support costs from published studies on infrastructure and resource capacity.

Once the models were established and provided reasonable forecasting performance to historical datasets, the models were then simulated with various perturbations in controlled experiments where individual variables are varied to understand their impact on output. For instance, if we assume the market is expanding, what impact does this have on the user growth compared with if we assume the market is shrinking?

Finally, some sensitivity analysis is performed on these key variables in two modes – the system dynamics model and a basic financial forecasting model – to assess the beneficial envelope for the most critical variables.

All of the models and results have been reviewed with subject matter experts from engineering, system dynamics, finance, and venture capital.

Section 1.5.6: Logistic Growth Models

A common approach to modeling population growth is using a logistics function, which yields an “S-curve,” or sigmoid curve (Wikipedia, 2015); that is, growth is near-exponential at first, until saturation begins to occur and then growth slows, eventually plateauing in a logarithmic-like fashion. The standard equation for this in ecology is provided in equation 1.1

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K}\right)$$

Equation 1.1 – Classic differential equation for logistic growth

| P | Population Size |
Species and disease populations have been very successfully represented with classic logistic growth models; however, in our case of online user communities, the traditional meanings of some of these variables are no longer valid, and we’ll need to recast them into something more meaningful. For instance, population size could be measured in user accounts, active user accounts, business accounts, etc. Carrying capacity is more a measure of the computing resources required to support the population; of course, in the age of cloud computing compute resources can often be elastically scaled (assuming the platform architecture can scale with compute resource expansion). Carrying capacity frequently varies with time in ecological as well as virtual environments, so the classic equation has been modified as shown in equation 2.

\[
\frac{dp}{dt} = rP * (1 - \frac{p}{K(t)}) \tag{Equation 1.2}
\]

These types of models have successfully been adopted for representing growth of online user communities (Kairam, Wang, & Leskovec, The Life and Death of Online Groups: Predicting Group Growth and Longevity, 2012). In fact, in 2011 Peter Cauwels and Didier Sornette – both of the Entrepreneurial Risks group at ETH Zurich – fit logistic-based population models to published Facebook and Groupon data with very accurate results, and used that as a basis to assess whether or not Facebook and Groupon were overvalued in the marketplace (Cauwels & Sornette, Quis Pendit Ipsa Pretia: Facebook Valuation and Diagnostic of a Bubble Based on Nonlinear Demographic Dynamics, 2012).

To explain this further, let’s take an example: in the figure 1.5 we see a common approach in System Dynamics to modeling this type of equation. We can see in the model that carrying capacity is represented as a constant expression encapsulated in the Adoption Fraction. Not only does this not
account for time-varying carrying capacity, it ignores any relationships between adopters and carrying capacity. In our environment, one could argue that carrying capacity is a function of cost, which needs to be supported by revenue generated; therefore, in order to scale the carrying capacity one would need more users, higher revenue per user, or both.

Chapter 1.5.7: Bass Diffusion Model

The bass diffusion model was first developed in 1963 by Professor Frank Bass at the Krannert School of Management at Purdue University (Bass, 1963) as a way to explain the diffusion of products through the marketplace. It has become one of the most popular models for new product growth and is widely used in marketing and strategy, among other things (Sterman, 2000, p. 332).

The core mathematical model, known as the Bass Model Principle, is described as follows (Bass’s Basement Research Institute, 2010):

\[
\frac{f(t)}{1-F(t)} = p + \frac{q}{M[A(t)]} \quad \text{Equation 1.3 Classic equation for the Bass Model Principle}
\]

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(M)</td>
<td>Potential Market</td>
</tr>
<tr>
<td>(t)</td>
<td>Time</td>
</tr>
<tr>
<td>(p)</td>
<td>Coefficient of Innovation</td>
</tr>
<tr>
<td>(q)</td>
<td>Coefficient of Imitation</td>
</tr>
<tr>
<td>(f(t))</td>
<td>Fractional Adoption Rate</td>
</tr>
<tr>
<td>(F(t))</td>
<td>Saturation</td>
</tr>
</tbody>
</table>

The Bass’s Basement Research Institute describes how this equation should be read: “The portion of the potential market that adopts at time \(t\) given that they have not yet adopted is equal to a linear function of previous adopters.” (Bass’s Basement Research Institute, 2010)

One of the key advantages of the Bass Diffusion Model is that it accounts for the genesis of the initial population. In the logistic growth model, the population growth relies entirely on contact
rate between potential adopters and existing adopters, but this does not account for situations where existing adopters are close to zero, such as a new technology or market. Originally described as "innovators and imitators" (Bass, 1963), the bass model includes constant or flat growth due to some other force and typically attributed to advertising (innovators) as well as growth through word of mouth (imitators).

The basic system dynamics representation of the bass model can be seen in figure 1.6 below and is a very good foundation, but still affords opportunity for improvement.

As with the logistic growth model, the carrying capacity is encapsulated in the Adoption Fraction as a constant extrinsic value. The bass model asserts that the total population is zero, which is an assumption that the size of the market is constant. It also does not include the response of the market to pricing, nor does it encompass financial characteristics such as operating costs or revenue generation.

**Chapter 1.5.8: Data and Model Sources**

All data presented in this document were taken or derived from materials available either publicly or from published literature. For instance, financial data in some cases comes from annual statements while user growth is taken straight from corporate publications or market research reports.
Much of this work is inspired by System Dynamicist Tom Fiddaman who built upon the work of Cauwels and Ddier (cit). The two models I reviewed and exercised from Fiddaman and the resulting analysis are freely available (Fiddaman, 2011). Much of this work is also inspired by System Dynamicist John Sterman. I have expanded and modified the basic Bass-diffusion model provided by John Sterman in his book, Business Dynamics (Sterman, 2000), which includes sample models on CD as a companion to the textbook.

In fact, published models on Facebook and Groupon from Fiddaman were explored as potential starting points for the modeling portion of this effort, but in the end were not a good fit. The objective was different, which led to fairly niche set of parameters. Still, these models were very informative, and an example can be seen below in figure 1.7.

![Figure 1.8 - Example System Dynamics model for looking at Groupon customer growth (Fiddaman, 2011)](image)

The final model shown below and used for experimentation was built from scratch for this effort. All works produced for this effort are freely available (Barkley, 2015) and are licensed under the terms and conditions of the Creative Commons Attribution License 4.0 (Creative Commons Corporation, 2015).
Chapter 2: Literature Review & Previous Research

In order to determine the right parameters and variables to be modeled for the analysis, an extensive review of existing research was conducted. Both primary and secondary sources were surveyed. Primary sources were most commonly reached using the MIT library online journal and periodical resources, but occasionally other sources were consulted, such as the Harvard Business School paper archive and public internet. Secondary sources were those that were referenced in the primary sources or referred from subject matter experts, and included not only journal articles but books, working papers, and interviews. One interview was conducted as well.

A perspective decomposition was performed on the initial core framework (Section 1.5.3), assessing different lenses through which to view a multi-sided platform and leading to the breakdown of the literature into several distinct categories. These were refined as articles reviewed led to deeper insights into the perspectives on how to view the core framework. In the end, books, papers, journal articles, interviews, etc. were reviewed and binned into five major categories: crowdsourcing & innovation; open source & multisource; multi-sided platforms; and psychology of user motives. These categories ultimately drove the keyword and search criteria for ongoing review.

Out of the millions of results, articles were sorted by relevance and then by date. The top ten or so articles were selected, and then augmented by the secondary sources. The tally can be seen in the table 1:

<table>
<thead>
<tr>
<th>Category of Research</th>
<th>Number of Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowdsourcing &amp; Innovation</td>
<td>45</td>
</tr>
<tr>
<td>Multi-Sided Platforms</td>
<td>12</td>
</tr>
<tr>
<td>Open Source &amp; Multi-Source</td>
<td>8</td>
</tr>
<tr>
<td>Psychology of User Motives</td>
<td>9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>74</strong></td>
</tr>
</tbody>
</table>
Table 2.1 – Literature sources consulted by “MSP perspective” category

For the most part these categories are self-explanatory, but a section on each is presented below. Each category section provides a definition of the category and a brief write-up on key findings, including pull-quotes and highlights from specific sources.

As each article was reviewed, key influencing factors that could be modeled were annotated and extracted. Although a formal codebook was not developed, factors were refactored into a higher-level “codes” that could be counted across the different literature sources. This process is sometimes referred to as “coding,” and is a classic approach to the systematic coding of text in qualitative data analysis (Corbin & Strauss, 1990).

Using this approach, the number of times a code appears is counted up by category, and then summed across all categories to show a final scoring matrix. For instance, one paper discussed “confidentiality” while others discussed issues around intellectual property – these were combined into a single code of “intellectual property” with a resultant score that was additive across the number of papers that mentioned it. Each variable was also mapped into one of the three legs of my framework (business strategy, technology strategy, and awareness).

The resultant scoring matrix was used to give a polynomial score for each variable, which were then rank-ordered and selected for modeling. In this case, all polynomial terms had equal weighting, but future research could weight the categories differently, perhaps resulting in different final variables. 73 total influencing factors were pulled from the literature review; the full scoring matrix can be viewed in Appendix B.

It is worth noting that there are a disproportionate number of references in the Crowdsourcing & innovation category. This is likely due to two things: the fact that “Crowdsourcing” and “Innovation” research were grouped into the same category, coupled with the fact that both of these topics independently are some of the most widely researched topics in business communities. These are big topics and many of the studies, conclusions, and parameters of this topic overlap with the other categories in my literature review. In fact, all of the categories frequently have some overlap among their references, and for this effort less concern was given
to even distribution across the categorical perspectives than in ensuring commonality of variables and importance across all references reviewed.

Section 2.1: Crowdsourcing & Innovation Research

The term "crowdsourcing", first coined by Jeff Howe in 2006 (Howe, The Rise of Crowdsourcing, 2006) has been a favorite topic of recent research studies, while studying "innovation" has long been a tradition in the research community.

Much of the literature in this space focuses on defining these concepts (Estelles-Arolas & Gonzales-Ladron-de-Guevara, 2012), while others try fitting various frameworks to the concepts in order to characterize them (Turian, 2012), assess the current market (Massolution, 2013), and provide guidance on when and how to use various instruments for business value (crowdsourcing.org, 2012). In addition, there are the occasional deeper research studies exploring thesis topics (Mekler, 2014) or using computer simulations to model crowd behavior (Prpic, Jackson, & Nguyen, A Computation Model of Crowds for Collective Intelligence, 2014). Finally, several case studies are available; Innocentive (Lakhani K., Innocentive Case Study, HBS, 2009) and NASA (Gliedman, Smith, & Burris, 2013) are two examples that were reviewed in this study.

Of the many references reviewed in this area, the most valuable were the in-depth research studies and the business value papers. The market analysis and case studies, while interesting and informative, offered little in the way of concrete variables that could be used in modeling and simulation. The body of knowledge in this space was focused largely on the usefulness of crowdsourcing for businesses and innovative potential as often measured by the effectiveness of distributed teams (Sukkasi, 2004), the effect of non-experts and team diversity in these projects (Mekler, 2010), and the ability to tap into larger cohorts for one-time activities.

Howe proposes that a “renaissance of amateurism” is fueling the potential of the crowd, and it is important to pick the right model, pick the right crowd, and offer the right incentives (Howe, Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business, TBD). These are recurring themes that raised a common set of questions such as:

- Internal vs. external crowds
- Reward incentives
- Pricing schemes
- Quality

Mekler writes about the importance of diversity: “When crowds are diverse, collaborative crowds outperform the nominal crowd. However, when crowds are highly homogenous, collaboration can reduce ideation performance” (Mekler, 2014). The results of a DARPA study support this by pointing to the importance of “openness” in crowds and platforms: “The openness of the marketplace will allow input from a wide range of contributors with varying skill levels” (Cao, Wallace, Beckmann, & Citriniti, 2012).

Quality of contributions and contributors is another key theme that we see, while others caution us about the challenges of intellectual property (Laasonen, 2014). Cheung gives us an interesting measure of what to expect in terms of contributing users versus regular users:

“Not all users were as interested in creating content as consuming them, however. In fact, Empirical observations suggested that typically only 1% of the users in a virtual community actually create new content, 10% would modify that content, while the remaining 89% would view the content without contributing at all [28]. While the actual percentages may vary depending on the user demographics and content type, these estimates nonetheless provide valuable insights regarding user participations. Although this “89:10:1 ratio” seems utterly unfavorable to websites that depend on user-generated content, that 1% is often a large number already for high-traffic websites with millions of visitors” (Cheung, 2012).

In total, 117 variables were pulled out of the references reviewed; these were rationalized into a final list of 51 total variables using the coding described in the methodology section. Some of the top factors from this area include:

- Crowd incentives
- Crowd selection (diversity, specialization, size, etc.)
- Platform selection (feature set, carrying capacity, etc.)
- Communication among the crowd
- Charging/pricing model
Section 2.2: Multi-Sided Platforms & Markets

Research

The topic of multi-sided platforms appears to be fairly niche, yet there is a healthy set of literature around the topic; some of this does not use that specific terminally, just instead casts the topic as the economics of online marketplaces. An MSP is a generalization of the two-sided platform concept where two or more sides of a marketplace or enabled to interact with each other.

Haigu, et al. discuss the differences in definitions, and tell us the defining characteristics are "they enable direct interactions between two sides of the platform" and "each side is affiliated with the platform". It is interesting to assess how this business model differs from others.

