Merging Prospect Theory with the Analytic Hierarchy Process: Applications to Technology Markets

by

Alexander Elkholy

A Thesis Presented in Partial Fulfillment
of the Requirement for the Degree
Master of Arts

Approved November 2014 by the
Graduate Supervisory Committee:

Dieter Armbruster, Chair
Karl Kempf
Hongmin Li

ARIZONA STATE UNIVERSITY

December 2014

ABSTRACT

This thesis presents a model for the buying behavior of consumers in a technology market. In this model, a potential consumer is not perfectly rational, but exhibits bounded rationality following the axioms of prospect theory: reference dependence, diminishing returns and loss sensitivity. To evaluate the products on different criteria, the analytic hierarchy process is used, which allows for relative comparisons. The analytic hierarchy process proposes that when making a choice between several alternatives, one should measure the products by comparing them relative to each other. This allows the user to put numbers to subjective criteria. Additionally, evidence suggests that a consumer will often consider not only their own evaluation of a product, but also the choices of other consumers. Thus, the model in this paper applies prospect theory to products with multiple attributes using word of mouth as a criteria in the evaluation.

DEDICATION

May Hall, thank you for being my grandmother.

Victoria, thank you for everything.

ACKNOWLEDGMENTS

I would like to say thank you to my committee for all of the help I received over the years.

Prof. Hongmin Li, I have always found your feedback helpful. Your assistance has been insightful and exactly relevant to my work.

Prof. Dieter Armbruster, I would like to thank you for everything you have done for me in my career as a student. You taught me how to think critically about my work and be more than just a student.

Karl Kempf, you have taught me more about being a working researcher than any class I ever took. I am constantly amazed and humbled by the discussions we have and could not have asked for a better role model.

Finally, I would like to thank Intel Corporation for their support.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

Manufacturers, firms and retail outlets have to make decisions on pricing, promotion and inventory, even if they are not selling directly to end-users. Because of this the literature is filled with economic models intended to help make better decisions. Most of these models, whether they are trying to optimize the amount of inventory to keep on hand or decide the best time to release a new product, in general rely on knowing or predicting an appropriate demand function. Often these models will assume the demand comes from a well known process (e.g. Bass (1969) diffusion) or make simplifying assumptions about the demand to do their analysis. One problem with this approach is that it does not allow for the possibility of social dynamics, which can greatly influence behavior of a potential customer. This is especially relevant in an age where the internet has transformed how people look for and communicate about new products. Social networks have been shown to have a great influence on subjects such as biology, neuroscience, computer science and economics, as Watts (1999) demonstrates (Watts and Strogatz, 1998, p. 440). In fact, word of mouth and network structure have been shown by Abrahamson and Rosenkopf (1997) and Aral and Walker (2011) to have significant effects on buying behavior.

In addition, Kiesling *et al.* (2012) suggests that while some aggregate models, such as Bass diffusion, can account for word of mouth effects, they do not allow for any heterogeneity in consumers or allow for one to ask what-if type questions. Having these features is especially useful for simulation studies and is particularly suited for agent based modeling. This thesis aims to model the decision process for an individual, which can be used as a framework for an agent based simulation.

Expected utility theory for a long time has been the reigning model on rational choice. However it has drawn criticism in that rational behavior of an individual is the exception, rather than the rule. The assumption of a perfectly rational agent has been criticized by some, including Kahneman and Tversky (1979), who proposed a generalization to expected utility theory called prospect theory. In the years since it has been considered one of the best models for decision making under risk (Barberis, 2012), uncertain outcomes (Tversky and Kahneman, 1992) and riskless choice (Tversky and Kahneman, 1991). The riskless choice model is especially useful as it provides a framework for extending prospect theory to multiple attributes. This problem was studied by Bleichrodt *et al.* (2009), Zank (2001) and Bleichrodt and Miyamoto (2003).

Applying prospect theory to products with more than one criterion raises the question of how the attributes of products are evaluated, and how to quantify the relative strengths and weaknesses. Stevens (1946) describes the ratio scale of measurement in which objects are compared as ratios, which serves as a basis for the analytic hierarchy process (AHP). Originally developed in the late 1970's, AHP is an approach to decision making in which individual preferences are converted into weights by measuring the preferences as ratios (Forman and Gass, 2001, p. 4). The process of converting to a ratio scale is also used to rank the products relative to each other with respect to different criteria (Saaty, 1990). Bernasconi *et al.* (2010) give empirical evidence to suggest that the method of ratio scaling mimics the cognitive process involved in decision making. They go on to suggest that the approach is applicable to prospect theory.

Prospect theory and AHP both provide a good framework for understanding the decision process, but neither take into account word of mouth effects. Because it is well known that communication between potential customers can greatly influence decisions, and the social network structure greatly influences the speed and completeness of the diffusion (Abrahamson and Rosenkopf, 1997, p. 290), word of mouth is an integral component to this model.

This thesis describes a decision model which merges these three concepts, prospect theory, AHP and networks. Prospect theory is used to capture the decisions made by people who are not perfectly rational, while AHP is used to provide a method for measuring the strength of different choices, serving as an input to the functions of prospect theory. Word of mouth effects are then used as an input which reflects the observation that success or failure of a new product depends heavily on who and how many people spread knowledge of it.

Prospect theory, AHP and networks are discussed in detail in chapters 2,3 and 4, respectively. Chapter 5 synthesizes these ideas in a single model and the influence of word of mouth and reference dependence are demonstrated through examples. Chapter 6 discusses the possibility that a technology market, as a whole, follows prospect theory in deciding the success of a product.

Chapter 2

PROSPECT THEORY

Prospect theory is widely accepted as one of the best explanations for violations of expected utility theory. Consider a gamble with 50% chance to win \$110, and a 50% chance to lose \$100. Expected utility predicts that the decision maker would evaluate $0.5 \times \$110 + 0.5 \times \$(-100) = \$5$ as the utility, making this gamble the rational choice to maximize utility. But in fact most people declined to take this gamble (Barberis, 2012, p. 4), violating expected utility. To demonstrate this and other violations, Kahneman and Tversky (1979) collected data of the following form by asking volunteers to choose between two prospects, *prospect* A: Win $\$X$ with a probability p or *prospect* B: Win $\$Y$ with probability q or Z with probability r. These can be represented as (X, p) and $(Y, q; Z, r)$, respectively. The decision maker is then assumed to evaluate each prospect and choose the one with the higher utility. Kahneman and Tversky show several cases where the volunteers were systematically irrational.

Endowment Effect and Status Quo Bias

Tversky and Kahneman (1991) describe an experiment in which participants were given a decorated mug and asked how much they would be willing to sell it for. Some of the participants were not given anything and were told they had the option of receiving a sum of money or a mug. They were asked their preferences on what amount of money would make them indifferent between the two options.

The price the mug owners were willing to sell for and the price the other participants would buy it for were different, with the mug owners choosing a higher price. This discrepancy between the prices can be interpreted as the sellers endowing the mug with more value simply by virtue of ownership.

Both the sellers and the participants who did not own a mug face the same problem but with different references, his is demonstrated in figure 2.1, where the mug owners are at state x, owning a mug, and the rest are at state r, choosing between x and y (cash). This status quo bias means the mug owners saw selling as a loss, where the buyers saw either option as a gain.

