Using Curvilinear Spline Regression To Empirically Test Relationships Predicted By Prospect Theory

Zinaida Taran, (E-mail ztaran@siena.edi), Siena College
Stephen C. Betts, (E-mail bettss@wpunj.edu), William Paterson University

ABSTRACT

Prospect theory (Kahneman & Tversky, 1979) suggests that decision makers compare decision criteria against a reference point when evaluating alternatives. Specifically it posits that decision makers are risk-seeking for losses (below the reference point) and risk-averse for gains (above the reference point). It further proposes that the degree of risk aversion above the reference point is greater than the risk seeking below it. This theory has received widespread acceptance due to intuitive appeal and theoretical support. However the theory does not have strong evidentiary support in actual practice because it is rarely empirically tested in non-experimental situations involving real market data. Often the type or amount of data available does not lend itself to the examination of relationships posited by prospect theory, however even if he data is appropriate, difficulties may arise in modeling and testing. In this paper, after a brief discussion of prospect theory and situations where it is applicable, we present an approach to the empirical testing of prospect theory predictions using curvilinear spline (piecewise polynomial) regression. Among the issues addressed are adequacy of data, choice of inflection point, modeling the curves and hypothesis testing.

INTRODUCTION

Prospect theory (Kahneman & Tversky, 1979) has been one of the most influential theories of decision making. It explains how people make choices among alternatives (prospects) by comparing them to reference points. Daniel Kahneman was awarded the Nobel Prize in 2002 "for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty" (The Nobel foundation, 2002). Before prospect theory, individual decision behavior was explained and predicted by expected utility theory (Von Neumann & Morgenstern, 1944). Expected utility theory posits that rational choices are made based on the levels of relevant decision criteria (Friedman & Savage, 1948). However, the rationality assumptions of expected utility theory have been questioned and have not found empirical support (Slovic & Tversky, 1974).

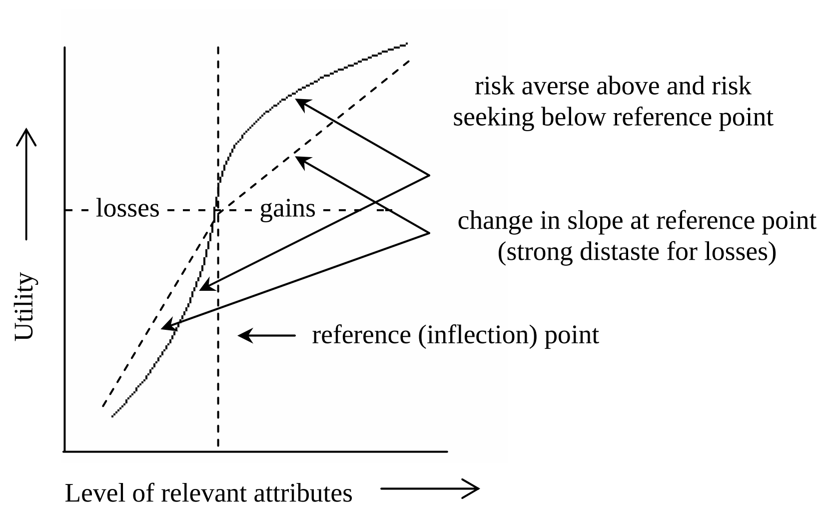
In particular, human choices often depend on the framing of a problem (Tversky & Kahneman, 1986; Hogarth & Reder, 1986), which cannot be explained by the expected utility theory (Tversky, 1969), but follow the propositions of the prospect theory (Kahneman & Tversky, 1979; Shafir, 1999). According to the prospect theory, people do not use the absolute value of the outcome but rather 'code' the alternative as a gain or a loss relative to a reference point. The resulting losses or gains are then weighed by their perceived probabilities of occurrence, forming a non-linear value function. People approach gains and losses differently, generally acting risk-averse on gains and risk-seeking on losses (Kahneman & Tversky, 1979). "We have an irrational tendency to be less willing to gamble with profits than with losses..." (Ghiglino & Tvede, 2000).

In this paper, we suggest using curvilinear spline (piecewise polynomial) regression to test prospect theory predictions. We outline prospect theory and the areas of its application first. Next, we discuss testing its predictions and the issues that may arise.