Numerous journal articles and business reports have discussed the challenges and tradeoffs with governance in these types of multi-sided platforms. Governance here is defined as a regulation of third-party action by the platform itself. At a high level, a platform's choice of tighter governance rules reflects a trade-off of quantity in favor of quality. An example of this is the difference between Apple iOS and Android in the marketplace. Here are three examples of differing governance policies:

- **Hardware**: Apple controls production in a vertically integrated fashion; Google, on the other hand, in addition to producing their own hardware platform, allows any hardware vendor to build an Android-compatible phone.
- **App store**: Apple has careful vetting of "apps" and app producers; Google only recently instituted app vetting which is still much more limited in scope.
- **App development**: Google allows 3rd party toolchains for building applications, while Apple requires you use only their SDK.

"We study the economic tradeoffs that drive organizations to position themselves closer to or further away from a multi-sided platform (MSP) business model, relative to three traditional
alternatives: vertically integrated firms, resellers, or input suppliers." Haigu and Wright tell us. “For instance,” they go on, “Amazon started off as a pure retailer but has moved closer to a MSP model over time by enabling third-party sellers to trade directly with consumers on its website.”

MSPs create value by enabling interactions among customer groups, and Andrei Haigu go on to tell us that the key value of MSPs is by decreasing the cost of transactions and searches (Haigu A., Successful Strategies for Multi-Sided Platforms, 2014). This can take place in service markets that are purely digital (e.g., iTunes) or knowledge products (e.g., Wikipedia), but many are about brokering goods between buyers and sellers, such as Airbnb, eBay, Xbox, Uber, Facebook, Amazon, and Rakuten.

Platforms are powerful concepts, and very different than other business models. Michael Cusumano proposes in his book (Cusumano, 2012):

“Managers (at least in industries affected by digital technologies as well as ‘network effects’ more broadly) should move beyond conventional thinking about strategy and capabilities to compete on the basis of platforms, or complements to another firm’s platform.”

Some interesting modeling approaches have been taken, including attempting to fit historical data from Facebook and Groupon using logistic growth models. Pricing and user motivations are themes that show up, and lots of good strategy advice is given. For instance, in his 2014 paper, Andrei Hagiu proposes key strategic points for successful MSPs (Hagiu, 2014). He suggests that the following key questions are the core drivers of success or failure in MSPs:

- How many sides to bring on board?
- What are the right design decisions for feature selection?
- What is the right pricing structure?
- What are the appropriate governance rules?

Some of the top factors emerging from this portion of the literature include:

- Crowd selection
- Transparency/openness
Section 2.3: Psychology of User Motives Research

A significant amount of work has been conducted to understand motivational forces and innovation capacity in individuals (Stock-Homburg, von Hippel, & Oliveira, 2014). One study examined Innocentive prize challenges and interviewed respondents (Lakhani K. R., 2006). Table 2.8, below, summarizes these findings.

Table 2.8: Correlations Between Variables Predicting Solver Being a Winner (N = 295 Respondents)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 RTP Problem Type</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Time to develop solution</td>
<td>0.09</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Money</td>
<td>0</td>
<td>-0.06</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Extrinsic motivation</td>
<td>0.01</td>
<td>0.07</td>
<td>-0.11†</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Intrinsic motivation</td>
<td>-0.05</td>
<td>0.05</td>
<td>-0.21*</td>
<td>0.32***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Beating other solvers</td>
<td>0</td>
<td>-0.07</td>
<td>-0.02</td>
<td>0.27***</td>
<td>0.26**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Unsatisfactory job</td>
<td>0.07</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.18***</td>
<td>-0.04</td>
<td>0.06</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Had free time</td>
<td>0.01</td>
<td>-0.13*</td>
<td>-0.08</td>
<td>0.05</td>
<td>0.08</td>
<td>0.03</td>
<td>0.25**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expertise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Interest Count (at registration)</td>
<td>-0.03</td>
<td>-0.06</td>
<td>0.12**</td>
<td>-0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.03</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10 Problem distance with field of expertise</td>
<td>-0.01</td>
<td>0.11†</td>
<td>-0.03</td>
<td>0.05</td>
<td>0.09</td>
<td>-0.01</td>
<td>0.11†</td>
<td>-0.01</td>
<td>0.13*</td>
<td>1</td>
</tr>
</tbody>
</table>

† significant at 10%; * significant at 5%; ** significant at 1%; *** significant at 0.1%

Table 2.2 – Correlations between variables predicting a crowd source challenge solver being a winner where N = 295 respondents (Lakhani K. R., 2006)
Eric von Hippel has also been known for his work on end-user motivation (von Hippel, 2005). In one of his most recent works he joined forces with Dr. Ruth Stock-Homburg, a psychologist and professor at Darmstadt University in Germany. Pushing von Hippel's previous work even further, they scientifically explored not only motives of end users, but also the personality traits behind those motives (Stock-Homburg, von Hippel, & Oliveira, 2014). One key finding is around the classic "hedonic" vs. "utilitarian" motives that von Hippel has pioneered. The relationship between hedonic user motives and utility are characterized by an inverted U-shape (pg. 2), meaning that hedonic motivations contribute toward utility as long as they are not the dominant motivator, but once they are they decrease the utility. These findings are summarized in the table below (Stock-Homburg, von Hippel, & Oliveira, 2014, pp. 17-20)³.

<table>
<thead>
<tr>
<th>Motivation</th>
<th>Innovative dimension</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilitarian</td>
<td>Utility</td>
<td>Positively correlated</td>
</tr>
<tr>
<td>Utilitarian</td>
<td>Novelty</td>
<td>Negatively correlated</td>
</tr>
<tr>
<td>Hedonic</td>
<td>Utility</td>
<td>Inverted U-shape</td>
</tr>
<tr>
<td>Hedonic</td>
<td>Novelty</td>
<td>Positively correlated</td>
</tr>
</tbody>
</table>

Table 2.3 – User motivation (hedonic vs. utilitarian) and correlations with innovative dimension (novelty vs. utility) (Stock-Homburg, von Hippel, & Oliveira, 2014)

In addition, they found that the key forces included:

- **Feedback** – Feedback is an important force for those who simply want to learn (advance their skills), are or seeking personal growth or social interactions for enjoyment.
- **Recognition** – Recognition is a form of social reward, and has long been known to be a motivating factor. Many people would rather have prestige over fortune.
- **Energy generation/drain** – Dr. Stock-Homburg did not go into great detail on this, nor does the paper, but I understand it to be related to a person’s extraversion. That is, extroverted people often are empowered by interaction with others, while introverts are often depleted by these types of interactions.

And the resistance forces included:

³ Note that this table was not taken directly from the paper, but rather synthesized from key findings in the paper.
• Genuine uncertainty
• Cannibalization
• Margin erosion
• Expectations of failure

As additional background, it is interesting to note that the background personality theories include:

• **Trait Theory** - In trait theory, personality is fundamentally a set of traits, which may or may not be immutable. A personality, then, is comprised of some number of traits. Dr. Stock-Homburg refers to these as "trait dimensions" and specifically described the five-factor model. In the five-factor model, there are five basic dimensions to a personality: openness to experience, extraversion, conscientiousness, agreeableness, and neuroticism (McCrae & Oliver, 1991).

• **Human Growth** – Human growth theory is about the personality as a key to self-fulfillment. Some seminal work in this model is the theory of personal behavior which describes two fundamental goal strivings: purposefulness, the awareness and purposeful directedness of an individual; and experienced meaningfulness, the "perceived significance" to an individual of their work activities (Barrick, Mount, & Li, 2013).

• **Cognitive Filtering** – Cognitive psychology filter theory treats says that we filter external stimuli. For example, the Cognitive-Affective Personality System (CAPS) as described by Walter Mischel. In this theory, the behavior of a person is a function of their perceptions, and not strictly core personality traits. What is interesting about this is that it gives us a plausible way to account for why people with the same personality trait dimensions (and interactions among dimensions) still behave differently when confronted with the same situation (Walter & Shoda, 1995).

Several authors point to the critical elements of the communal infrastructure that holds together the fabric of shared motivations and group behaviors. For instance, "commons" are the shared

---

4 Genuine uncertainty is a term meaning that an individual, in addition to not knowing the probabilities of possible outcomes, does not even know the possible outcomes themselves (Loasby, 1976).
resources of a community, and frequently play a critical role in user motivations for creative works (Sukkasi, 2008).

The top parameters in this research area include:

- Personal motivations and incentives
- Transparency/openness
- Concerns for Intellectual Property
- Profit
- Communication
- Cost of participation (a.k.a. charging/pricing model)
- Low barrier to entry
- Operating costs

**Section 2.4: Open Source and Multi-Source Research**

Open source software, and now hardware, has become the subject of a great deal of scrutiny and study. These terms have become quite loaded, and for this research I'd prefer to think of these types of activities as co-development of creative works. This has been a successful paradigm in many industries (Golden, 2005), but is new in the world of hardware, which is governed by patents more than copyrights (Open Source Hardware Association, 2015). Open source strategies can often yield a better result in terms of quality and efficacy, but are still sometimes perceived to provide less business value because of the sharing of Intellectual Property.

Karl Fogel, in his landmark book “Producing Open Source Software”, offers several pearls of wisdom about what makes open source software projects function from a motivational and societal perspective:

“Free software is a culture by choice. To operate successfully in it, you have to understand why people choose to be in it in the first place. Coercive techniques don't work. If people are unhappy in one project, they will just wander off to another one. Free software is remarkable even among intentional communities for its lightness of investment. Many of the people involved have never actually met the other participants face-to-face. The normal conduits by which
humans bond with each other and form lasting groups are narrowed down to a tiny channel: the written word, carried over electronic wires. Because of this, it can take a long time for a cohesive and dedicated group to form. Conversely, it's quite easy for a project to lose a potential participant in the first five minutes of acquaintanceship. If a project doesn't make a good first impression, newcomers may wait a long time before giving it a second chance." (Fogel, 2006)

Fogel, Golden, and others offer several practical suggestions about how to succeed in co-development or co-engineering activities under an open source arrangement. The shared intellectual property and tasking responsibilities can be set up in a variety of different ways, and it remains true that some of the biggest drivers are of success in open source are around transparency and intellectual property, and accessibility remains important for participants to join and contribute easily.

A recent interview with Richard Stallman regarding the Digital Manufacturing Commons raised some interesting questions (Stallman, 2015)5:

“Running models makes me worry. Are the models free software? And whose computer do they run on? Who chooses which version of the model to run?”

He also points out the challenge of speaking broadly about Intellectual Property (IP):

“It is a mistake to refer collectively to those three things [copyright, licenses, and patents]. That's what creates confusion.”

This fundamentally is about management of Intellectual Property (IP), and relates to not only the user motivations described earlier, but organizational motivations. Successfully enabling both the individual and the organization is difficult, and the right set of policies for managing open

5 Please note that citing this interview in no way implies that Dr. Stallman is in any way supporting open source. Dr. Stallman has made it clear that he supports the free software movement, and fundamentally disagrees with "open source".
source, shared source, and private source data and tools is a very precarious and important challenge.

Some of the top parameters in this category include:

- Transparency/openness
- Intellectual property
- Cost/pricing model
- Low barrier to entry
- User incentives
- Communication
- Social responsibility
- Legal matters and licensing
- Operating costs
- Quality
- Growth rate

**Section 2.5: Key Findings of Previous Research**

The literature review presented highlights several key areas of concern that are fitting parameters to model as variables. Many of the findings were too fine-grained to be effective for the level of our model, but when generalized give us a good set of parameters. As an example, crowd diversity is a good and important concept that we can model. Several of the references took this to a deeper level and discussed cognitive diversity, social diversity, skill diversity, etc. If we were focusing our model solely on diversity and its impact all of these would be good factors, but since are model is much higher-level and holistic, those have simply been distilled into “diversity”.

Each of these parameters was scored across each literature review category and a summed score was generated for each. In rank order, the top eleven can be seen in the scoring matrix below:
I included the top 11 even though there appears to be a clear cutline at the 7th. This is because the remaining four items I believe are very important in understanding financial performance overall of this type of platform.

In addition to the scoring matrix, I have mapped each parameter into the framework categories of business strategy, technology strategy, and awareness. The top parameters for each of these pillars of the framework can be seen below:
Table 2.6 – Top factors in the business strategy pillar of the core framework

<table>
<thead>
<tr>
<th>Type</th>
<th>Parameter</th>
<th>Crowdsourcing &amp; Innovation</th>
<th>Multi-Sided Platforms &amp; Markets</th>
<th>Psychology of User Motives</th>
<th>Open Source &amp; Multi-Source</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>business strategy</td>
<td>incentives</td>
<td>14</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>24.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>crowd selection (diversity, specialization, population size, etc.)</td>
<td>13</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>19.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>communication among crowd</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>16.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>charging/pricing model</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>15.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>transparency/openness</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>13.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>profit</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>10.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>quality</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>9.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>operating costs</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>8.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>labor pool (private, community)</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>7.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>socially responsible</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Table 2.6 – Top factors in the “Technology Strategy” pillar of the core framework

<table>
<thead>
<tr>
<th>Type</th>
<th>Parameter</th>
<th>Crowdsourcing &amp; Innovation</th>
<th>Multi-Sided Platforms &amp; Markets</th>
<th>Psychology of User Motives</th>
<th>Open Source &amp; Multi-Source</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>business/technology strategy</td>
<td>intellectual property</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>16.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>Platform selection (feature selection, advanced capabilities, etc.)</td>
<td>9</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>12.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>low barrier to entry</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>9.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>security threat</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>regulation</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>controlling the crowd</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>testing/real-time feedback</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>packaging, releasing, dev</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>shared commons</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

There were also several parameters that did not fall into one of these categories. They all ranked low and tended to be one-offs from any particular references, and so were not included in the models. Nevertheless, I've included them here:
Table 2.7 – Top factors that did not fall into one of the pillars of the core framework

The full scoring matrix can be found in Appendix B.