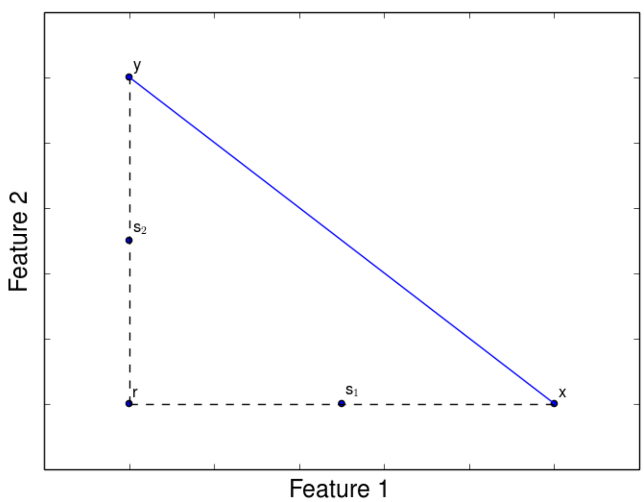


Figure 2.1: Different Reference Points Give Different Utilities.

Improvements versus Tradeoffs

In figure 2.1, both options x and y are assumed to be the same utility regardless of the reference point. This is true when one is at r because either is seen as a gain. However, when s_1 is used, the choice of x is a gain in feature 1 combined with a loss in feature 2, while a choice of y is the opposite, with a gain in feature 2 and a loss in feature 1. This is reversed when evaluating from s_2.

Kahneman and Tversky hypothesize that when a decision maker is choosing between x and y from the position of s_1, x will be preferred over y and vice versa for s_2. Indeed, evidence was found to support this. In a second study done with volunteers, one group was given a coupon for a free dinner (s_1), and the second group was given a free professional photo portrait (s_2). The subjects were then asked if they wanted to exchange their gift for either two free dinners (x) or several more professional portraits (y).

Almost nobody kept their original gift, with most at s_1 choosing the second dinner, and most with s_2 choosing the portraits, confirming their hypothesis.

Value Function

From this evidence a new formulation for evaluating prospects was proposed where the utility of a good is expressed with a value function $v(\cdot)$. The value function $v(\cdot)$ captures the irrational behavior described by specifying a nonlinear equation which represents the three axioms of prospect theory: (i) gains and losses are defined relative to a reference point; (ii) losses are perceived as larger than equivalent in magnitude gains; (iii) the sensitivity to marginal increases in gains or losses is diminished with larger magnitudes. These properties imply that the value function is concave/convex when above/below the reference point, respectively, and that it is steeper for losses then gains. One possible function is shown in figure 2.2 (Kahneman and Tversky, 1979, p. 281).

In their subsequent paper Tversky and Kahneman (1992) extended prospect theory to uncertain outcomes. Data was used to estimate the parameters of the value

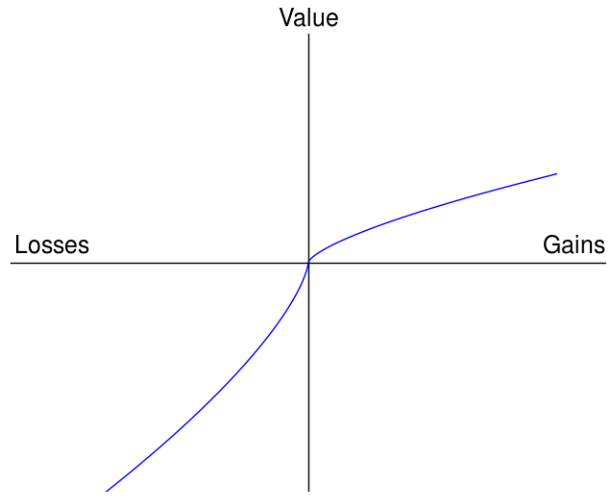


Figure 2.2: Kahneman and Tversky's Example of a Value Function.

function. Since then there have been numerous papers and functional forms consistent with the original axioms proposed (Stott, 2006, pp. 105-106). Tversky and Kahneman (1991) dealt with choices for which the outcomes are known completely without risk, accounting for reference dependence, loss aversion and diminishing returns. The riskless choice model is used as a basis for this thesis, using the value function proposed by Tversky and Kahneman (1991).

While prospect theory is an important model for explaining how individuals are not perfectly rational, there are difficulties in knowing how to apply it. This is because there are not many well known applications of prospect theory (Barberis, 2012, p. 2). Kahneman and Tversky also did not put forth any explanation for what the reference should be.

One of the problems addressed by this thesis is the difficulty of measuring the strength/weaknesses of choices. This is especially true when choices are judged on entirely subjective criteria (e.g. color, style). Additionally, decisions are made without accounting for social effects, which is considered in the coming chapters.

Chapter 3

ANALYTIC HIERARCHY PROCESS

AHP is a methodology used in making complex choices. Originally formulated in the 1970's by Thomas Saaty (Saaty, 1977), the intention of AHP was to create a practical and methodical system for making decisions (Forman and Gass, 2001, p. 4). Problems are formulated such that for a decision, the user selects from multiple alternatives which are judged with respect to several criteria. An example is choosing what computer to buy, where some of the criteria they are evaluated on are portability, speed and cost. The defining feature is that the alternatives are judged not on absolute measurements, but instead are judged relative to each other. This allows one to effectively give rankings on criteria which are difficult to assign a numeric value to.

AHP suggests that when approaching a problem one should break it down into a hierarchy. At the highest level of the hierarchy is the goal, or decision which one is trying to make. At the bottom level are the alternatives one needs to make a decision on. Between the two are the criteria with which the alternatives are judged, and each level represents a cut of the problem representing different factors which go into the problem.

Once the problem is organized as such, each level of the hierarchy is worked through from the top down. The criteria are ranked and assigned weights with respect to their relative importance to the problem. The alternatives are then ranked on how well they meet each criteria. Comparing criteria or alternatives is done pairwise on a scale from 1 to 9. Two alternatives a and b, for example, are compared by saying a is four times better than b. In doing so objects are compared on a ratio scale, as opposed to assigning a numeric value (Stevens, 1946, p. 679). This way of comparing is what allows AHP to rate things which are not easy to assign a number to, such as style. In addition, it ensures that all measurements are on the same scale (Saaty, 1990, p. 10).

Comparison Matrix

Saaty (1990) describes how comparisons are made by creating a matrix. Suppose one is given n objects to compare, $O_1, ..., O_n$, and the goal is to rank the relative strengths, $s_1, ..., s_n \in \mathbb{R}$. A matrix O is created containing the pairwise ratios of the strengths

$$O = \begin{pmatrix} s_1/s_1 & s_1/s_2 & \cdots & s_1/s_n \\ s_2/s_1 & s_2/s_2 & \cdots & s_2/s_n \\ \vdots & \vdots & & \vdots \\ s_n/s_1 & s_n/s_2 & \cdots & s_n/s_n \end{pmatrix}. \tag{3.1}$$

The i, jth entry of O is the relative advantage of O_i over O_j. One would say that O_i is x times better than O_j and hence the i, j entry in the matrix would be x. By using the ratios, one is essentially using the objects as units in measuring each other.

If this matrix, $O = (o_{ij})$, is consistent (i.e. $o_{jk} = o_{ik}/o_{ij}$ for $i, j, k = 1, ...,\text{dim}(O)$), then the largest eigenvalue, λ_{max}, is equal to the dimension of the comparison matrix, $\lambda_{\text{max}} = n$, and the vector $(s_1, s_2, ..., s_n)^t$ is an eigenvector

$$\begin{pmatrix} s_1/s_1 & s_1/s_2 & \cdots & s_1/s_n \\ s_2/s_1 & s_2/s_2 & \cdots & s_2/s_n \\ \vdots & \vdots & & \vdots \\ s_n/s_1 & s_n/s_2 & \cdots & s_n/s_n \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} = n \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}. \tag{3.2}$$

The normalized eigenvector, which is unique, associated with λ_{max} is called the priority vector. If the comparison matrices are perfectly consistent then calculating the scores is simple. However in practice the values of s_i/s_j will not necessarily be accurate.

As long as the matrix is reciprocal ($o_{ji} = 1/o_{ij}$), then the principal eigenvalue is at least as large as the dimension, i.e. $\lambda_{\max} \geq n$ (Saaty, 1990, p. 13). Reciprocity is a much easier condition to enforce and ensures that the principle eigenvector is stable to small perturbations.