PROSPECT THEORY

Nobel-winning Prospect theory (Kahneman & Tversky, 1979) is one of the most influential theories of decision-making. It was developed to explain instances where the traditional expected utility theory failed to explain people's choices (Tversky, 1969). It can be briefly summarized as follows: people 'code' the outcomes of various prospects (alternative outcomes) as either gains or losses relative to some reference point. Then they weigh the resulting gains and/or losses by their subjective probabilities. When weighing the outcomes, they weight gains differently from losses, acting risk-seeking for losses, and risk-averse for gains (See Figure 1). This concept has received its wide application as reference price in pricing, where prospect theory is most widely used (Erdem, Mayhew, & Sun, 2001; Niedrich, Sharma, & Wedell, 2001). Tversky and Kahneman (1986) also observed what they referred to as a "strong distaste for losses", when people derive more disutility from a loss than they derive utility from an equivalent amount of gain.

Figure 1 - Prospect Theory Effect



Prospect theory has become one of the most widely cited theories and is utilized in finance, economics, management, decision theory and political science and many more disciplines. Marketing is among such disciplines. People's decisions involving money (prices, discounts, coupon promotions, advertising, monetary incentives including sales force compensation, product bundling (Jagpal, 1998; Stremersch & Tellis, 2002; Johnson, Herrmann, & Bauer, 1999) have been frequent applications of the prospect theory in marketing. Somewhat less frequently, prospect theory was applied to decisions regarding time (Mowen & Mowen 1991; Leclerc, Schmitt & L. Dube´, 1995) and very rarely to quality (Ong, 1994; Betts & Taran, 2003; Betts & Taran, 2006).

Prospect theory has been elaborated upon and expanded into cumulative prospect theory which allows for cumulative utility driving choices among any number of outcomes, including continuous distributions (Tversky & Kahneman, 1992). It accomplishes this by employing cumulative, rather than separate, decision weights, and also allowing to accommodate uncertainty (when probabilities are unknown) in addition to risk (when probabilities are known). This development allows to apply the theory to a greater variety of domains, including taxpayer decisions (Anderson, 1997), lotteries (Donkers, Melenberg & Van Soest, 2001), racetrack betting (Jullien & Salanie, 2000), and consumer choices among many products and features. It also leads to increased appropriateness of utilizing methods such as regression analysis to research the predictive power of its postulates in various situations.

USING REGRESSION TO FIT MODELS BASED ON PROSPECT THEORY

Most empirical research involving the prospect theory relies on designed laboratory experiments (e.g. Salminen & Wallenius, 1993), field experiments (e.g. List, 2004; Bolton & Lemon, 1999), surveys (e.g. Donkers, Melenberg & Van Soest, 2001) or panel data (e.g. Mayhew & Winer, 1992). Such reliance on experimental/survey generated data raises concerns related to external validity (generalizability) of the findings. While there have been a great many such studies, and they have shown strong support for the basic tenets of prospect theory, we wish to suggest a way to expand into modeling observational, field data based on prospect theory predictions.

Such field data can be analyzed utilizing non-linear regression methods. Since we are expecting the model to be changing its behavior past reference points, spline (piecewise) regression could be an appropriate tool (Smith, 1979). Spline regression (piecewise polynomial regression) is a regression models in which the function changes at one or more points (knots) along the range of the predictor (Seber & Wild, 1989, Ch. 9).

We suggest using curvilinear spline regression using the following steps:

1. Define an objective measure of utility corresponding to different levels of your variable under investigation. (For example, price that people are willing to pay as evidenced by their maximum bid on eBay, or price that people are willing to pay to a third party for a used vehicle).
2. Define reference points (for example, average in its class – see discussion below) and compute deviations from it: this is the new independent variable.
3. Introduce an indicator variable I = 0 below the reference point, and =1 above the reference point.
4. Apply functional transformation (see discussion below) to the deviations from the reference point. This is X_{curve} - the term that models the curve predicted by utility theory.
5. Multiply I by the transformed deviation X_{curve}. This is the term that models the disutility of loss relative to the utility if gain.
6. Run regression with both, X_{curve} and $I*X_{curve}$ among the independent variables.

Let us discuss methodological/conceptual issues arising while applying piecewise polynomial regression to field data to fit models based on prospect theory. Particularly, we will discuss aggregation of data, fitting of the curve in line with prospect theory predictions, shift in slope, and choice of reference point.