Chapter 3: System Models

As described in the Section 1.5 (Methodology & Approach), the variables are broken down into two groups:

- **Core model variables** – These variables represent the basic features of the foundational model like “users” as well as most off the output variables which are desirable to predict, such as “revenue.”

- **Key influencing variables** – These variables are mostly input and intermediate variables that represent the levers in our core framework that have influence over the output variables which are desirable to predict or control.

Also as described in Section 1.5 (Methodology & Approach), two core mathematical models were investigated, but ultimately the Bass-diffusion model is the one which has been modeled and extended. There are some freely available models online that were reviewed (Fiddaman,
but ultimately were deemed unusable for this research due to the fact that these models were niche cases different than the ones in this study (e.g., total valuation of a company) and consequently had different variables and relationships.

### Section 3.1: Core Model Variables

The following are the 19 variables needed to represent the core mathematical Bass-Diffusion model as well as the cost and revenue measures that we care about.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Units</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Period (months)</td>
<td>Special</td>
<td>Time in Vensim is represented as an overall system parameter and is used to integrate during simulations as its step function.</td>
</tr>
<tr>
<td>Adopters (Users)</td>
<td>Persons</td>
<td>Stock</td>
<td>This variable is our key stock for measuring the number of users on the platform</td>
</tr>
<tr>
<td>Potential Adopters (Potential Users)</td>
<td>Persons</td>
<td>Stock</td>
<td>This variable is our key stock of people who have not yet adopted the product but may yet.</td>
</tr>
<tr>
<td>Adoption Rate</td>
<td>Persons/Period</td>
<td>Flow</td>
<td>Key flow of potential users to actual users.</td>
</tr>
<tr>
<td>Total Population</td>
<td>Persons</td>
<td>Exogenous</td>
<td>Represents the total number of potential users.</td>
</tr>
<tr>
<td>Initial Population Fraction</td>
<td>Dimensionless</td>
<td>Constant</td>
<td>The initial population represented as initial value of &quot;adopters&quot;.</td>
</tr>
<tr>
<td>Adoption Fraction</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>The fraction of times a contact between an active adopter and</td>
</tr>
<tr>
<td></td>
<td>Persons/Period</td>
<td>Exogenous/Auxiliary</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Contact Rate</strong></td>
<td>Persons/Period</td>
<td>Exogenous</td>
<td>The rate at which active adopters come into contact with potential adopters.</td>
</tr>
<tr>
<td><strong>Adoption from Awareness</strong></td>
<td>Persons/Period</td>
<td>Auxiliary</td>
<td>This is how many people become adopters due to advertising or other non-word-of-mouth channels.</td>
</tr>
<tr>
<td><strong>Awareness Effectiveness</strong></td>
<td>Persons/Period</td>
<td>Exogenous</td>
<td>This is the effectiveness of advertising and other non-word-of-mouth channels of recruitment or visibility.</td>
</tr>
<tr>
<td><strong>Carrying Capacity</strong></td>
<td>Persons</td>
<td>Implicit</td>
<td>Represents the max number of people the platform will sustain. Carrying Capacity is not represented implicitly in the model, but rather is a function of development costs and the Scale-Up Loop.</td>
</tr>
<tr>
<td><strong>Cost per User</strong></td>
<td>Dollars/Person</td>
<td>Auxiliary</td>
<td>Operating costs per user.</td>
</tr>
<tr>
<td><strong>Revenue per User</strong></td>
<td>Dollars/Person</td>
<td>Exogenous</td>
<td>Generated revenue per user.</td>
</tr>
<tr>
<td><strong>Total Revenue</strong></td>
<td>Dollars/Period</td>
<td>Output</td>
<td>Total generated revenue.</td>
</tr>
<tr>
<td><strong>Support Costs</strong></td>
<td>Dollars/Period</td>
<td>Auxiliary</td>
<td>Total platform operating costs.</td>
</tr>
<tr>
<td><strong>Gross Margin</strong></td>
<td>Dollars/Period</td>
<td>Output</td>
<td>Gross margin expressed as a percentage of total revenue. Gross Margin = (Total Revenue – Support Costs) / Total Revenue</td>
</tr>
<tr>
<td><strong>Profit</strong></td>
<td>Dollars/period</td>
<td>Auxiliary</td>
<td>Total profit.</td>
</tr>
</tbody>
</table>
Section 3.2: Core Model Development

The initial model using only the core variables is shown below in figure 3.1.

Figure 3.1 – Initial model using only core model variables

There are several noteworthy features to this initial model. First, note that all of the variables from Section 3.1 are represented, if at a crude level (excepting Time, Carrying Capacity, and Initial Population Fraction which are implicitly defined through other variables or system parameters as described in the table). The key stocks are “Potential Adopters” and “Adopters”, connected by the key flow “Adoption Rate.”
Each variable relationship is shown with an arrow indicating direction, and a plus or minus symbol indicating polarity. Polarity is the description of whether one variable has a reinforcing effect or a balancing effect on another (Sterman, 2000). For instance, if one variable going up makes the other variable go up, then this relationship is represented as a plus symbol. As an example, as the number of users increases, so do the overall support costs.

There are also a number of loops represented in the model. A loop is indicated with a clockwise or counterclockwise directional arrow and an “R” for “Reinforcing” or a “B” for “Balancing,” along with a description of the loop. Each loop represents an acausal feedback chain. For instance, the “Word of Mouth” loop reflects the traditional network effects in our MSP framework and causes exponential growth of users, in turn causing exponential growth of Word of Mouth. This Word of Mouth loop would lead to infinite exponential user growth, except that it is limited by the total population. In fact, Word of Mouth becomes less effective as the number of Adopters grows, because the Adopters are increasingly talking to more Adopters and less Potential Adopters. This is represented by the two Market Saturation loops showing the ever-depleting number of Potential Users accessible to both Word of Mouth and awareness campaigns.

The two loops represented above are not actually complete loops. The loops are not closed between variables, and therefore are not true loops. Rather, these loops are placeholders to represent important phenomenon that will need to be accounted for in the final model, and are ideal candidates for linking some of the second set of variables from the key influencing factors. Here is a brief description of each of these loops:

- **Economies of Scale Loop** – This loop represents the substantial benefits of Economies of Scale that are present in many industries, including software platforms. Building a software platform that supports 200 users is no more expensive than supporting 100 users. This is a key way of growing revenue in superlinear fashion even with linear user growth. However, at certain points of scale the costs will jump in a stepwise manner, as indicated in the Scale-Up loop.

- **Scale-Up Loop** – The “Scale-Up” loop will eventually model a step function in supportability costs when supporting orders of magnitude scale increase. Costs per user
are fixed up to a point, say 100,000, and then the infrastructure has to be expanded and possibly even re-architected. This will happen again at the next "breaking point" of scale in the system; perhaps one million users, one hundred million users, and so on.

These loops are fundamental in System Dynamics, and referenced frequently throughout Sterman’s textbook on the subject. Sterman explains that scale economies are powerful drivers of reducing per unit cost (among other things) through consolidation of general, overhead, and administrative functions (Sterman, 2000, p. 369). This force is even stronger in the software platform industry, since the cost of adding more users is typically negligible until the carrying capacity is reached and the platform needs to expand the infrastructure.

As a first-order validation, the model was with initial conditions for Facebook and compared against historical data for user growth, operating costs, and revenue. Our initial results in predicting users, costs, and revenue can be seen in the figures below, and a complete analysis and case study of Facebook is presented in Section 4.2.

![Figure 3.2 - Simulated user growth from Facebook (blue line) plotted against actual historical data for user growth (red line). This trend is a surprisingly good fit for a first order model.](image)
Figure 3.3 – Simulated operating costs (red line) for Facebook plotted against historical reporting of operating costs (blue line). The shape of the simulated curve isn’t a terrible fit, but clearly needs a much slower growth pattern with a steeper inflection point. In addition, we can see a couple of stepwise changes. These both point to the need for better representation of the Scale-Up loop.

Figure 3.4 – Simulated revenue (red line) for Facebook plotted against actual historical revenue data (blue line). As we can see, it’s not a good fit at all. Our simulated line rises too early and tapers off way too early. It’s not clear from this graph that Economies of Scale were actually achieved even in the historical data, though a closer look at the actual numbers will be required to tell. It is also quite interesting that there are several spikes at what appears to be
seemingly period intervals in the actual historical data where revenue flattens out or goes down before it picks back up. Perhaps this is a clearer reflection of the "Scale Up" loop than seen in the operating costs?

All-in-all, this is not a bad first model. To first-order, user growth fits very well, operating costs sort-of well, and revenue worst of all three, though the general shape is reasonable. Given the assumptions made in initial conditions, the incomplete loops, and the lack of other confounding factors, this seems a reasonable model.

By closing the two additional scale loops described above and adding simple cost and revenue scaling factors, we can achieve an even better fit. The final model with some clean-up is shown in Figure 3.5.

Figure 3.5 – Core model with highlighted loops and simulated vs. actual data graphs.

Section 3.3: Key Influencing Variables

The 11 key influencing factors are described below, and are broken into two groups: Factors which are represented in the model in a straightforward fashion, and those which will be more complex to represent.
Section 3.3.1: Directly Represented Factors

The following influencing factors are either already represented in our model or have been represented in as a single variable. The variables here are largely controllable factors by the owners/operators/implementers of MSPs and therefore are not involved in any feedback loops. Rather, they are considered tunable model elements used to represent a real-world scenario.

As an example, the variable “Low Barrier to Entry” represents how easy or hard it is for a potential adopter to leverage the platform, which is typically a function of the cost of entry, technical difficulty of entry, or other qualifications on participants for entry (such as a valid driver’s license, in the case of Uber drivers). In most cases, these can be drastically changed as a matter of policy (e.g., lowering the cost of entry is often entirely controlled by the platform owner/operator). Even in cases where there are other constraints – for example, the regulatory constraint of a valid driver’s license in the case of Uber – it does little good from a modeling perspective to represent these at a finer level of granularity. Simply representing the barrier to entry on a 0 to 1 scale is sufficient to represent the influence of barrier to entry on the adoption fraction.

- **Crowd selectivity** – This variable is a floating point number from 0 to 1, and represents how “exclusive” this platform is, which will limit the rate of adoption and max potential adoption, but increase the quality of user-created content.
- **Low Barrier to Entry** – This variable will be tied directly to Adoption Fraction in the core model. It is just called Barrier to Entry in the final model and is a floating point number from 0 to 1 representing the difficulty of being able to join the MSP.
- **Platform Selection** – This variable is really about the attractiveness of the platform, and represents the ability to “get it right” when thinking about the platform as a product. This is a very straightforward truism, and will feed into adoption rate. Though there are many factors that have potential to influence this, such as market research, customer segments, etc. these are outside of the scope of this research. Here, it is represented simply as an exogenous variable which is called “Platform Attractiveness”.
- **Profit** – This is already covered in our core model variables.
- **Operating Costs** – This is already covered in our core model variables.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Units</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowd Selectivity</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Represents exclusivity and is a fraction from 0 to 1 which determines which percentage of the total population the platform is accessible to.</td>
</tr>
<tr>
<td>Barrier to Entry</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Feeds into adoption fraction. Adoption fraction is now an exogenous variable equal to a constant value plus the barrier to entry value.</td>
</tr>
<tr>
<td>Platform Attractiveness</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Feeds into adoption fraction.</td>
</tr>
<tr>
<td>Profit</td>
<td>Not Applicable</td>
<td>Not Applicable</td>
<td>Already covered by the core model variables (Section 3.1).</td>
</tr>
<tr>
<td>Operating Costs</td>
<td>Not Applicable</td>
<td>Not Applicable</td>
<td>Already covered by the core model variables (Section 3.1).</td>
</tr>
</tbody>
</table>

Table 3.2 – Extended model variables for factors directly representable in the model

Section 3.3.2: Complex Factors for Representation

The following factors don’t translate directly into single variables in the model, or are connected in complex ways to other model variables. For instance, the “quality” factor is about the value that individual adopters bring to the marketplace, and so a “value” variable also needed to be introduced. The “value” variable, in turn, is now a function of the number of adopters as well as the quality of the adopters, giving us a nice representation of the “content creators” to “users” ratio described by Cheung (Cheung, 2012). This “value” variable also affects the price sensitivity of the adopters and potential adopters, and so on.
- **Incentives** – Incentives here impact the number of adopters and contributors, but presumably also draw on resources to enact. We’ve broken out incentives into the top ones referenced in the literature review on this category: love, money, glory, utility. All of these influence a new variable, “Creation Urge” which indicates what fraction of adopters are going to create value in the platform beyond just using the platform. In some MSPs like Uber, these will be low or even zero, while in other platforms like Facebook where user-created content is highly encouraged, the values of these variables will be more nuanced. Another important concept is crowd diversity, which is represented here through the initial values of the incentive variables.

- **Intellectual Property** – The ability to control generated IP during usage of the platform is a key drivers for businesses. Rather than a multiplier that enhances adoption or usage, it is more often viewed as a risk management question. Therefore, this has been modeled as IP Risk, and is fractional value from zero to one which is influenced by crowd communication (see next bullet) and openness.

- **Communication Among Crowd** – Communication among the crowd has network effects which increase the quality as well as rate of adoption; however, this will also have implications on the intellectual property. For instance, co-collaborators must share IP rights or licenses, whereas in silo environments like Innocentive, Inc., it is much more straightforward to manage IP. Here it is represented as a variable called “Coopetition”.

- **Charging/Pricing Structure** – This is a big part of how different business models can be represented in the system, and is also a nice way to tie revenue to adoption rate (through barrier to entry).

- **Transparency/Openness** – Openness is the name of the variable in the model, and represents the transparency with which the platform operates. This includes transparency into business process, crowd control mechanisms, and even source code.