AHP follows a three step process to arrive at a decision using the comparison matrix, preference elicitation, comparison of alternatives and determination of rank. This can be visualized as a hierarchy with the decision problem at the top, the alternatives $A_1, ..., A_n$ on the bottom, and the criteria $C_1, ..., C_m$ with which to judge the alternatives in the center

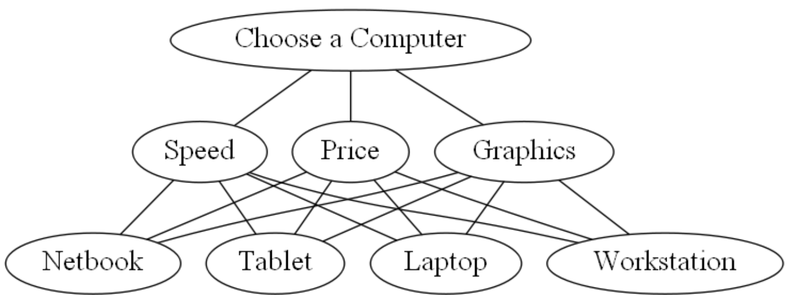


Figure 3.1: Building a Hierarchy for Choosing a Computer.

Preference Elicitation

Once the problem has been formulated as a hierarchy the first step is find the relative importance of the criteria to the user. This is done without considering the alternatives. For laptops (figure 3.1) the criteria could be price, hard drive space, speed, style, etc, which are put into the comparison matrix to derive the set of weights w_k. These represent the relative importance of the criteria with respect to the decision. In this context equation (3.1) is used to compare the criteria to find the values of w_i/w_j (the objects O_i are replaced with the criteria C_i). The end result of this process gives the equation (3.2). This is used to solve for the vector consisting of the weights of each criterion,

$$\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix}, \tag{3.3}$$

which is the priority vector associated with $\lambda_{\max}$.i

Comparison of Alternatives

The next step is to evaluate the alternatives by comparing them pairwise on how strong they are in each criterion. Using equation (3.1), m comparison matrices of size $n \times n$ are completed, where n is the number of alternatives and m is the number of criteria. The kth matrix compares the alternatives with respect to the kth criterion. The goal is to find the set $(a_{k1}, ..., a_{kn})$ for each of the criteria, which represent the rankings of the alternatives with respect to the kth criterion. Thus for each of the criteria, the priority vector

$$\begin{pmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kn} \end{pmatrix} \tag{3.4}$$

is calculated.

Determining Rank

Let M be an $n \times m$ matrix such that $M_{ij} = (a_{ij})$. By multiplying M with the weight vector $\vec{w}$ from equation (3.3), a new vector of size $n \times 1$ is created, where the ith entry in the vector is the final ranking of the ith alternative

$$M \cdot \vec{w} = \begin{pmatrix} \sum_{i=1}^{m} w_i a_{1i} \\ \sum_{i=1}^{m} w_i a_{2i} \\ \vdots \\ \sum_{i=1}^{m} w_i a_{ni} \end{pmatrix}. \tag{3.5}$$

Each entry is the weighted sum of the strengths of the alternatives in each criterion. The final vector gives the rankings of the products where a higher ranking is a more suitable alternative.

This chapter has demonstrated the process used by AHP. Using the comparison matrix, the strengths of the alternatives are measured relative to each other. Subjective judgements which may not have any intuitive method of measurement can now be ranked by using the ratio scaling method. This also ensures that all judgements are scaled in the same way, so as to allow for meaningful comparisons. But while AHP solves the problem of measuring the strengths of alternatives, it is still necessary to account for how word of mouth can influence a decision.

Chapter 4

DIFFUSION AND NETWORKS

Diffusion

Bass (1969) originally proposed a highly successful model for the diffusion of new products and ideas. The likelihood of purchase at time T with no purchases made yet is given by (see figure 4.1)

$$\frac{f(T)}{1-F(T)} = p + qF(T),$$

where $F(T)$ is the total number of adopters at time T, $f(T)$ is the likelihood to purchase at time T and p, q represent the coefficients of innovation and imitation, respectively. It is commonly assumed that the coefficient of innovation represents how many adopt independently or through advertising, while the coefficient of imitation represents the bandwagon effect, where an increasing number of adopters creates a positive feedback which in turn creates a larger pressure to adopt. Rogers (2010) describes the innovation coefficient as starting the process of diffusion with the early adopters being the innovators. Once there are enough adopters, word of mouth effects (imitation) dominate the process and almost all potential adopters will decide to adopt.

While diffusion has seen widespread success, the process of diffusion through word of mouth is aggregated and does not allow one to understand what underlying dynamics are at work. Bass diffusion originally described a simple diffusion process, even when there could be many factors which influence the rate, or even completeness of the diffusion. In the Bass model, and others like it, all potential adopters are

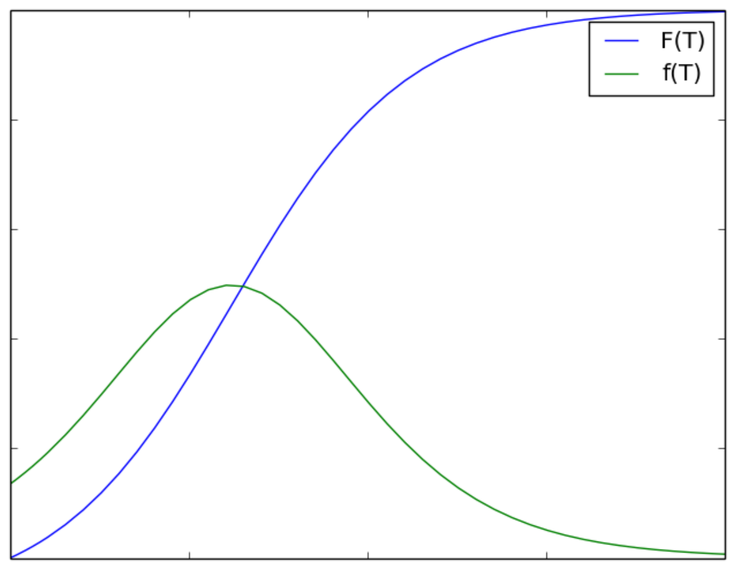


Figure 4.1: Plot of F and f, where F is the Cumulative Number of Adopters, and f is the Number of New Adopters Each Time Period.

assumed to be homogeneous, and it is not specified how communication influences their decisions (Delre *et al.*, 2007, p. 187). It has also been criticized for not being suitable for answering what-if type questions (Kiesling *et al.*, 2012, p. 184).

Bandwagons and lock-in

Arthur (1989) proposed a model with a market of two competing technologies. One of these technologies can become locked-in, creating a situation where it becomes dominant, essentially forcing the other out of the market. This framework has been applied to studying VHS vs. Betamax video formats (Katz and Shapiro, 1986, p. 822)(Liebowitz and Margolis, 1995, p. 205), QWERTY vs. Dvorak keyboards (David, 1985, p. 332) (Liebowitz and Margolis, 1995, p. 213) and AC vs. DC electricity (David and Bunn, 1988, p. 165), for example.

Abrahamson and Rosenkopf (1993) describe bandwagons as a pressure which is caused by the increasing number of adopters. Even if the product is not the best available, individuals and organizations will jump on the bandwagon simply because they see so many others doing so and wish to keep up. Katz and Shapiro (1985) assume that with more adopters the usefulness and profitability of the product increases. Telephones or fax machines are good examples of products whose utility increases with more adopters. Bikhchandani *et al.* (1992) presents a model describing fads/fashions with bandwagons as an informational cascade, where the quality of a new product is uncertain to the individual making a decision; the decision maker will receive a signal about the product as being either good or bad with equal probability. Each individual also observes the decision of all those who decided before. Thus over time the pressure to conform comes from the collective decision of the decision makers, which can lead to either a profitable or unprofitable product to be locked in.