- **Quality** – Quality here refers to the quality of user-generated content. This is fundamentally tied to value-creation in the marketplace and operates as an attractor or detractor to potential users becoming users.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Units</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money Incentives</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Represents how effectively the platform satisfies monetarily-incentivized adopters. The incentive variables also collectively represent the crowd diversity.</td>
</tr>
<tr>
<td>Love Incentives</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Represents how effectively the platform satisfies love-incentivized adopters. The incentive variables also collectively represent the crowd diversity.</td>
</tr>
<tr>
<td>Glory Incentives</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Represents how effectively the platform satisfies glory-incentivized adopters. The incentive variables also collectively represent the crowd diversity.</td>
</tr>
<tr>
<td>Utility Incentives</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Represents how effectively the platform satisfies utility-incentivized adopters. The incentive variables also collectively represent the crowd diversity.</td>
</tr>
<tr>
<td>Creation Urge</td>
<td>Dimensionless</td>
<td>Auxiliary</td>
<td>Summed score of all Incentive variables.</td>
</tr>
<tr>
<td>Base Content Per User</td>
<td>Content/Persons</td>
<td>Exogenous</td>
<td>Expected baseline contribution ratio.</td>
</tr>
<tr>
<td>Content Creation</td>
<td>Content/Persons</td>
<td>Auxiliary</td>
<td>Expected contribution ratio of</td>
</tr>
<tr>
<td>Fraction</td>
<td>Base Creation Quality</td>
<td>Content Creation Quality</td>
<td>Content Collaboration Enablement / Coopetition⁶</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>1/Content</td>
<td>1/Content</td>
<td>Dimensionless</td>
</tr>
<tr>
<td></td>
<td>Exogenous</td>
<td>Auxiliary</td>
<td>Exogenous</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Represents how much the platform allows content creators and adopters to interact with each other.</td>
</tr>
</tbody>
</table>

⁶ Coopetition is a portmanteau of "cooperation" and "competition" which describes the paradoxical nature of competitor collaboration for a greater overall outcome (Nalebuff & Brandenburger, 1996).
Table 3.3 – Expanded set of variables needed to represent complex factors in the model

<table>
<thead>
<tr>
<th>Cost to User</th>
<th>Dollars</th>
<th>Exogenous</th>
<th>Represents the pricing structure of the platform.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>1/dollars</td>
<td>Auxiliary</td>
<td>Classic definition of benefit per unit cost is how this variable is defined. The units, 1/dollars, represents the return on investment of every dollar for platform adopters.</td>
</tr>
<tr>
<td>Price Sensitivity</td>
<td>1/dollars</td>
<td>Auxiliary</td>
<td>Represents how much a user is willing to per in fractionated dollars</td>
</tr>
<tr>
<td>Market Competitiveness</td>
<td>Dimensionless</td>
<td>Exogenous</td>
<td>Percentage from 0 to 1 representing the amount of competition in the marketplace, such as threat of new entrants.</td>
</tr>
</tbody>
</table>

Section 3.4: Extended Model Development

The first and most straightforward step was to extend the core model (Section 3.2) with the directly representable factors (Section 3.3.1). The resultant model can be seen in the figure below, with the additional variables shown in fuchsia (near the bottom right of the figure).
Figure 3.6 - Core model incorporating directly representable influencing factors. These factors – Crowd Selectivity, Barrier to Entry, and Platform Attractiveness – are the exogenous variables highlighted in fuchsia. Note that Total Population and Adoption Fraction have now had to become auxiliary variables in the model to accommodate these additions. Profit and Operating Costs did not change since they were both extension variables and core variables.

Filling in the remaining variables adds a lot of additional complexity to the model. Just representing the variables and their first-order interactions yields a largely treelike structure with exogenous variables flowing into endogenous variables\textsuperscript{7}. This can be seen in in Figure 3.7 below.

\textsuperscript{7} Exogenous variables represent things that are outside the system boundary, but have important implications. Endogenous variables are inside the system boundary (Sterman, 2000, pp. 95-96). For example, the Contact Rate of adopters with potential adopters is an exogenous variable over which there is often little control. Adoption Rate, however, is an endogenous variable that is influenced by our exogenous variable Contact Rate.
Figure 3.7 – Extended model variables representing the complex key influencing criteria discussed in Section 3.3. Incentives are on the far right, and influence the content (or services) provided by platform adopters, which in turn feeds into benefit and ultimately value, as does the quality of content (or services) provided. IP, Openness, and Collaboration Enablement are all toward the bottom of the diagram, with pricing represented as Cost to User.

These two variable sets have been modeled in different views of VenSIM, but have also been combined into an integrated view, as shown below in Figure 3.8.
Figure 3.8 – An integrated model value with the core model on the left and the extended model elements mostly on the right. Extended model variables are all shown in fuchsia, and red circles with labels are used to diagram the various elements.

The model now contains all core variables for a Bass-diffusion model with financial measures, as well as an extended set of variables representing aspects of all key influencing factors from the literature review. In addition, a few basic loops are covered and all variables are connected to the rest of the model through one or more directed inputs. Lastly, all model variables have appropriate units and equations, and the model is fully functional for simulation.

By tuning the exogenous variables, we can still see reasonable fit to the Facebook data set, as shown in Figure 3.9.
Figure 3.9 - User growth, operating cost, and revenue forecasts compared with actual historical data from Facebook, using the System Dynamics model shown in Figure 3.8.

The operating costs and revenue forecasts rise too quickly and taper off too soon. In addition, although the initial conditions of the model as set by the exogenous variables pass a basic “sanity” check, there has been no real validation yet (deeper data fit exercises are developed in Chapter 4). Still, this simple experiment provides some initial feedback that the model seems to be on the right track.

Revisiting the model structure, several of the new variables introduced to represent the key influencing factors appear incomplete. There are some clear connections which have not yet been represented between the new variables, and no new loops have been revealed.
First of all, there is no relationship between Cost Per User and Cost To User. These variables are connected via a pricing policy which dictates how much of the cost burden is passed on to the user. This might be very small for a startup with VC funding, but substantial for an established incumbent.

Also, the incentive factors actually increase the more Adopters there are. This is because there is greater opportunity for an incentivized Adopter to affect more Adopters, or have more Adopters affect them. In business nomenclature, this represents the concept of "reach," and will affect in a reinforcing fashion the number of Adopters.

Value and Market Competitiveness are also under-represented in the model. In the case of Value, it will impact Platform Attractiveness, representing a convergence toward market leaders as their market share grows. In addition, Value will affect Price Sensitivity: the more Value realized (represented in dollars returned on every dollar investment) the more people are willing to tolerate fees and fixed costs. Market Competitiveness, is something that cannot be controlled directly, and is simply a factor of timing and the market. It influences the Price Sensitivity and the Platform Attractiveness because of the increased threat of substitutes (Porter, 1979).

Revisiting all of these factors results in a more robust model as shown in Figure 3.10.
There are several notable changes in the final model. Several auxiliary variables have been added or tweaked, such as changing “Collaboration Enablement” to Coopetition and adding a Per User Cost variable, which is just a normalized version of the Cost to User variable. These changes have relatively little effect on the model and are mostly for clarity or convenience. In programming terminology, many of these changes could be thought of as “refactoring” the model.

Perhaps more interestingly, are the addition of a small set of influencing variables as well as a number of additional connections between variables. In order to represent the market forces, a “Market Competitiveness” variable was added along with a “Price Sensitivity” for the reasons discussed above; in addition, the Price Sensitivity variable was factored into the Revenue Per User to support the logic that with less Price Sensitivity, the more market offerings can inflate their prices or complement them with add-on services. Another key variable added was “Reach” to tie the number of Adopters back to user incentives for creating content/services in the marketplace. Another interesting key variable is Pricing Policy, which represents how much of the Per User Cost (operations) is passed on directly to the user through pricing. The last particularly interesting addition is the tie-in of Content Creation to the Cost Per User, under the
logic that as users accrue more content, upload more files, and leave comments for each other, there will be additional storage and processing costs (e.g., for search) which will not go away. The variable “Content Creation Fraction” was used for this rather than “Content Creation” because the fractional rate already represents a value that is normalized to a per-user value.

Finally, and perhaps most interesting in the complete model, are the addition of three new reinforcing loops:

- **Opportunity Loop** – This loop represents the opportunity for Adopters of the platform. The more users that exist on the platform, the higher the reach by introducing data, services, apps, or marketing into the platform; therefore, the stronger the incentives. When there are more users on the platform, all of our forms of incentives also go up: for monetary incentives, the opportunity to generate revenue is enhanced; for love incentives, the ability to affect more people’s lives is increased; for those motivated by glory, the opportunity for reputation increase goes up; for people with utilitarian incentives, more users means more content, which means a higher likelihood of findings data and tools to solve a problem.

- **Market Convergence Loop** – This loop anecdotally represents the “all markets converge to three” rule of business (Lampe, 2015). As the number of adopters grows on a platform, the more market share and thus market power a single company will have. The smartphone industry is a good example of this principle, where there are two or three (depending on how you count) dominant platforms (iOS, Android, Windows Mobile). The importance of this loop in the model is that simply annotates the relationships where more users leads to more content, and thus more value, which in turn increases the adoption fraction of the platform (through Platform Attractiveness). It also has the added benefit of reducing Price Sensitivity because the number of substitutes in the marketplace is reduced.

- **Cost Competitiveness Loop** – This loop shows the impact of increased operating costs on Value and therefore Adoption Fraction. Since Value is defined as Benefit divided by Cost, then when the Per User Cost increases the Value goes down. On the other hand, as more users adopt the platform, the operational costs go down due to the Economies of Scale loop, which in turn reduced the Per User Cost, allowing the organization to be more competitive.
A full export of all model variables with their units and mathematical relationships (equations) can be seen in Appendix C, and the full models themselves are downloadable from http://cdimm.net/.

Chapter 4: Simulation Experiments

Section 4.1: Using the Model

Now that the model is complete, it can be used for a variety of purposes. It can be tuned and fit to historical data in order to study a particular case. This could be because of a root cause analysis of failure, to make decisions about new platforms or adjustments to existing ones, or just to understand how a specific platform evolved over time.

In addition, certain modes of analysis can be performed to inform business choices. For example, a sensitivity analysis can be performed to understand the top parameters on which to focus. Statistical analysis can also be applied to understand which variables are correlated, and which may be simply “noise,” confounding the model.

Lastly, the model can be used to forecast growth patterns, financial measures, and other points of interest when launching a new platform into the market or when considering strategic acquisitions. In this scenario, a user would set up the model with the initial values and see how it performs. In addition, optimization algorithms can be applied, or a design of experiments (DoE) could be used for the purpose of controlling a specific outcome, whether it be maximizing revenue, minimizing cost, optimizing cost-benefit or cost-revenue ratios, or to keep risk low.

Section 4.1.1: Model Interface

In order to facilitate these uses of the model, a special view has been developed called “Model Interface.” This model interface shows the key input parameters – mostly exogenous variables that must be defined as initial conditions – as well as key output parameters that a typical user
would care about. In addition, a set of key graphs are in the view so that it is quick and easy to see the impact of tweaking a model input variable. Lastly, a set of reference variables are available for linking historical data into the graphs for comparison. This model interface can be seen in Figure 4.1, and a brief description of each column follows.

<table>
<thead>
<tr>
<th>REFERENCE DATA</th>
<th>INPUTS</th>
<th>OUTPUTS</th>
<th>GRAPHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>* users raw data</td>
<td>* Cost Per User</td>
<td>* Support Cost</td>
<td>* operating costs</td>
</tr>
<tr>
<td>* users raw data</td>
<td>* Base Div Costs</td>
<td>* Variable Costs</td>
<td>* revenue</td>
</tr>
<tr>
<td>* revenue raw data</td>
<td>* Base Revenue Per User</td>
<td>* Total Revenue</td>
<td>* User Growth</td>
</tr>
<tr>
<td>* revenue raw data</td>
<td>* Baseline Adoption</td>
<td>* Adoption Rate</td>
<td>* Profit</td>
</tr>
</tbody>
</table>

Figure 4.1 – The model interface with input variables, output variables, and graphs of key measures grouped by category (financial, adoption, policy, or other). Reference data sets are also configurable from this view on the far left of the table.

- **Reference Data** – The variables in this column represent historical data. The only purpose is to pull in historical data points to plot against the model forecasts to fit the model to historical data or simply gauge the fit. The default value in the equation box is calling a Vensim method which pulls the data from Excel. For example, "users raw data" is defined as GET_XLS_DATA('facebook', 'users', 'g', 'i3').

- **Inputs** – The variables in this column are exogenous variables that are either constraints on the model, or initial conditions reflecting your technology strategy, business strategy, and awareness campaigns. For example, in our Facebook study the default value for Barrier to Entry is minuscule at .0001, reflecting the fact that Facebook makes it incredibly easy for anyone with an email address to join.

- **Outputs** – The variables in this column are the things a model user will most likely care about, such as Adoption Rate and Total Revenue. By highlighting these variables, a user can easily view the value at every time step in either tabular or graphical fashion, as well as get statistics (e.g., mean) or identify causal relationships.
• **Graphs** – This column contains six key graphical views, which will be updated every time a simulation is run.

**Section 4.1.2: Advanced User Options**

More advanced users can pull up the full view and manipulate some of the intermediate (auxiliary) variables as they see fit, as well as even change the specific mathematical relationships between the model elements. This may be useful in cases where the “default” value is a constant that wasn’t pulled out into a separate variable. For instance, the variable Money Incentives is defined in the model as ".001 + Reach”. However, Innocentive deals in price challenges and therefore naturally has a strong monetary incentive; this could lead a user doing an Innocentive analysis to change the Money Incentives variable to “0.02 + Reach” or something even higher.

Advanced users also have the option to make the variables more nuanced. In the default model, the variables are defined as constants that can be changed along a real number or integer scale. However, this is often not representative of the real world, where these variables may be more of a function. In this case, the user has the ability to define a variable in one of the following ways:

- Constant (default)
- Step change function at specific time intervals
- Ramp function
- Exponential growth function
- Exponential decay function
- Sin wave function
- Pulse function at specific time intervals
- Pulse train function at specific time intervals
- Square wave function

One example of when this would be useful would be when taking a multi-year view of a platform that has attempted to adapt to a changing business climate. Facebook, for instance, changed
their business model in 2008, leading to a large spike in their operating costs and a smaller spike in their revenue.

Vensim also provides additional advanced features that may be useful to users. First, Vensim provides a capability known as Synthesim, which provides sliders and sparkline graphs on each variable to allow rapid tuning and experimentation with the model. Vensim also provides Monte Carlo simulation capabilities which offer the ability to run the model treating input variables as distributions and therefore resulting in distributions of output variables. Finally, Vensim provides an optimization utility which allows you to optimize the performance of one or more variables. These features are best understood by visiting the Vensim Help online (Vensim Inc., 2015).

Section 4.1.3: External Data and Model Time Steps

An important set up characteristic of the dynamic model is the special variable “Time.” Time set up is done using the Model > Settings option from the pull down menu, and allows a user to configure a variety of options for the integration. The base date at Time=0 and Time display options are configurable, as are the timescale (total period) along with initial time and final time. A user can set the interval at time steps from seconds to years, and it is critical to set the proper time increment before any model execution.

If a user is just forecasting a theoretical model, then they should use whatever time step makes the most sense. For models of this nature, months or quarters are probably the most appropriate when forecasting three or more years. Days or weeks could be useful for smaller timescales (e.g., 18 months), or for better fidelity. However, if a user is trying to fit or compare a simulation with historical data, then it is imperative they use the right timescale. For instance, in the Facebook case shown in Section 4.2, we are comparing user growth and revenue data at specific months, so we need to set the time step to months. Once specified in the model, a user needs to ensure they have specified an appropriate time interval (called TIME STEP in the options menu) over which to integration. At a high-level, this is a trade-off between fidelity and simulation speed. Higher time intervals will integrate faster but are more inclined to common errors like rounding propagation.

---

8 Note that available features may depend on which version of the software has been licensed. The available product pricing and options are available at http://vensim.com/purchase/.
Finally, once the time step interval has been specified, the external data sets need to be laid out in this format. For instance, if you have monthly data, you need to reference by a time step column in a spreadsheet which maps the progression of time to actual month to actual data. An example of this can be seen in Figure 4.2, and template spreadsheets are available at http://cdimm.net/.

<table>
<thead>
<tr>
<th>Time in timestep</th>
<th>Period in months</th>
<th>Date</th>
<th>Facebook Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4375</td>
<td>1</td>
<td>12/1/2004</td>
<td>1,000,000</td>
</tr>
<tr>
<td>1.5</td>
<td>2</td>
<td>1/1/2005</td>
<td></td>
</tr>
<tr>
<td>1.5625</td>
<td>3</td>
<td>2/1/2005</td>
<td></td>
</tr>
<tr>
<td>1.625</td>
<td>4</td>
<td>3/1/2005</td>
<td></td>
</tr>
<tr>
<td>1.6875</td>
<td>5</td>
<td>4/1/2005</td>
<td></td>
</tr>
<tr>
<td>1.75</td>
<td>6</td>
<td>5/1/2005</td>
<td></td>
</tr>
<tr>
<td>1.8125</td>
<td>7</td>
<td>6/1/2005</td>
<td></td>
</tr>
<tr>
<td>1.875</td>
<td>8</td>
<td>7/1/2005</td>
<td></td>
</tr>
<tr>
<td>1.9375</td>
<td>9</td>
<td>8/1/2005</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>9/1/2005</td>
<td></td>
</tr>
<tr>
<td>2.0625</td>
<td>11</td>
<td>10/1/2005</td>
<td></td>
</tr>
<tr>
<td>2.125</td>
<td>12</td>
<td>11/1/2005</td>
<td></td>
</tr>
<tr>
<td>2.1875</td>
<td>13</td>
<td>12/1/2005</td>
<td>6,000,000</td>
</tr>
<tr>
<td>2.25</td>
<td>14</td>
<td>1/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.3125</td>
<td>15</td>
<td>2/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.375</td>
<td>16</td>
<td>3/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.4375</td>
<td>17</td>
<td>4/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>18</td>
<td>5/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.5625</td>
<td>19</td>
<td>6/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.625</td>
<td>20</td>
<td>7/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.6875</td>
<td>21</td>
<td>8/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.75</td>
<td>22</td>
<td>9/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.8125</td>
<td>23</td>
<td>10/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.875</td>
<td>24</td>
<td>11/1/2006</td>
<td></td>
</tr>
<tr>
<td>2.9375</td>
<td>25</td>
<td>12/1/2006</td>
<td>12,000,000</td>
</tr>
<tr>
<td>3</td>
<td>26</td>
<td>1/1/2007</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.2 – A snippet of an example Excel spreadsheet mapping historical data by dates into Vensim time steps.

Section 4.2: Facebook Case Study

Facebook is an interesting case study because they are one of the most classic and most successful instances of an MSP; in addition, they have publicly available data for their finances as well as number of users over time. A historical picture of Facebook has been pieced together using various public data sources, and our simulation has been tuned to get the best fit of the simulation to this historical data. The simulation, parameters, and fitness is described below, followed by a discussion.

Section 4.2.1: Facebook Parameters & Fit
The historical Facebook data elements are:

- **Number of users** – The number of users was taken from Facebook’s own published numbers on the Company Info section of their website (Facebook, Inc., 2015). Although this data set is sparse, it is the most direct source available and goes back the longest. It is difficult to know how Facebook isn’t counting, and the assumption here is it is by user account. The number of users go back to 2004.

- **Operating costs** – The operating costs come from directly from the 10k statements as reported by Ycharts (Ycharts, Inc., 2016), and were spot-checked with the actual statements for verification (Securities & Exchange Commission, 2015). In financial reporting, expenses are broken into many categories, but in this model they are all rolled up into a single aggregate number. The operating costs in the spreadsheet are an aggregation of Cost of Goods Sold, Sales and Marketing Expense, General and Administrative Expenses, SG&A Expenses, and Total Operating Expenses. The operating costs numbers go back to 2011.

- **Revenue** – The revenue information also comes directly from the 10k statements as reported by Ycharts (Ycharts, Inc., 2016). The revenue numbers are directly from the “Revenue” line in their quarterly reports going back to 2011.

The number of users is cumulative and populated at sparse, sometimes random intervals, while the operating costs and revenue are reported quarterly. Therefore, for this model the time units were set to month and the time step to 0.0625. The data is not reported every month, but has been overlaid onto a monthly time schedule in an Excel spreadsheet for the raw data.

One challenge is that this data set does not account for abandoned accounts or for users with multiple accounts. Two separate experiments were run: one using the raw Facebook reported numbers under the assumption that abandoned accounts and users with multiple accounts cancel each other out; the other using “adjusted users” where we assume that one in every ten users has more than one account (this was done by dividing the raw number of users by a 1.1 users-per-account factor).
The input differences across experiments caused different degrees of model tuning. The input data that yielded the best results when fit was the cumulative cost numbers (not monthly) with adjusted users. The initial conditions that resulted in this best fit are shown below.

- **Timestep and Model Settings** – The timestep for integration was set monthly from 2003 out to 2015 using an Euler and Difference solver (results were nearly identical).

- **Base Cost Per User** – This was $2.46 per user.

- **Base Dev Costs** – This was set at $4 million dollars.

- **Base Revenue Per User** – This was set at $0.41 (41 cents) per user.

- **Total Population** – The initial total population was set at the total population of connected households (World Wide Web Consortium, 2016), with a 1.5x multiplier per household, and with the entire population of China subtracted.

- **Awareness Effectiveness** – This is set at 1 in 2,000, meaning that 1 in 2,000 people who are exposed to advertising results in a new member.

- **Contact Rate** – This is set at 35, which indicates that every month an adopter of Facebook comes in contact with (and talks about Facebook) 35 other non-adopters.
- **Base Adoption Fraction** – This is set at 0.00212, which means that about 1 out of every 472 people who come in contact with friends who are users are converted.

- **Adopters** – Adopters starts with the initial condition of 1.

- **Pricing Policy** – This is set at 0.2.

- **Attractiveness Factor** – This is set at 0.00026, or 13/50,000.

- **Openness** – This is set at 0.3, given that it is easy to join though Facebook is not transparent about their code or processes.

- **Coopetition** – This is set at 0.5 because interaction among participants is encouraged, although users typically stick to their circles (“friends”) through privacy controls.

- **Crowd Selectivity** – This is set at 1.0 because the platform is open to all.

- **Barrier to Entry** – This is set at 0.0001, or 1 in 10,000, since anyone can join in an automated fashion.

- **Base Content Per User** – This is set at 0.124, based off of the 1:89:10 ratio (Cheung, 2012).

- **Base Creation Quality** – This is set at 0.3.

- **Market Competitiveness** – This is set at 0.4.

The results of the simulation with these initial conditions are the best fit to user growth, operating costs, and revenue, which are shown below.
Figure 4.4 – User growth simulation versus historical data. A very nice fit.

Figure 4.5 – Historical data of operating costs versus simulation prediction of operating costs. The simulation curve is smooth, but a decent regression fit. The spikes and adjustments in are hard to account for.
Figure 4.6 – Historical data of Facebook revenue versus simulation prediction of revenue.

Figure 4.7 – Simulated profit curve of Facebook data.

Section 4.2.2: Discussion

The data fit here is very good for users, and more of a regression fit for the cost curves. The user growth patterns are remarkably accurate. The cost curves are a decent fit, but don’t
account for all of the peaks and valleys, indicative of market changes and adjustments in the Facebook business model. The revenue curve is also a nice fit, but appears to level out too soon, though it is difficult to say this with certain given that we don't know what the future holds for the Facebook market in 2016.

As mentioned previously, this data set does not account for abandoned accounts or for users with multiple accounts, and this model ended up comparing against an “adjusted user rate” to try and account for this. Adjusted user rate is a crude way of representing this dynamic element, and in reality a defection loop is also probably necessary to more accurately account for abandoned accounts.

The number of users is cumulative and populated at sparse, sometimes random intervals, while the operating costs and revenue are reported quarterly. Therefore, for this model the time units were set to month and the time step to 0.0625. The data is not reported every month, but has been overlaid onto a monthly time schedule in an Excel spreadsheet for the raw data. Experiments have been run comparing both the quarter-over-quarter numbers as well as cumulative data. The differences were largely negligible.

The model settings to result in this fit seem reasonable. The Base Cost Per User probably changed over time, but as a fixed average $2.46 per user seems like a reasonable number to account for user support and storage space. The Base Dev Cost at $4 million dollars is consistent with what platform investors suggest is appropriate for this type of utility, at least at the outset. The Base Revenue Per User, set at $0.41 (41 cents) is harder to justify, because Facebook doesn’t publish sources of revenue. They put most of the burden on advertisers and businesses, which makes it difficult to normalize into a per-user type of metric. Also, the Pricing Policy is set at 0.2 which is high, given that Facebook users typically do not pay anything.

The Awareness Effectiveness, representing response to advertising campaigns, hackathons, and other awareness efforts is set at 1 in 2,000. This rate is reasonable, but strongly debated in business circles. Additionally, the contact rate and base adoption fraction appear reasonable as well, although again the debate among proper numbers for these variables is a source of much speculation and debate. The attractiveness of the platform is set quite low, as are several of the other variables feeding into Adoption Fraction. This is a reflection of the sensitivity of the model.
to Adoption Fraction, where small changes can make a huge impact due to the traditional network effects.

The Openness factor is set at 0.3, since Facebook allows a lot of open interaction and control over individual privacy. At the same time, they are secretive with their business plans, feature rollouts, and source code. Similarly, Coopetition is set at $\frac{1}{2}$ because Facebook allows a lot of interactions while still keeping their key revenue sources isolated and allowing users individual privacy controls.

The content creation and related set of variables seem reasonable: quantity is fairly high while quality is somewhat low, since anyone can easily contribute content for any reason at any time. The market forces at 0.4 seems a bit high, and in reality probably varied a great deal over time.

The model resulted in good data fit, but there are definitely some weaknesses of the model that could be accounted for to achieve better realism. Model strengths and weaknesses are discussed more thoroughly in Section 4.7.

**Section 4.3: Twitter Case Study**

Twitter is another interesting case study because they are a less traditional but one of the most popular MSPs; in addition, they have had very struggles with their infrastructure and making money. A historical picture of Twitter has been pieced together using various public data sources, and our simulation has been tuned to get the best fit of the simulation to this historical data. The simulation, parameters, and fitness is described below, followed by a discussion. The Twitter case study was set up and executed in much the same way that the Facebook case was.

**Section 4.3.1: Twitter Parameters & Fit**

The historical Twitter data elements are:
• **Number of users** – The number of users was taken from a variety of online sources such as Statista, which holds data that was collected with Twitter’s announced number of “active users” with quarterly earnings (Statista, 2016).

• **Operating costs** – Like with the Facebook example, the operating costs come from directly from the 10k statements as reported by Ycharts and were spot-checked with the actual statements for verification. The operating costs here come directly from their “Total Operating Costs” line in their quarterly reports going back to 2012.

• **Revenue** – The revenue information also comes directly from the 10k statements as reported by Ycharts and are directly from the “Revenue” line in their quarterly reports going back to 2012.