Abrahamson and Rosenkopf (1997), stresses the need for individuals to communicate in a social network. The authors argue that small differences in the structure can affect the speed and completeness of the diffusion process. Watts (1999) showed that networks are applicable to a wide variety of fields. In particular many social networks have been found to conform to a small world structure, where most individuals do not have a large number of connections but can be connected by only traveling a few links. Janssen and Jager (2003) further showed the effects of social networks on the adoption of new products based on word of mouth effects. In their model the utility of a product is given by

$$\text{Utility} = \beta S + (1 - \beta)x, \tag{4.1}$$

where $\beta \in [0, 1]$ is a parameter weight for the sum of the social component and the strength of the product. The product's strength is represented by $S = 1 -$

$|d - p|$, which is the difference between the personal preference, p, and the dimension d (strength of product). The social component, x, is the fraction of friends the individual has who have adopted the product.

Having reviewed the applications of networks in studying diffusion, fads and word of mouth effects, it will now be applied to the decision making process in the next chapter.

Chapter 5

DECISION MAKING PROCESS

Following AHP, comparison matrices are generated to calculate weights and rank products. The matrix M is formed again where each column vector is the priority vector for the rankings of products with respect to the criteria. As before each row in M is associated with a product and each column entry is the score of an alternative with respect to the kth criteria.

In choosing between n products, assume there are m criteria. These are properties of a product that the decision maker will consider when evaluating their utility, such as safety when buying a car, battery life for a laptop or price. Each of the criterion can be either positive or negative. I.e. the utility of a product is increased the stronger a product is in a positive criterion, and decreased when it is stronger in a negative criterion. An example of a negative criterion is price, as higher price is less desirable. Associated with each of the criteria is a weight w_k as before, which are calculated using the comparison matrix.

The value function $v(\cdot)$ from Kahneman and Tversky (1979) will also be used. To fulfill the three axioms, the function must be concave above the reference point ($v'' < 0$ for $x > 0$) and convex below ($v'' > 0$ for $x < 0$) (diminishing returns and reference dependence) and be steeper for positive arguments than negative arguments ($v(x) < v(-x)$ if $x > 0$). This model will use the value function from Tversky and Kahneman (1992),

$$v_k^+(x) = \begin{cases} \alpha(x-r)^\gamma & \text{if } x \geq r \\ \beta(-(x-r))^\delta & \text{if } x < r \end{cases}, \tag{5.1}$$

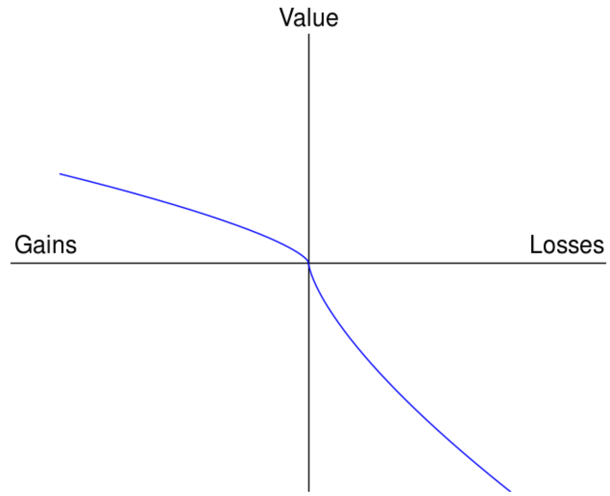


Figure 5.1: Value Function for Negative Criteria.

for its simplicity and because it has been shown to fit experimental data well (Stott, 2006, pp. 118-119). Note that the $\alpha, \beta, \gamma, \delta$ parameters and reference r are dependent on the criteria. This is the most general case and the actual values are calibrated from data. For negative criteria, because higher scores should give a decrease in utility, the function $v_k^-(x) = v_k^+(-x)$ is defined, which is a reflection about the y-axis.

The zone of insensitivity captures the effect that up to a threshold, small changes in an criterion (such as price) about a reference does not affect the overall utility. Several authors (Kalwani and Yim (1992), Bridges *et al.* (1995), Raman and Bass (2002)) have found evidence to support the existence of such a region. This effect can be included in the value function by denoting the minimum threshold needed to cause a positive change in utility as Z_h and the minimum needed to cause a negative change as Z_l. Assuming both constants are positive, the positive value function is

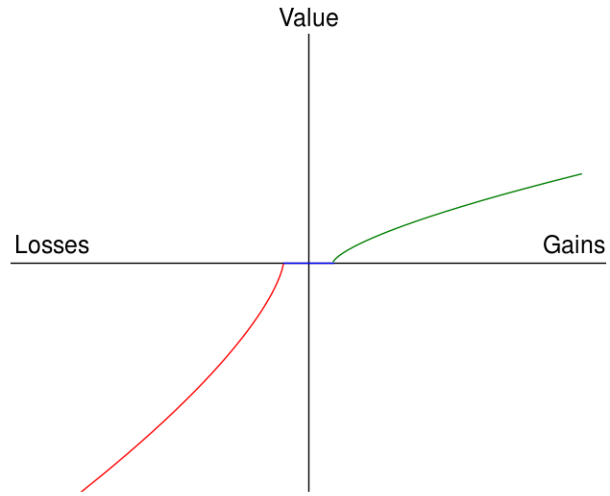


Figure 5.2: Modified Positive Value Function Showing Insensitivity.

modified as such

$$v_k^+(x) = \begin{cases} \alpha(x - r - Z_h)^\gamma & \text{if } x \geq r \\ \beta(-(x - r + Z_l))^\delta & \text{if } x < r \end{cases} . \tag{5.2}$$

This will have the effect of creating an interval $(r - Z_l, r + Z_h)$ about the reference where the utility does not change.

Word of Mouth

To incorporate word of mouth, assume the decision maker is in a social network, meaning they have some number of connections to other agents, their friends. When faced with a choice, said decision maker will seek feedback from their friends about the product, which can either be positive or negative. Positive and negative word of mouth are naturally positive and negative criteria, respectively. A vector is formed for positive word of mouth,

$$\vec{p} = \begin{pmatrix} a \\ f_1 \\ \vdots \\ f_n \end{pmatrix}. \tag{5.3}$$

Here $f_1, ..., f_n$ represent the (normalized) number of friends of the agent who are saying positive things about alternatives 1 to n, respectively, and a is the reference value. Write the negative word of mouth similarly as

$$\vec{n} = \begin{pmatrix} 0 \\ g_1 \\ \vdots \\ g_n \end{pmatrix}, \tag{5.4}$$

where the reference is zero for negative word of mouth is always zero.

Integration

One of the challenges of prospect theory is that it does not specify what a reference should be. Two common references, for example, would be a product one already owns and what one expects to own in the future; each reference can lead to significantly different choice outcomes. For this theory a generic reference product comparable to the products in the decision is assumed. The alternatives and the reference will be evaluated with respect to each criteria (except for word of mouth, which uses $\vec{p}$ and $\vec{n}$) which is used to create each matrix R_k, where $(R_k)_{ij}$ represents the relative advantage of alternative i over alternative j using AHP. Note the first row and column $(R_k)_{1,j}$

and $(R_k)_{i,1}$ correspond to the reference product in this instance,

$$R_k = \begin{pmatrix} u_{r_k}/u_{r_k} & \cdots & u_{r_k}/u_{n_k} \\ u_{1_k}/u_{r_k} & \cdots & u_{1_k}/u_{n_k} \\ \vdots & & \vdots \\ u_{n_k}/u_{r_k} & \cdots & u_{n_k}/u_{n_k} \end{pmatrix}. \tag{5.5}$$

Solving for the principal right eigenvector gives the rankings of the products for this criteria $(u_{r_k}, u_{1_k}, \cdots, u_{n_k})^t$, where u_{r_k} is the ranking for the reference product. Iterating through each criterion and completing this process for each of the m criteria will give the values for the relative strength of each product with respect to each of the criteria. From this the matrix M is formed and functions v^+ and v^- are applied as appropriate. At this point the matrix is different from AHP in two fundamental ways, first the row for reference product was removed and used to parameterize the value functions. Second, the word of mouth criteria is the sum of the positive and negative word of mouth. This means that enough negative feedback on a product will cancel the positive feedback. However negative feedback is weighted more than the equivalent amount of positive feedback because the value function weighs losses more than gains,

$$M = \begin{pmatrix} v_1^{+/-}(u_{1_1}) & \cdots & v_n^{+/-}(u_{1_m}) & \\ \vdots & \vdots & \vdots & v^+_{\text{social}}(\vec{p}) + v^-_{\text{social}}(\vec{n}) \\ v_1^{+/-}(u_{n_1}) & \cdots & v_n^{+/-}(u_{n_m}) & \end{pmatrix}, \tag{5.6}$$

where each $v_k^{+/-}(\cdot)$ the value u_{r_k} is used as the reference

$$v_k^+(x) = \begin{cases} \alpha\sqrt{x - u_{r_k} - Z_h} & \text{if } x \geq u_{r_k} + Z_h \\ \beta\sqrt{-(x - u_{r_k} + Z_l)} & \text{if } x \leq u_{r_k} - Z_l \\ 0 & \text{else} \end{cases}. \tag{5.7}$$

Multiplying this matrix and the weight vector $\left(w_1, ..., w_m, \frac{w_s}{2}\right)^t$, where w_s is the weight for word of mouth, gives the overall rankings of

$$\begin{pmatrix} \sum_{i=1}^{m} w_i v_i^{+/-}(u_{1_i}) \\ \vdots \\ \sum_{i=1}^{m} w_i v_i^{+/-}(u_{n_i}) \end{pmatrix} + \frac{w_s}{2}\left(v_{\text{social}}^{+}(\vec{p}) + v_{\text{social}}^{-}(\vec{n})\right). \tag{5.8}$$

The weight $\frac{w_s}{2}$ is used to avoid weighing the word of mouth too much relative to the other criteria. The sum in the jth row is the score for the jth product. This contains the ranking of the products as in AHP, relative to the reference. Also note that the value function applied to the vectors $\vec{p}$ and $\vec{n}$ is applied element-wise to the entries of said vectors.

Example

To see how reference dependence and word of mouth have an effect on the decision, consider three alternatives, computer A, computer B and computer C. Each will be compared on the criteria speed and graphics. To do this, the weights for the criteria must be calculated. Here the graphics are considered three times as important as speed is, and word of mouth twice as important as speed. This leads to a hypothetical comparison matrix

$$\begin{array}{l} \\ \text{Speed} \\ \text{Graphics} \\ \text{Word of Mouth} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Speed} & \text{Graphics} & \text{Word of Mouth} \end{array} \\ \begin{pmatrix} 1 & 1/3 & 1/2 \\ 3 & 1 & 1/2 \\ 2 & 2 & 1 \end{pmatrix} \end{array}. \tag{5.9}$$

This matrix is slightly inconsistent, which makes the largest eigenvalue 3.1357 just over the expected eigenvalue of 3, which a consistent 3×3 matrix would give. Calculating the priority vector by finding the eigenvector and normalizing gives the weights as: $\vec{w} = (0.1677, 0.3487, 0.4836)^t$.

For each criterion the products are compared, with the exception of word of mouth, a comparison matrix is generated where each entry is the relative strength of one product over another. As before, the following are hypothetical matrices which are slightly inconsistent

$$\text{Speed} \quad \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 1 & 1/5 & 3 \\ 5 & 1 & 2 \\ 1/3 & 1/2 & 1 \end{pmatrix} \end{array} \tag{5.10}$$

$$\text{Graphics} \quad \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 1 & 1/3 & 1/4 \\ 3 & 1 & 1/4 \\ 4 & 4 & 1 \end{pmatrix} \end{array}. \tag{5.11}$$

Because the decision maker is also considering the decisions of his/her friends, two more vectors are created, $\vec{p}$ for positive word of mouth, and $\vec{n}$ for negative word of mouth

$$\vec{p} = \begin{pmatrix} 20 \\ 70 \\ 60 \end{pmatrix} \qquad \vec{n} = \begin{pmatrix} 30 \\ 20 \\ 20 \end{pmatrix}. \tag{5.12}$$

Calculating the priority vectors from (5.10) and (5.11) along with the (normalized) word of mouth vectors, a new matrix is formed where each column is one of the calculated vectors

$$\begin{array}{c} A \\ B \\ C \end{array}\begin{pmatrix} 0.2377 & 0.1130 & 0.1333 & 0.4286 \\ 0.6072 & 0.2351 & 0.4667 & 0.2857 \\ 0.1551 & 0.6519 & 0.4000 & 0.2857 \end{pmatrix}. \tag{5.13}$$

Using only AHP with word of mouth, the ranking of the products is calculated by multiplying this matrix by the weight vector $\vec{w}$ from (5.9)

$$\begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \end{array}\begin{pmatrix} 0.2377 & 0.1130 & 0.1333 & 0.4286 \\ 0.6072 & 0.2351 & 0.4667 & 0.2857 \\ 0.1551 & 0.6519 & 0.4000 & 0.2857 \end{pmatrix}\begin{pmatrix} 0.1677 \\ 0.3487 \\ 0.2418 \\ 0.2418 \end{pmatrix} = \begin{pmatrix} 0.2151 \\ 0.3657 \\ 0.4191 \end{pmatrix}. \tag{5.14}$$

According to this calculation, product C is the best choice. However, if the decision maker already owns one of these products, say product A, and uses it as a reference, then prospect theory can be applied. Note that the negative word of mouth vector, $\vec{n}$, now has to change to $\vec{n} = (0, 20, 20)^t$, because product A is now the reference, and for simplicity no reference will be used for negative word of mouth.

$$\begin{array}{c} \text{B} \\ \text{C} \end{array}\begin{pmatrix} v^+_{\text{Speed}}(0.6072) & v^+_{\text{Graphics}}(0.2351) & v^+_{\text{social}}(0.4667) + v^-_{\text{social}}(0.2857) \\ v^+_{\text{Speed}}(0.1551) & v^+_{\text{Graphics}}(0.6519) & v^+_{\text{social}}(0.4000) + v^-_{\text{social}}(0.2857) \end{pmatrix}. \tag{5.15}$$

The value function is used with the following parameters and r depending on the criteria

$$v^+(x) = \begin{cases} \sqrt{x-r} & \text{if } x \geq r \\ -2\sqrt{-(x-r)} & \text{if } x < r \end{cases}. \tag{5.16}$$

Applying the value function gives

$$\begin{array}{c} \text{B} \\ \text{C} \end{array}\begin{pmatrix} 0.6079 & 0.3494 & -0.4917 \\ -0.5748 & 0.7341 & -0.5526 \end{pmatrix}. \tag{5.17}$$

When the matrix is multiplied by the weight vector,

$$\begin{array}{c} \text{B} \\ \text{C} \end{array}\begin{pmatrix} 0.6079 & 0.3494 & -0.4917 \\ -0.5748 & 0.7341 & -0.5526 \end{pmatrix}\begin{pmatrix} 0.1677 \\ 0.3487 \\ 0.2418 \end{pmatrix}, \tag{5.18}$$

the following scores are calculated

$$\begin{array}{c} \text{B} \\ \text{C} \end{array}\begin{pmatrix} 0.1049 \\ 0.0260 \end{pmatrix}. \tag{5.19}$$

This shows a reversal of rank due to the reference. This is partially due to the endowment effect, where the decision maker owns product A which is almost twice as fast as C. Indeed, deciding to purchase C would mean a loss of speed, which is much more important than gains in graphics, for which both B and C are better.