Like the Facebook model, the time units were set to month and the time step to 0.0625. The data is not reported every month, but has been overlaid onto a monthly time schedule in an Excel spreadsheet for the raw data. Also, like the Facebook model, four different experiments were run, with adjusted users, non-adjusted users, month-over-month financials, and cumulative financials.

The input differences across experiments resulted in different levels of fit. In general, the Twitter data curves are much more “hockey stick” than “S-curve” in shape, and have a poorer fit with against the simulations. The input data that yielded the best results when fit was the cumulative cost numbers (not monthly) with adjusted users. The model resulting in the best fit was month-over-month financials with adjusted users, though it was not as clear a choice as it was with the Facebook data set. For Twitter, accounts per user were set to 1.5, since it is more common for a user to have several different accounts. The initial conditions that resulted in this best fit are shown below.

• **Timestep and Model Settings** – The timestep for integration was set monthly from 2003 out to 2015 using an Euler and Difference solver (results were nearly identical).

• **Base Cost Per User** – This was $2.55 per user.

• **Base Dev Costs** – This was set at $1.75 million.

• **Base Revenue Per User** – This was set at $0.3375 (~34 cents) per user.
- **Total Population** – The initial total population was set at the total population of connected households (World Wide Web Consortium, 2016), with a 1.5x multiplier per household, and with the entire population of China subtracted.

- **Awareness Effectiveness** – This is set at 1 in 2,000, meaning that 1 in 2,000 people who are exposed to advertising results in a new member.

- **Contact Rate** – This is set at 37.5.

- **Base Adoption Fraction** – This is set at 0.00187, which means that about 1 out of every 535 people who come in contact with friends who are users are converted.

- **Adopters** – Adopters starts with the initial condition of 1.

- **Pricing Policy** – This is set at 0.

- **Attractiveness Factor** – This is set at 0.00026, or 13/50,000.

- **Openness** – This is set at 0.15.

- **Coopetition** – This is set at 0.6375 because interaction among participants is encouraged and is most commonly public.

- **Crowd Selectivity** – This is set at 1.0 because the platform is open to all.

- **Barrier to Entry** – This is set at 0.0001, or 1 in 10,000, since anyone can join in an automated fashion.

- **Base Content Per User** – This is set at 0.124, based off of the 1:89:10 ratio (Cheung, 2012).

- **Base Creation Quality** – This is set at 0.3.

- **Market Competitiveness** – This is set at 0.2625.

The data fit of the various models is shown below.
Figure 4.8 – Operating costs comparison of various models and data fit. The month-over-month with adjusted users is the black line, and clearly the best fit.

Figure 4.9 – Revenue forecasts shown with raw data (blue). None of these are particularly good fit, although the cumulative financials with adjusted users is the best.
Figure 4.10 – User growth forecasts next to actual Twitter user growth data (blue). Our month-over-month financials with adjusted users is again the best fit. Although, the actual data is so noisy, that a fit here won’t be great and will have high variance and total mean squared error.

### Section 4.3.2: Discussion

The model in the Twitter case is much less of a clean fit than in the Facebook case. The historical data is very noisy and very sparse. This may be due to Twitter’s growth strategy, which was not always accessible to anyone with a web browser. In addition, they’ve had very radical changes in business models to try and generate revenue effectively over the years.

For user growth forecasting, an added wrinkle with this data set is that it is looking at active accounts, which does take into consideration defectors and abandoners. This may account for some of the noise in user growth, which suggests that adoption curves may not be as smooth as for typical product adoption like VCRs.

The model parameters mostly seem reasonable, with reasonable adoption variables, although it is clear that these variables should be changing over time rather than constants. Cost per user was a little higher than expected, since user-contributed content is not typically a lot of space on disk or processing power. The revenue per user was rather small, which fits the challenge Twitter has had with monetizing the platform.
Openness, Coopetition, Pricing Strategy, and Market Competitiveness were tuned to reflect realistic conditions for Twitter, but in effect had little outcome on the overall model results. The incentive characteristics were left at their defaults for this set of experiments.

One noteworthy observation is that the model which best fit the data, did not show the best Profit curve, and in fact may be the worst since it levels off earliest, as can be seen in Figure 4.11.

![Figure 4.11 - Projected profit curves from Twitter simulation experiments.](image)

**Section 4.4: Hypothetical Case Study: Manufacturing Technology Platform (MTP)**

This case study is different than the other two, because rather than trying to fit the model to historical data, the model is being set up with expected initial conditions to try assess the business viability of a new platform.

In this hypothetical example, an entrepreneur is considering introducing a new technology platform into the manufacturing sector, and wants to understand the business case. The Manufacturing Technology Platform, or MTP, has raised many questions. When can the
entrepreneur expect to start seeing profits? How much will the operating costs be, and how much capital will they have to raise to finance the venture? What can they do to optimize their costs and chances of success?

Section 4.4.1: MTP Parameters & Fit

The initial conditions for this experiment are described below, along with their rationale.

- **Timestep and Model Settings** – The timestep for integration is monthly, forecasting out through a five-year period.

- **Base Cost Per User** – The Base Cost Per User is initially set at $1.00. Although this may seem low, the entrepreneur reasons that the base costs are fixed for the first 100,000 users. In addition, the entrepreneur knows that through strategic partnerships he/she will be able to lower their operating costs.

- **Base Dev Costs** – This is set to $6 million, which is the initial funding available to kickstart the project.

- **Base Revenue Per User** – This is initially set at $2.50 on the basis of a cost calculation for paid support; however, this could be adjusted to explore different business models.

- **Total Population** – The initial total population is set at 245 million (number of U.S. adults).

- **Awareness Effectiveness** – This is set at 1 in 2,000 (.0005), meaning that 1 in 2,000 people who are exposed to advertising results in a new member. This seems reasonable, though may be more effective in this sector than normal.

- **Contact Rate** – This is set at 20; since this is a limited community representing only 5% of the U.S. population, the rate of contact is expected to be fewer than in the general population.

- **Base Adoption Fraction** – This is set at 0.002, which means that 1 out of every 500 people who come in contact with platform users are themselves converted to users. This appears reasonable, especially considering that this is without other influencing factors like platform attractiveness or barrier to entry.

- **Adopters** – Adopters starts with the initial condition of 1.
• **Pricing Policy** – This is set at 0, since the entrepreneur expects to generate revenue through sales of complementary services.

• **Attractiveness Factor** – This is set at 0.00026, or 13/50,000. This is a very low number, and one avenue for exploration is cost of increasing the attractiveness versus the return on the additional investment.

• **Openness** – This is set at 0.9, since this will be an open source project with very transparent business processes.

• **Coopetition** – This is set at 0.75 because interaction among participants is encouraged through a marketplace.

• **Crowd Selectivity** – This is set at 13.24% (.1324) since that will get us to about 18.5 million potential adopters. This is based on the number of total manufacturing jobs – 1 in 6 – in the United States (National Association of Manufacturers, 2015), which is our target user base.

• **Barrier to Entry** – This is set at 0.0001, or 1 in 10,000, since anyone can join in an automated fashion and with ease. This means about 1 in 10,000 people will have trouble trying to join.

• **Base Content Per User** – This is set at 0.124, based off of the 1:89:10 ratio (Cheung, 2012).

• **Base Creation Quality** – This is set at 0.5, as quality is emphasized and incentivized in the platform.

• **Market Competitiveness** – This is set at 0.6 because, with the Internet of Things renaissance, there are many new market entrants as well as numerous incumbents looking to expand their portfolio.

The initial model run with these parameters shows a drastically accelerating user adoption, costs, and revenue, as shown in Figure 4.12.
Looking at these results, it appears to be an unrealistic business scenario. The user adoption curve isn’t outlandish – The model forecasts around 11,000 users by the end of the first year, and around 1 million users by year 3. This is ambitious, but achievable. Tweaking the model also allows the entrepreneur to look at a less ambitious case – a lower bound on expected growth rate.

Far more concerning here is the explosive operating costs and the negative profit. The initial guess for Cost Per User may have seemed reasonable, but will clearly not be sustainable. This is where the model can be tuned to constraints, in this case investor funding, to see what the per capitated cost needs to be in order to stay within the budget.

Lastly, it appears that the Revenue Per User may be overestimated – a simple thing to correct.
Making these various adjustments shows a much more reasonable set of results, as shown in Figure 4.13.

Operating costs in the low millions is a more reasonable forecast; not pennies, but sustainable. Revenue still appears to be inflated, which is why Profit rises so sharply. Still, at least now the shape of these curves are where they need to be to warrant a venture. User growth appears more linear, but this is likely just because it is at the very early edge of the adoption curve before the inflection point.

In this way, the model can be used to identify constraints, optimize outcomes, and understand uncertainties of a platform in the marketplace.
Section 4.4.2: Discussion

In the hypothetical scenario, the model is used to attempt to forecast performance and tune real-world business decisions to meet constraints and optimize outcomes. The likelihood that all of the input variables would be static constants is quite low, and in a real-world scenario our entrepreneur would want to tweak these as step functions, ramp functions, or other forms of variable parameters.

The model doesn’t have a single or simple way to reflect different business strategies for comparison, and would need to be reset for each case in the comparison. For instance, charging all users an up-front cost versus a freemium model versus relying on advertising for revenue are three very different business approaches. The Vensim model could be tuned to represent each of these business scenarios, thought it would be laborious to do so and may even require adjustments to the model itself and not just the interface parameters. In many cases, a hybrid model would be desirable, and not only is this even more difficult to represent, but nearly impossible to optimize against if using different models for comparisons and simulations.

It is also unclear how well the model performs in a fresh-stage platform with a small user base, such as the one in the hypothetical scenario. In both the Twitter and Facebook case studies, the platforms have had more than a decade to mature. In addition, the user base is the connected global population, rather a narrow subpopulation. The result is that the numbers for Twitter and Facebook is that they are dealing with multiple order of magnitudes more users and financials than our hypothetical scenario.

Nevertheless, the hypothetical scenario does enable one to explore the relationships between cost, revenue, adopters, and control factors to a first order. In the final simulation run for the hypothetical scenario, the numbers are reasonable and none of the output variables appear to contradict one another. Even if they are incorrect by 50% or more margin of error, they are still useful in assessing the overall magnitude, optimal control factors, and general shape of the platform progression over time.

Section 4.5: Sensitivity Analysis
The model, although it has its weaknesses, reflects the prioritized parameters from the literature review along with auxiliary financial, market, and user adoption variables. In addition, it functions for simulations and has been explored through a number of case studies. One final aspect in this research is to understand which of these parameters are most important to focus on understanding, and getting right in a real-world use case.

A great tool for this is sensitivity analysis, a form of uncertainty modeling (de Neufville & Scholtes, 2011). Fortunately, the Vensim tool suite allows for running Monte Carlo simulations, where an input variable can be represented as a distribution rather than a static constant. The model can be exploited in a number of useful ways using this type of technique, though the scope of this research is limited to understanding to what degree a change in an input variable impacts the output variables.

It is important to note that these sensitivities are contingent upon the setup of other model variables, and this type of multi-variable sensitivity study is beyond the scope of this research. Therefore, the input variables and their impact on output variables have been explored one at a time and using our Facebook base case. In order to make this context clear, two of our Facebook output data sets are shown embedded with the sensitivity graphs. In addition, the focus is on Operating Costs, Revenue, User Growth, and Profit – sensitivity impact on other variables is beyond the scope of this research.

All Monte Carlo runs executed 1,000 simulations each, and all were done using a random uniform (normal) distribution with latin hypercube as the sampling technique. For succinctness, only the interesting results are shown. For instance, if there is no sensitivity between an input and an output variable, this is mentioned but the graph is not shown.

**Section 4.5.1: Sensitivity of Financial Factors**

Three financial variables were tested for sensitivity: Cost Per User, Revenue Per User, and Base Dev Costs. The Cost Per User was varied from $0.10 to $10.00. Not surprisingly, the Support Costs are highly sensitive to this variable, as shown in Figure 4.14. Its influence on profit was similar, though did not impact revenue nearly at all. In addition, the Cost Per User influenced user adoption, but only at higher amounts, as seen in Figure 4.15.
Base Dev Costs were varied from $500,000 to $30 million, to see their impact on the outcomes. These costs are ultimately amortized over the life of the project, although come back into play in the scale-up loop. The Base Dev Costs had no influence on the Adopters or Revenue, but did have small sensitivity and influence on the Support Costs and therefore Profit, as shown in Figures 4.16 and 4.17.

Lastly, Base Revenue Per User was varied from $0 to $20.00. This had no output on Support Costs or Adopters, but Revenue and therefore Profit were highly sensitive to this variable, as can be seen in Figure 4.18 and 4.19.

Figure 4.14 – Sensitivity graph of Support Costs given variability of Cost Per User. Changes in Cost Per User result in very different results of Support Costs.
Figure 4.15 – Sensitivity graph of Adopters given variability of Cost Per User. Interesting to note that this entire window is contained in the 95th percentile of Cost Per User variations.

Figure 4.16 – Sensitivity graph of Support Costs given variability of Base Dev Costs. The sensitivity is a fairly narrow band, but clearly affected by moderate changes in the Base Dev Costs.
Figure 4.17 – Sensitivity graph of Profit given variability of Base Dev Costs.

Figure 4.18 – Sensitivity graph of Revenue given variability of Revenue Per User. The Revenue is highly sensitive, even at small changes, and gets less sensitive at larger scales.
Section 4.5.2: Sensitivity of Adoption Factors

Awareness Effectiveness, Contact Rate and Base Adoption Fraction were all explored using sensitivity analysis. The Awareness Effectiveness was varied from 0 to 0.01, meaning that advertise has zero effect to influencing 1 out of every 100 people to join. The sensitivity of this variable on Support Costs is minimal, but has a huge impact on Adopters, Revenue, and Profit. The sensitivity graph for Profit is shown in Figure 4.21 – Adopters and Revenue graphs look very similar but have not been shown here.