Chapter 6

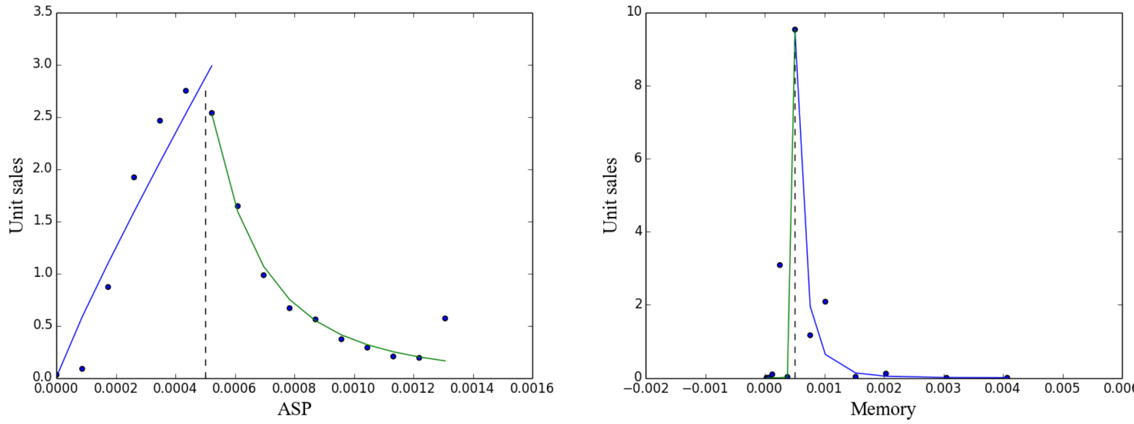
MARKET DATA

In the original paper describing prospect theory, data was collected from students at various universities (Kahneman and Tversky, 1979, p. 264). Later in their paper on cumulative prospect theory, similar experiments were conducted with Stanford and Berkeley graduate students (Tversky and Kahneman, 1992, p. 305). In each case, the individuals were asked to choose which of two prospects (outcome x with probability p or outcome y with probability q) would give the highest utility. This data was used to estimate parameters for the weighting and value functions.

This thesis, however, analyzes data from retail sales of technology to find evidence that the market behaves in the same way as an individual. Specifically, since there is no uncertainty in specifications of the technology (the criteria), the data is examined for evidence that the market as a whole behaves in the same way as an individual in determining the success of a product.

Because sales is one of the clearest indicators of utility, the assumption is made that the number of units sold of a product corresponds to the utility of that product. One would then expect that similar products which differ in only one dimension will show loss aversion. I.e. with respect to a reference, decreases in that dimension will show a greater loss of utility than the equivalent gain. Additionally, larger changes in that dimension will eventually level out, showing decreasing sensitivity to marginal changes.

Anonymized data for retail sales of notebook computers by price was supplied by Intel Corporation for years 2013 and 2014, and included the average selling price (ASP) binned to the nearest $100 dollars, the amount of RAM in the laptop, and the



(a). ASP (b). Memory

Figure 6.1: Unit Sales Plotted Against ASP and Memory.

size of the display.

Figures 6.1a and 6.1b show the sum of the sales of laptops for each criterion from both years, with a nonlinear regression done to find the parameters of the value function (equation 5.1). Because the data is sensitive, the actual numbers have not been included. Additionally, all results presented have been re-scaled to mask the actual numbers without changing the trends.

The reference chosen is the weighted mean of the criteria where the weights are the unit sales. This is shown with the dashed vertical line in figures 6.1a, 6.1b and 6.2. Decreases in memory and display, along with increases in price show decreases in sales, implying loss aversion. Additionally, the decrease in unit sales levels out the larger losses, implying a decreasing sensitivity. Gains in the criteria, price discounts, larger memory capacity, and bigger displays, unexpectedly also show fewer unit sales.

However, where the sales are expected to increase or stay the same for gains (price discounts, larger memory capacity, and bigger displays), they instead decrease sharply, implying that any change relative to the reference is perceived as a loss. While this is not what prospect theory predicts, it bares resemblance to the results

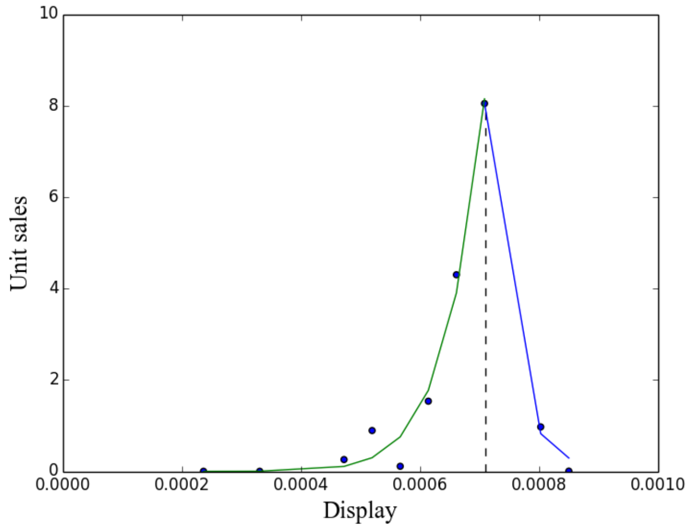


Figure 6.2: Unit Sales Plotted Against Display Size.

of Bridges *et al.* (1995), who find the same sharp decrease in any deviation from the reference. The explanation offered is that potential customers will judge the quality of a product based on price, and infer lower priced units must be inferior in quality. This does not explain why small discounts from the expected price are perceived as a loss of utility.

In addition to laptops, sales data for smartphones and tablet computers were provided by Intel Corporation. This data had much more detail provided, both unit sales and ASP of each product was given per quarter from 2010Q1 to 2014Q1. Because the products were available at different times and for different lengths (some were introduced late, or reached the end of their life earlier), the average sales per quarter is used as the basis for comparisons.

In figure 6.3, the average selling price varies by small amounts. It is necessary to bin the data in a useful way. By looking at figure 6.3, one can identify the clusters of prices with a high volume of sales. To identify these clusters and use them as bins to aggregate the sales, an algorithm was written which is described in the appendix.

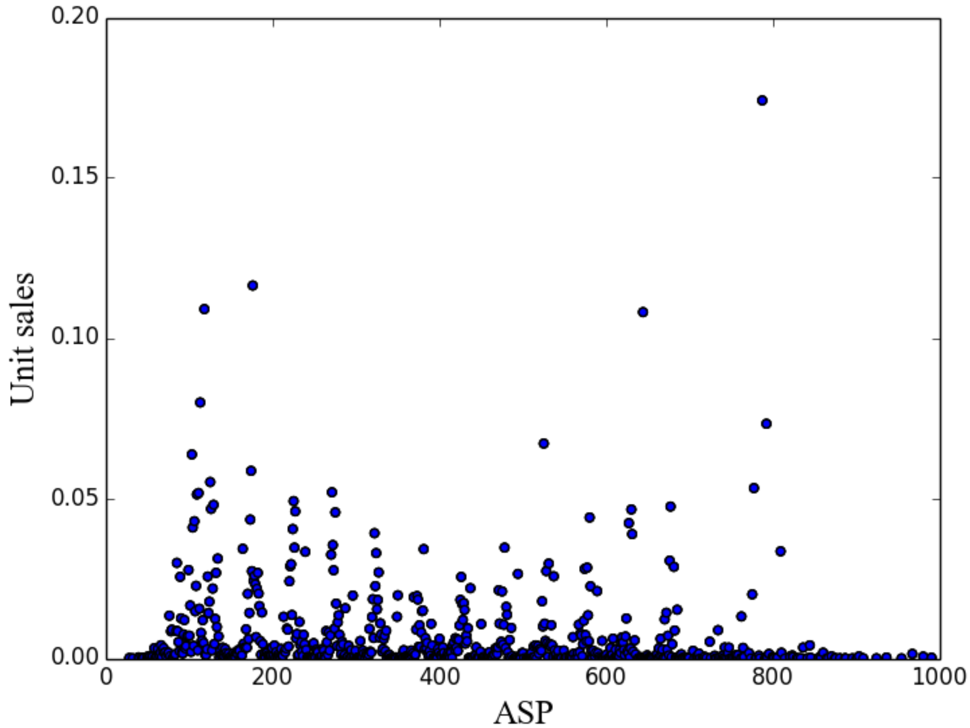
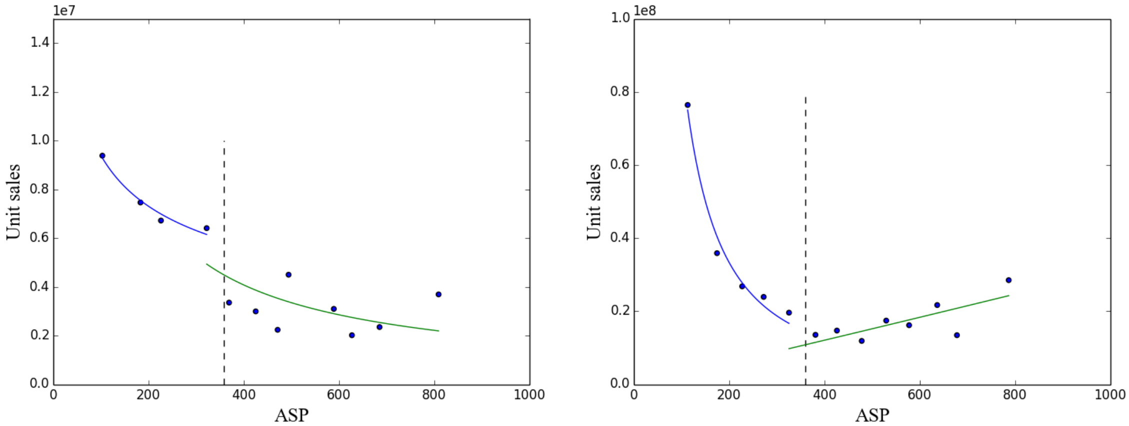


Figure 6.3: Smartphone and Tablet Sales by ASP.

Figures 6.4a and 6.4b are created by summing the sales at each bin and plotting that against the ASP. Using the weighted mean of ASPs again as the reference, indicated by the dashed lines, the data points are fitted to the function in equation 5.3 using nonlinear least squares, with a regression done for gains and losses separately.

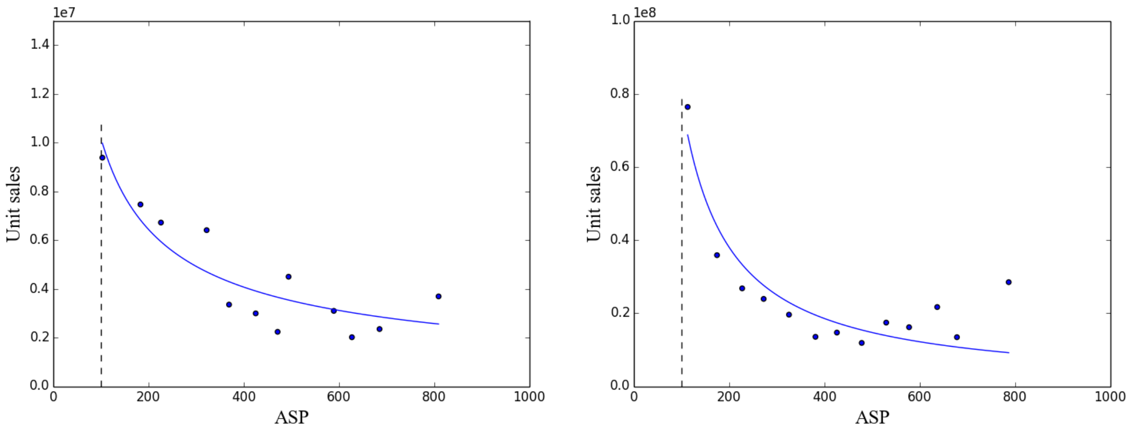
The analysis on the smartphone and tablet data show that utility is increasing with cheaper products with respect to the reference. However the utility from decreases in price does is not concave, which is predicted by prospect theory.

Figures 6.5a and 6.5b are generated using a different reference, using a price lower than the smallest ASP bin. In this context all the smartphone and tablets could be seen as losses relative to this hypothetical cheap laptop and provides a better fit for the data. While using such a low priced laptop does not necessarily reflect reality,



(a). Tablets (b). Smartphones

Figure 6.4: Binned Sales for Tablets and Smartphones.



(a). Tablets (b). Smartphones

Figure 6.5: Binned Sales for Tablets and Smartphones with a Reference ASP Lower Than What is Offered.

it does demonstrates one of the difficulties of prospect theory, that using different references give different results (Barberis, 2012, p.178).

While sometimes there is evidence that the market obeys some of the features of prospect theory, it is not clear if it is an overall trend. This is partly because of the difficulty in knowing what the reference should be, and also because there are other factors which can skew the data, such as extremely low priced laptops.

Chapter 7

CONCLUSIONS

This thesis has presented a model for the decision process of a consumer. Because demand is an essential input to many models intended to help managers optimize inventory control or decide prices, understanding the dynamics of how the consumer makes a decision is important.

Prospect theory is a widely successful theory which captures the irrationality of consumers. It accounts for the effects of reference dependence, where gains or losses are relative to a reference, loss aversion, where losses are perceived as larger than gains, and diminishing sensitivity, where marginal changes in utility have diminished influence.

Kahneman and Tversky's value function is the basis for this thesis, but there is an inherit problem with measuring what is meant by gains or losses in a product attribute, especially with attributes that are purely subjective. This is resolved by applying the analytic hierarchy process, which allows for measuring these intangible properties. These measurements are used as an input to the functions of prospect theory to account for people's subjective biases in thinking.

While AHP and prospect theory are intended to reflect an individual's decision process, it is not complete without accounting for social dynamics in the decision, This is supported by a wealth of evidence (Abrahamson and Rosenkopf, 1997, pp. 289-290)(Delre *et al.*, 2007, pp. 186-188) showing that people routinely base decisions on word of mouth. The hugely successful Bass model is itself predicated on this idea. The effects of social interaction and the word of mouth on the decision process was added as an attribute in the AHP process.

This model was applied to retail sales data provided by Intel. The unit sales, which was assumed to be an indicator of utility, was analyzed for evidence that the market as a whole behaves similarly to prospect theory. However there was not a lot of evidence to support this hypothesis. This failure can mostly by attributed to the fact that prospect theory, and also this model, are intended to describe an individual and those individual consumers may vary widely in their value functions and reference points.

Further research with data may solve the question of market sales arising from many irrational individuals. It should focus classifying the individuals in a market and accounting for their differences. Aggregating the individuals to find the distributions of prospect theory's parameters can help test the hypothesis that the market is composed of many individuals making judgements consistent with prospect theory.

Another extension of the model is to create a more sophisticated approach to the social component. People may weigh the reliability of the network partners differently, which can be captured by using the weighting method in AHP. Additionally, positive and negative feedback may only be part of what goes into a person's decision, and many papers detailing complex behaviors associated with fads, fashions and diffusion may be relevant. Different networks can also be used to classify different groups of consumers. For example, the network of industry professionals may be quite different than the network of teenagers.