Contact Rate was varied from 5 to 150, and may be the most sensitive element of the entire model. Its influence on Adopters, Profit, and Revenue is huge, and it even has more influence on Support Costs than most of the other variables. Base Adoption Fraction yields less variability, but still substantial, and was varied from 0 to 0.5, under the rationale that 50% conversion rate of uninitiated is the best that any platform could hope for. Selected sensitivity graphs for these two variables can be seen in Figures 4.22 through 4.24.
Figure 4.20 – Sensitivity graph showing impact of Awareness Effectivity on Support Costs.

Figure 4.21 – Sensitivity graph showing impact of Awareness Effectivity on Profit.
Figure 4.22 – Sensitivity graph showing influence of Contact Rate on Support Costs.

Figure 4.23 – Sensitivity graph showing influence of Contact Rate on Profit. Sensitivity graph of Contact Rate for Adoption and Revenue look similar. Small changes to this variable in all the bottom three bands (50th percentile, 75th percentile, 95th percentile) can result in huge differences.
Figure 4.24 – Sensitivity graph showing impact of Base Adoption Fraction on Profit. It is interesting to note all the sensitivity is influencing early and in narrow bands. Beyond a particular number of Adopters, this variable doesn’t have much impact, which is why we see our two simulation runs (red line and navy line) largely unaffected.

Section 4.5.3: Sensitivity of Policy Factors

Pricing Policy, Attractiveness Factor, Openness, Coopetition, and Barrier to Entry were all tested using sensitivity analysis. None of these variables except for Barrier to Entry (varied from 0 to 1) had more than negligible impact on Adopters or Support Cost, and the Attractiveness Factor (varied from 0 to 0.5) didn’t have an impact on any of the variables, possibly indicated a weakness of the model.

Pricing Policy, Openness, and Coopetition were all varied from 0 to 1, and all had moderate impact on Revenue and Profit. The sensitivity of Coopetition was the greatest amongst this group of variables, and it is notable that Barrier to Entry also has a big influence at larger values and at later time stages. Selected graphs demonstrating these findings can be seen in Figures 4.25 through 4.29.
Figure 4.25 – Sensitivity graph showing influence of Pricing Policy on Total Revenue.

Figure 4.26 – Sensitivity graph showing influence of Openness on Profit.
Figure 4.27 – Sensitivity graph showing influence of Coopetition on Support Costs.

Figure 4.28 – Sensitivity graph showing influence of Coopetition on Total Revenue.
The key influencers of “Value” in the model are the quantity and quality of user-contributed content. Therefore, Base Content Per User and Base Creation Quality were both studied with a sensitivity analysis, as was the Market Competitiveness. All three of these variables are fractionated, representing a percentage, and so were all varied from 0 to 1.

The Base Content Per User had a small impact on Support Costs, and a large impact on Revenue and Profit (Figures 4.30 and 4.31) and Base Creation Quality had only a moderate influence on Revenue and Profit (Figure 4.32). Market Competitiveness was the most interesting variable in this group, influencing all four variables, primarily through its influence on the Adoption Fraction.

This different impacts of Market Competitiveness are interesting because they aid in understanding the relationship between Adopters, Support Costs, and Revenue. Costs are typically a constraint which cannot be reduced by much. In a highly competitive market margins
typically get smaller due to fixed costs, higher price sensitivity by consumers and lower revenue. These effects can be seen quite nicely in Figures 4.34 through 4.36.

**Figure 4.30** – Sensitivity graph showing influence of Base Content Per User on Support Costs.

**Figure 4.31** – Sensitivity graph showing influence of Base Content Per User on Total Revenue.
Figure 4.32 – Sensitivity graph showing influence of Base Creation Quality on Profit.

Figure 4.33 – Sensitivity graph showing influence of Market Competitiveness on Adoption. It is interesting to note that it gets less sensitive at higher values.
Figure 4.34 – Sensitivity graph showing influence of Market Competitiveness on Support Costs.

Figure 4.35 – Sensitivity graph showing influence of Market Competitiveness on Profit.
Section 4.6: Model Strengths and Weaknesses

Overall, the model is very well-conceived intellectually. A persistent challenge with modeling has always been: “what to model?” This manifests itself most commonly in three ways: defining the system boundary; determining which system elements should be represented; and at what fidelity they should be represented.

The rigorous approach taken to identify the key factors through the core framework and literature review has ensured that the key system elements are represented. Likewise, using the Bass-diffusion model as a foundation ensures that fundamental growth factors are accounted for properly. The system boundary is clearly identified and encompasses the model elements well. The fidelity of the model elements varies, depending on the nature of the phenomenon in the model, and while adequate, this is an area for model refinement.

The model makes sense intuitively, with strong variable relationships and loops that are grounded in the research and literature. The different variable groups and loops have been annotated and color-coded, and all arrows have polarity indicators. Multiple views have been developed making use of shadow variables for a clean display.
All of the units resolve appropriately when the model elements are combined, which also lends validity to the model. For example, Total Revenue is in Dollars, and is mathematically defined as Adopters times Revenue Per User. Since Adopters is in Persons and Revenue Per User is in Dollars Per Person, multiplying the two cancels out Persons only leaving Dollars, which are the appropriate units for Revenue.

Furthermore, the model functions properly. Simulations can be run, variables can be tuned, it can be fit to data, and sensitivity analysis can be performed. In addition, a special view has been developed to serve as a straightforward interface into the model for performing all of the aforementioned operations. Initial experimentation of the model and historical data fit with Twitter and Facebook also demonstrate that the model functions properly; that is, it simulates the real world with promising verisimilitude. The shapes of the curves are generally what is expected, the influence of most of the input characteristics appear proportional with real-world behavior, and the data is able to be fit to historical use cases with values that are believable.

The dynamics of MSPs and especially the indirect network effects are central to the thesis of this research. The model does not, however, provide direct observation of the indirect network effects, nor does it allow for easy experimentation with the number of sides of an MSP. Instead, the model takes into account indirect network effects through other factors such as Coopetition, Platform Attractiveness, and Barrier to Entry. This is sufficient to account for indirect network effects, but not to observe how different sides of the marketplace interact with each other directly. In order to experiment with how many sides to bring to an MSP, for instance, one would need to run multiple experiments with different configurations of these input variables to represent the different options. Representing different user groups with individual cost burdens, adoption rates, and revenue would greatly enhance the models ability to assess indirect network effects, but would add a good deal of complexity as well.

The mathematics used to relate variables are mostly intuitive and represented as simple algebraic expressions. There are two big exceptions, however. The Economies of Scale loop uses a base 10 logarithm to assess the scale of operations, and uses less intuitive mathematics in general in the variables surrounding this loop. The mathematics of this loop could benefit from further scrutiny.
The other exception is the Adoption Fraction. The model is incredibly sensitive to this variable, and consequently, all of the variables that feed into it. This means a user ends up with very small numbers, (such as 0.00026 for Platform Attractiveness in the Facebook case study) which, while appearing to function adequately in the model, can leave one scratching their head about what the number actually reflects in the real world.

The model would benefit from expansion on financial and market variables. The expenses, revenue, and market factors are generally represented pretty well, but at a very low fidelity. Support Costs and Revenue are single real parameters, as is Market Competitiveness, but these factors are difficult to reason about without breaking them down further. For instance, a quick look at a balance sheet will often show five or more lines for different types of costs. In addition, there are many more subtle market influences that are unaccounted for, such as regulation, supplier power, and threat of new entrants.

Several of the model variables are represented as static constants. Even if they were changed to represent functions, there are several additional connections and feedback loops that could enhance the precision overall. For instance, market competitiveness is a dynamic force that is changing, in part due to the market convergence and number of adopters. Another example is the Total Population: in the case of Facebook this is expanding as the world becomes more connected, and in the case of our hypothetical MTP scenario this marketplace is expanding as we face a renaissance of manufacturing through IoT, grassroots makers, and government investment in the sector.

Finally, the model has not withstood a truly rigorous validation. The results of our experiments are promising, but without more detailed data comparisons, and without testing against smaller scale and more heterogeneous platforms, it is difficult to trust the answer beyond a wide margin of error. The financial data was taken from 10k statements, which does not provide a strict way of rolling up and reporting costs and revenue. As a result, it is difficult to compare these numbers across companies, and often even within the same company from year to year as their reporting changes. For better confidence in the financial forecasts, the model should be explored using raw numbers from a company’s internal books.

Nevertheless, the forecasts fit reasonably well, and with input numbers that are believable. The model appears quite useful for gauging orders of magnitude of platform adoption, support costs,
and revenue, and their general progression over time. For understanding growth trends and timing, it appears to function well.

Chapter 5: Results

Chapter 5.1: Summary of Experiment Results

Several model experiments were conducted, involving two historical case studies and one hypothetical case study. The two historical case studies were Facebook and Twitter, and in each case the model had the input variables arranged to fit historical growth and financial data and then assessed for veracity of the observed values. Each case was explored for data fitness across four different expressions of the data, summarized in the table below.

<table>
<thead>
<tr>
<th>USERS</th>
<th>FINANCIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported users</td>
<td>Monthly cost and revenue</td>
</tr>
<tr>
<td>Reported users</td>
<td>Cumulative cost and revenue</td>
</tr>
<tr>
<td>Adjusted users</td>
<td>Monthly cost and revenue</td>
</tr>
<tr>
<td>Adjusted users</td>
<td>Cumulative cost and revenue</td>
</tr>
</tbody>
</table>

Table 5.1 – Design of experiments for historical case studies

Adjusted users was explored as a way to account for users with multiple account, since “reported users” is really “reported user accounts.” In addition, for both raw users and adjusted users, different experiments were conducted to assess fitness of data when looking at cumulative costs and revenue over time, or just monthly growth. Using the model interface, the various inputs were adjusted to get the best fit for Adoption, Support Costs, and Revenue. The scenario which produced the best data fit was analyzed and documented. For Facebook, this was Adjusted users and Cumulative financials, and for Twitter it was adjusted users with monthly financials.

In the Facebook case, the data fit was quite good, especially for Adopters, which showed a fairly steady growth pattern. The financial data for Facebook was unevenly distributed and so the actual simulation fit more like a regression curve, finding the average slope between the actual data points. For the Twitter data, the fit was not as good. This was in part because there was not enough data, and in part because the data set was very noisy, with high variance at
unevenly reported intervals. Operating costs and revenue simulations fit the historical data reasonably well, but the forecasted trend line was not able to inflect quickly enough to match the user growth well.

The Facebook parameters resulting in the best fit mostly appear very realistic, especially for Cost Per User, Revenue Per User, Awareness Effectiveness, Contact Rate, Adoption Fraction, Barrier to Entry, and Content Creation Per User. The ones that seemed questionable were Pricing Policy, Attractiveness Factor, and Market Competitiveness.

The Twitter parameters resulting in the best fit appeared mostly realistic as well. The Content Creation variables seemed questionable, and the Attractiveness Factor seemed low, but all other parameters seemed reasonable.

The model was used to explore a hypothetical scenario as well. The hypothetical scenario was a new platform in the manufacturing sector, therefore more niche and with its own unique characteristics. For this set of experiments, the model was initialized to expected parameters and then simulated in an attempt to forecast growth and profit. These parameters resulted in a reasonable forecast of user growth and profit loss over time, but operating costs skyrocketed. After iterating over various adjustments, the model suggests that operating costs need to be held to $1.20 per user, and generating at least twenty cents per user in revenue. User growth in this model was more linear than the historical examples and forecasted to be at about 12,000 in the first year, which is ambitious but reasonable. It also forecasted that Profit would go from negative to positive around month 11, which seems unrealistic.

The last experiment was to perform a sensitivity analysis on the key input variables. The model was shown to be most sensitive by far to Contact Rate, and the all results of the sensitivity analysis are shown in Table 5.2 grouped into sensitivity tiers.

<table>
<thead>
<tr>
<th>SENSITIVITY</th>
<th>VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Contact Rate, Base Adoption Fraction</td>
</tr>
<tr>
<td>Moderate</td>
<td>Coopetition, Content Per User, Awareness Effectivity, Revenue Per User, Cost Per User, Market Competitiveness, Base Creation Quality</td>
</tr>
</tbody>
</table>

97
Chapter 5.2: Key Findings & Discussion

This research indicates that direct network effects are still the biggest drivers of adoption, which in turn drives costs and revenue, value creation, and marketplace power. Indirect network effects are murkier, and the research indicates that their impact on the dynamics of platform adoption are largely through influencing the Adoption Fraction and Contact Rate. This research supports the hypothesis that these factors are worth careful consideration, and that they can be predicted and, to some extent, controlled for improving outcomes.

The key factors culled from the literature review provide practical insight into the most influential factors, and can be used to guide top-level decision-making. By translating the key factors from the literature review into both a conceptual and runnable model (see Figure 3.10), these relationships revealed themselves in an intellectually intuitive way, and can serve as a roadmap for understanding the complex dynamics of an MSP in the marketplace.

The results of the simulation experiments lend validity to the model and indicate that these dynamics are predictable and, to some extent, controllable. This finding indicates that it is worthwhile to perform some analysis in order to make the business case and inform strategic decisions. Some of the inputs can be treated as constraints (e.g., Market Competitiveness), while others need to be thought of as decision points (e.g., Incentives). In effect, being aware of what you can change and deciding how to change it is as important as understanding what you can't change and how to formulate a strategy around those constraints.

An important discovery in this work is that the indirect network effect factors often manifest themselves as coefficients to traditional network effect variables. As an example, “incentives” was the top factor from the literature review, but its real impact is to make the platform more attractive through content creation and reach. This “Opportunity Loop” is a reinforcing power that ultimately amplifies the percentage of people who will adopt the platform through word of mouth.
The sensitivity analysis revealed that Adoption Fraction and Contact Rate are the most important variables; however, these are not things that can be directly controlled. Therefore, the key elements to focus on for understanding, forecasting, or optimizing an MSP in the wild are those which are sensitive but controllable: Coopetition, Content Per User, Awareness Effectivity, Revenue Per User, and Cost Per User, Market Competitiveness, and Base Creation Quality.

**Chapter 5.3: Implications for Business Leaders**

The key factors identified in the literature review, and their conceptualization in a connected model should be used as a reference guide of how to approach risk analysis and strategic decision-making for MSPs. Their impacts should be explored when assessing strategic questions for MSPs, quantitatively if possible.

Qualitatively, the core framework is a nice conceptual model to understand why these factors are important, and since they were mapped into the three pillars of the core framework (business strategy, technology strategy, net positive awareness) it is straightforward to see which part of the organization should own the analysis and/or decision. For example, the revenue model should be ultimately determined as a business strategy function involving the CEO and finance, while pricing policy would be owned by the Chief Marketing Officer and finance.

To the extent possible, the project should allocate appropriate resources for a project analysis to do more quantitative explorations. This will help strengthen the business case and inform decision-making.

Market research can lend insight into the how competitive the marketplace is and whether the platform is an early or late entrant. Our model suggests that a competitive marketplace will reduce the Adoption Fraction through word of mouth, but this can be offset by investing more resources into awareness campaigns, from traditional advertising to guerilla marketing to prize challenges. The impact of doing this will have larger cumulative effects when done earlier.

This research suggests that Adoption Fraction and Contact Rate via word-of-mouth are the most important things to affect, and every effort should be made to do so. Things like
ambassador programs, badges, and referral rewards should be employed to amplify the effect on these variables.

In addition, this research suggests that value creation by user-contributed content is a fundamental driver for increasing the total value and therefore adoption of a platform. To capitalize on this effect, it should be made easy for users to contribute content and to discover and reuse shared content. Also, extra resources can be allocated to incentivize users to create content, and the earlier this is done the more overall impact it will have. Balancing the quality of contributions and the quantity will be a strategic decision depending on target market. For instance, the hypothetical case was focused on the professional engineering services in the manufacturing sector, and an emphasis on quality is appropriate for this somewhat niche community, as opposed to the Facebook and Twitter cases where success is defined by volume of content rather than accuracy, meaningfulness, usefulness, or interestingness.

Where pricing is concerned, it is helpful when forecasting costs, revenue, and profit to think about costs on a per-user basis as a bottom-line financial metric. In some cases, where every user pays a monthly fee, this is a straightforward undertaking. In other cases the cost burden will be unevenly distributed. For instance, the costs may be borne out by advertisers while the service is free to the users.

When assessing these types of tradeoffs, business leaders should remember this fundamental tenet: enhance adoption incentives and limit adoption inhibitors. The side of the marketplace that stands to benefit the most from interactions and transactions with other sides of the marketplace is most likely to tolerate fees.

Chapter 5.4: Implications for Future Research

This research was largely concerned with business dynamics, but was not large enough in scope to truly shed insight into the cycles of creative destruction on a macroscale or in the market at large.

User Incentives was the most critical factor revealed from the literature review, and is an entire category of research by itself. Its inclusion in this model, while conceptually sound, was limited.
A deeper exploration into the motivational forces of people and their system impacts would add a lot of value to the body of existing knowledge on this topic.

Market forces were not fully explored, with Market Competitiveness being the only variable. Inclusion of Porter forces (Porter, 1979) or similar framework would be useful in understanding how much control one can really exert over the successful adoption of a product or platform. The “Market Convergence Loop” in the model does represent some of the changing market dynamics as a product or platform becomes more popular, but explicitly representing threat of substitutes, buyer power, supplier power, etc. would help understand the cycle of creative destruction on a macro-scale.

In addition, further exploration to understand the “tipping point” effect would be very useful to assess risk. At what point, does the market begin to converge? Many people say 100,000 users is the magic number, but do the numbers support this conclusion? How many adopters are needed to really enhance supplier power?

As discussed in Section 4.6, the model developed and used in this research has areas for improvement:

- **Accounting for change** – How do companies react to changes in the marketplace? Treating variables as step functions, ramp functions, distributions, etc. would provide a much more realistic and accurate picture.

- **Financial veracity** – What are the real financial numbers to be concerned with? Support costs and revenue numbers taken from public financial statements are not trustworthy sources of “ground truth” financial performance.

- **Mathematical validity** – What are the mathematics behind these complex interactions? Actual proofs and better approximation functions would yield increased accuracy, precision, and correctness of this type of simulation.

**Appendix A: Bibliographic References**


Creative Commons Corporation. (2015). Creative Commons Attribution 4.0 International (CC BY 4.0). Retrieved from Creative Commons: https://creativecommons.org/licenses/by/4.0/


Howe, J. (TBD). *Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business.* TBD: TBD.


Lampe, J. (2015, 8 14). Vice President of New Business Solutions, UI LABS. (J. A. Barkley, Interviewer)


Stallman, R. (2015, May). Open Source Hardware Marketplace Email Interview. (J. A. Barkley, & Z. Li, Interviewers)


## Appendix B: Full Scoring Matrix

<table>
<thead>
<tr>
<th>Type</th>
<th>Parameter</th>
<th>Crowdsourcing &amp; Innovation</th>
<th>Multi-Sided Platforms &amp; Markets</th>
<th>Psychology of User Motives</th>
<th>Open Source &amp; Multi-Source</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>business strategy</td>
<td>incentives</td>
<td>14</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>24.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>crowd selection (diversity, specialization, population size, etc.)</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>19.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>intellectual property</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>16.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>communication among crowd</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>16.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>charging /pricing model</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>15.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>transparency/openness</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>13.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>Platform selection (feature selection, advanced capabilities, etc.)</td>
<td>9</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>12.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>profit</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>10.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>low barrier to entry</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>9.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>quality</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>9.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>operating costs</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>8.0</td>
</tr>
<tr>
<td>strategy</td>
<td>labor pool (private, community, public, hybrid) internal vs. external</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>7.0</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----</td>
</tr>
<tr>
<td>business strategy</td>
<td>socially responsible</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>5.0</td>
</tr>
<tr>
<td>business/technology strategy</td>
<td>security threat</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>task size</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>controlling the crowd</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>task difficulty / complexity</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>legal matters and licensing</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>add capacity (grow workforce)</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>task type</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>growth rate</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>business strategy</td>
<td>geographic focus</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>industry focus</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3.0</td>
</tr>
<tr>
<td>spillover</td>
<td></td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>marketplace</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3.0</td>
</tr>
<tr>
<td>awareness</td>
<td>diverse range of clients</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2.0</td>
</tr>
<tr>
<td>awareness</td>
<td>market leader</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>testing/real-time feedback</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2.0</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----</td>
</tr>
<tr>
<td>business strategy</td>
<td>company reputation</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2.0</td>
</tr>
<tr>
<td>awareness</td>
<td>having contacts in key markets</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>packaging, releasing, dev</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>contracting</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>government involvement</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>awareness</td>
<td>events sponsoring</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>awareness/business strategy</td>
<td>political infrastructure</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>awareness/business strategy</td>
<td>social infrastructure</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>shared commons</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>feedback</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>recognition</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>energy drain</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>cannibalization</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>margin erosion</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>utilitarian motives</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>business strategy</td>
<td>hedonic motives</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>technology strategy</td>
<td>multi-tenancy</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>strategy</td>
<td>real-time data from mfg.</td>
<td>provide all four XaaS layers</td>
<td>intelligent search</td>
<td>quoting engine</td>
<td>carrying capacity</td>
<td>baseline number of clients per professional</td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>technology strategy</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>BPO</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>coordination</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>core competence</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>good internal processes</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>idea fixation</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>implementation/governance</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>increase worker</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>satisfaction and loyalty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>innovation model</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>leading work management</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>capabilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of submissions</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>production blocking</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>review format</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>understanding of context</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Vensim Export with Units and Mathematical Definitions

Adopters= INTEG (  
    Adoption Rate,  
    1)  
Units: Persons

Adoption Fraction=  
    Base Adoption Fraction - (Market Competitiveness/1000) - Barrier to Entry  
+ Platform Attractiveness  
Units: Dimensionless  
The fraction of times a contact between an active adopter and a  
potential adopter results in adoption.

Adoption from Awareness=  
    Awareness Effectiveness*Potential Adopters  
Units: Persons/Month  
Adoption can result from advertising according to the  
effectiveness of the advertising effort with the pool of  
potential adopters.

Adoption from Word of Mouth=  
    Contact Rate*Adoption Fraction*Potential Adopters*Adopters/Total Population  
Units: Persons/Month  
Adoption by word of mouth is driven by the contact rate between  
potential adopters and active adopters and the fraction of times  
these interactions will result in adoption. The word of mouth  
effect is small if the number of active adopters relative to the  
total population size is small.
Adoption Rate =
Adoption from Awareness + Adoption from Word of Mouth
Units: Persons/Month
The rate at which a potential adopter becomes an active adopter.
This is driven by advertising efforts and the word of mouth effect.

Attractiveness Factor =
0.00026
Units: dollars [0,0.001,0.0001]

Awareness Effectiveness =
0.0005
Units: 1/Month
Advertising results in adoption according the effectiveness of the advertising.

Barrier to Entry =
0.0001
Units: Dimensionless

Base Adoption Fraction =
0.00212
Units: **undefined** [0,0.01]

Base Content Per User =
0.124
Units: Content/Persons
Note, this initial value comes from the 1:89:10 ratio found in the literature review. Here we are lumping in new content and modified existing content together, so we get 11/100 = 0.124.

Base Cost Per User =
114
2.46
Units: dollars/Persons [0.5, 0.1]

Base Creation Quality =
0.3
Units: 1/Content

Base Dev Costs =
4e+006
Units: dollars

Base Revenue Per User =
0.41
Units: **undefined**

Benefit =
(Content * Content Creation Quality) - ((Content * Content Creation Quality) * IP Risk)
Units: Dimensionless

Contact Rate =
35
Units: 1/Month
The rate at which active adopters come into contact with potential adopters.

Content =
Adopters * Content Creation Fraction
Units: Content

Content Creation Fraction =
Base Content Per User * Creation Urge * Coopetition
Units: Content/Persoms
Content Creation Quality =
   Base Creation Quality * (1 + (1 - Crowd Selectivity))
Units: 1/Content

Coopetition =
   0.5
Units: Dimensionless

Cost Per User =
   (Base Cost Per User^2 + Content Creation Fraction) ^ Scale of Operations
Units: dollars/Persons

Cost to User =
   Cost Per User * (1 + Pricing Policy)
Units: dollars/Persons

Creation Urge =
   Glory Incentives + Love Incentives + Money Incentives + Utility Incentives
Units: Dimensionless

Crowd Selectivity =
   1
Units: Dimensionless

Development Costs =
   Base Dev Costs + (Base Dev Costs * Scale of Operations)
Units: dollars

Glory Incentives =
   116
$0.001 + \text{Reach}$

Units: Dimensionless

Gross Margin =

Profit / Total Revenue

Units: Dimensionless

IP Risk =

IF THEN ELSE( (Coopetition - Openness) > 0, (Coopetition - Openness), 0 )

Units: Dimensionless

Love Incentives =

$0.001 + \text{Reach}$

Units: Dimensionless

Market Competitiveness =

0.4

Units: Dimensionless

Money Incentives =

$0.001 + \text{Reach}$

Units: Dimensionless

Openness =

0.3

Units: Dimensionless

Per User Cost =

Cost to User * User

Units: dollars

Platform Attractiveness =

117
IF THEN ELSE (0.01 - (1/(Attractiveness Factor*Value)))<0, 0, 0.01 - (1/(Attractiveness Factor*Value))

Units: Dimensionless

0.2 - 1/(Attractiveness Factor*Value)

Potential Adopters= INTEG (Adoption Rate, Total Population - Adopters)

Units: Persons

The number of potential adopters is determined by the total population size and the current number of active adopters.

Price Sensitivity=

Value * (1+Market Competitiveness)

Units: 1/dollars

Pricing Policy=

0.2

Units: Dimensionless

Profit=

Total Revenue - Support Costs

Units: dollars

raw total population::INTERPOLATE::=

1.50025e+009

Units: Persons

GET XLS DATA("?facebook','users','1','d2")

Reach=

IF THEN ELSE((Adopters * Reach Factor)>0.1, 0.1, Adopters * Reach Factor)

Units: Dimensionless
Reach Factor =
\[ \frac{1}{(\text{Total Population} \times 10)} \]
Units: 1/Persons

Revenue Per User =
\[ \text{Base Revenue Per User} + (\text{Base Revenue Per User} \times \text{Price Sensitivity} \times 10) \]
Units: dollars/Persons

Scale Factor =
\[ 1 \]
Units: 1/Persons

Scale of Operations =
\[ \log_{10}(\text{Adopters} \times \text{Scale Factor}) \]
Units: Dimensionless

Support Costs =
\[ \text{Development Costs} + (\text{Cost Per User} \times \text{Adopters/1e+007}) \]
Units: dollars

Total Population: INTERPOLATE :=
\[ \text{Crowd Selectivity} \times \text{raw total population} \]
Units: Persons
The size of the total population.

Total Revenue =
\[ \text{Adopters} \times \text{Revenue Per User} \]
Units: dollars

User =
\[ 1 \]
Units: Persons

119
Utility Incentives = 0.001 + Reach
Units: Dimensionless

Value = Benefit / Per User Cost
Units: 1/dollars