Finally the work presented in this thesis can be implemented in an agent based simulation to study the dynamics of the demand and the sensitivity of the model to parameters.

REFERENCES

Abrahamson, E. and L. Rosenkopf, "Institutional and competitive bandwagons: Using mathematical modeling as a tool to explore innovation diffusion", Academy of management review **18**, 3, 487–517 (1993).

Abrahamson, E. and L. Rosenkopf, "Social network effects on the extent of innovation diffusion: A computer simulation", Organization science **8**, 3, 289–309 (1997).

Aral, S. and D. Walker, "Creating social contagion through viral product design: A randomized trial of peer influence in networks", Management Science **57**, 9, 1623–1639 (2011).

Arthur, W. B., "Competing technologies, increasing returns, and lock-in by historical events", The economic journal pp. 116–131 (1989).

Barberis, N. C., "Thirty years of prospect theory in economics: A review and assessment", Tech. rep., National Bureau of Economic Research (2012).

Bass, F. M., "A New Product Growth for Model Consumer Durables", Management Science **15**, 5, 215–227 (1969).

Bernasconi, M., C. Choirat and R. Seri, "The analytic hierarchy process and the theory of measurement", Management Science **56**, 4, 699–711 (2010).

Bikhchandani, S., D. Hirshleifer and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades", Journal of political Economy pp. 992–1026 (1992).

Bleichrodt, H. and J. Miyamoto, "A characterization of quality-adjusted life-years under cumulative prospect theory", Mathematics of Operations Research **28**, 1, 181–193 (2003).

Bleichrodt, H., U. Schmidt and H. Zank, "Additive utility in prospect theory", Management Science **55**, 5, 863–873 (2009).

Bridges, E., C. K. Yim and R. A. Briesch, "A high-tech product market share model with customer expectations", Marketing Science **14**, 1, 61–81 (1995).

David, P. A., "Clio and the economics of qwerty", The American economic review pp. 332–337 (1985).

David, P. A. and J. A. Bunn, "The economics of gateway technologies and network evolution: Lessons from electricity supply history", Information economics and policy **3**, 2, 165–202 (1988).

Delre, S. A., W. Jager and M. A. Janssen, "Diffusion dynamics in small-world networks with heterogeneous consumers", Computational and Mathematical Organization Theory **13**, 2, 185–202 (2007).

Forman, E. H. and S. I. Gass, “The analytic hierarchy process-an exposition”, Operations research **49**, 4, 469–486 (2001).

Janssen, M. A. and W. Jager, “Simulating market dynamics: Interactions between consumer psychology and social networks”, Artificial Life **9**, 4, 343–356 (2003).

Kahneman, D. and A. Tversky, “Prospect theory: An analysis of decision under risk”, Econometrica: Journal of the Econometric Society pp. 263–291 (1979).

Kalwani, M. U. and C. K. Yim, “Consumer price and promotion expectations: an experimental study”, Journal of marketing Research pp. 90–100 (1992).

Katz, M. L. and C. Shapiro, “Network externalities, competition, and compatibility”, The American economic review pp. 424–440 (1985).

Katz, M. L. and C. Shapiro, “Technology adoption in the presence of network externalities”, The journal of political economy pp. 822–841 (1986).

Kiesling, E., M. Günther, C. Stummer and L. M. Wakolbinger, “Agent-based simulation of innovation diffusion: a review”, Central European Journal of Operations Research **20**, 2, 183–230 (2012).

Liebowitz, S. J. and S. E. Margolis, “Path dependence, lock-in, and history”, JL Econ. & Org. **11**, 205 (1995).

Raman, K. and F. M. Bass, “A general test of reference price theory in the presence of threshold effects”, Tijdschrift voor Economie en management **47**, 2, 205–226 (2002).

Rogers, E. M., *Diffusion of innovations* (Simon and Schuster, 2010).

Saaty, T. L., “A scaling method for priorities in hierarchical structures”, Journal of Mathematical Psychology **15**, 3, 234–281 (1977).

Saaty, T. L., “How to make a decision: the analytic hierarchy process”, European journal of operational research **48**, 1, 9–26 (1990).

Stevens, S. S., “On the theory of scales of measurement”, Science **103**, 2684, 677–680 (1946).

Stott, H. P., “Cumulative prospect theory’s functional menagerie”, Journal of Risk and Uncertainty **32**, 2, 101–130 (2006).

Tversky, A. and D. Kahneman, “Loss aversion in riskless choice: A reference-dependent model”, The Quarterly Journal of Economics **106**, 4, 1039–1061 (1991).

Tversky, A. and D. Kahneman, “Advances in prospect theory: Cumulative representation of uncertainty”, Journal of Risk and Uncertainty **5**, 4, 297–323 (1992).

Watts, D. J., *Small worlds: the dynamics of networks between order and randomness* (Princeton university press, 1999).

Watts, D. J. and S. H. Strogatz, “Collective dynamics of small-worldnetworks”, Nature **393**, 6684, 440–442 (1998).

Zank, H., “Cumulative prospect theory for parametric and multiattribute utilities”, Mathematics of Operations Research **26**, 1, 67–81 (2001).

APPENDIX A

BINNING ALGORITHM

This algorithm takes as an input the ASP values and the associated unit sales at each ASP. It returns a two dimensional array, one dimension being ASP, the second being ones and zeros, which are the entries where there is a high volume of sales. It takes two parameters BAR and SENSITIVITY. BAR is the cut-off threshold where anything above is considered a high sale, and below is considered a low sale. SENSITIVITY is how sensitive the algorithm is to detecting clusters.

In creating this array the algorithm puts each entry as either a one, if sales at that entry is above the BAR parameter, or a zero, if sales are below BAR. It then takes that array and counts the number of sequential ones and zeros. If it reads a one, it assumes it is in a high-sales cluster and does nothing. If it detects a sequence of zeros which is shorter in length than SENSITIVITY it changes those entries to ones. If it detects a sequence of zeros longer than SENSITIVITY then it assumes it is in a low-sales cluster and does nothing. This is because there can be individual entries of low-sales within a cluster of high-sales, for which it is desirable to still be considered a part of a high sales cluster.

The output is an array of alternating sequences of ones and zeros (and each entry is associated with a price point), each sequence is at least as long as SENSITIVITY. For each sequence of ones, the weighted arithmetic mean of the corresponding ASPs is taken as the bin. Each weight is the number of sales at that ASP from before the algorithm started. The calculated mean is then used as a bin. Increasing the BAR parameter or decreasing the SENSITIVITY will give more bins and vice versa.

```
Data: Unit sales by ASP
Result: Locations of clusters
SET parameter SENSITIVITY as an integer
SET parameter BAR as a floating point
SET variable NUMBER-OF-CONSECUTIVE-ZEROS as 0
SET variable NOT-IN-CLUSTER as TRUE
SET variable LAST-ONE as 0
SET variable CURRENT-INDEX as 0
for Each ASP do
    if Units at current ASP ≤ BAR then
        Set unit sales to 0
    else
        Set unit sales at ASP to 1
    end
end
while CURRENT-INDEX is not at the end of the data do
    while NOT-IN-CLUSTER do
        if Units at CURRENT-INDEX of ASP == 1 then
            Set NOT-IN-CLUSTER to FALSE. Set LAST-ONE as
            CURRENT-INDEX.
        end
        Increment CURRENT-INDEX. BREAK if at end of data.
    end
    while |LAST-ONE - CURRENT-INDEX| < SENSE do
        if Units at CURRENT-INDEX of ASP == 1 then
            Set units sold of all ASP from LAST-ONE to CURRENT-INDEX
            to 1. Set LAST-ONE to CURRENT-INDEX.
        end
        Increment CURRENT-INDEX. BREAK if at end of data.
    end
end
```