Aggregation And Other Data Issues

Publicly available sources (Consumer Reports, etc.) oftentimes report data that is aggregate, for instance, national average over some period. This may cause difficulties in modeling if heterogeneous segments within the population are suspected (Jagpal, 1998). A disaggregated data set and explicit consideration of group memberships and separate reference points may be the optimal route. Another route may be to limit categories. In some instances, possible heterogeneities are confounded with categories (brands), and once the data is centralized, can be considered under control (Betts & Taran, 2003).

It is important that the data contains a variable which can be considered a reasonable approximation to people's utilities associated with different outcomes.

Curve

The shape proscribed by the prospect theory is that the utility function is concave over gains and convex over losses (see Figure 1.) The *exact* functional relationship is not proscribed by the prospect theory and such parameterizations have been subject to examination in the literature (Neilson & Stowe, 2002). The task is to find a functional transformation that possesses the general features – concave past reference point and convex below reference point. One such function is cubic root $Y=\sqrt[3]{X}$. A possibility is actually any root of degree of 3, 5, 7, 9,… etc. as long as it's odd and thus allows for the function of a negative X to be negative. Another function that describes similar relationship is logistic function, with proper modifications to fit its assumptions. For the sake of parsimony, we opt for the cubic root; however, if other transformations fit the data better, they should be used instead.

The curve term has to be statistically significant and its sign has to be in the direction predicted for the model to be considered fit.

Inflection Point(s)

Reference point is the cornerstone of human evaluation of alternatives, according to the prospect theory. Determining the location of the reference point requires much consideration. It can be exogenous to the study, provided by a supplemental study or existing sources. In particular, a survey related to perceptions of such a reference point ("fair price", "fair compensation", "just right fit") can be a beneficial further research to studies that didn't have an exogenous reference point (Betts & Taran, 2003). If no such given reference point is available, several choices can be made.

Given the category (for example, age of the car, or certain brand, or type of the product or activity), the reference point can be the average – mean, median or mode depending on the nature of the data. For example, average (mean) quality rating of a 5-year-old vehicle, or average (mean) quality rating among all cars. We believe this is a defensible choice of the reference point location, especially in a situation where consumers are highly likely to have all the necessary information and comparison data. Category averages are oftentimes reported and discussed in various situations by the information sources from the Consumer Reports to the evening news to the children's school report cards, and so we propose that it is a rather likely baseline for the consumers to form their judgments around.

The reference point may be located at one of the ends of the distribution, especially in the presence of a clear leader among the choices (for example, anecdotally, Hellmann's mayonnaise is the standard of a mayonnaise, and any other mayonnaise is judged against it; Hellmann's thus will serve as the reference point). Information on the presence of the leader, in lieu of a known exogenous reference point, should point in the direction of the location of the reference point. The reference point may be also located deciles up and down from the average. Differences in model fit would suggest superior choice among. Unless additional information indicating a leader driving consumer preferences is known, the average seems the most logical choice to start.

Formal methods of finding a reference point are available as well. A log-likelihood function L of all the model parameters (regression coefficients and the location of the reference point) is maximized with respect to the above parameters. The reference point is chosen such that it maximizes the L (Kmenta, 1971, pp.568-569). There is certain degree of difficulty involved in implementing such optimization. Alternatively, finite mixture models can be used to find the best fitting parameters when some individuals adopt a gain frame, and others adopt a loss frame (Wang & Fishbeck, 2004).

It is also possible for there to be multiple points - for example, within a kind – for cars, luxury cars vs. small cars. As pointed out above, heterogeneity of consumer segments might lead to several reference points upon aggregation. This problem will require a disaggregated data set, with a possible two-step approach: clustering analysis helping to form disparate groups, with separate reference point for each obtained in the regression on step 2. If such group membership is known, then cluster analysis is not necessary, and group membership may be included into the analysis explicitly. Such disaggregated data set needs to be large enough to accommodate the number of variables.

So far, we have considered the category in which the reference point to be chosen, as given. However, there are various possibilities that need to be considered carefully. Adaptation-level theory (Helson, 1964) addresses the choice of an appropriate basis for reference, specifically proposing that consumers can shift their reference point if adequately compelled to do so. This indicates that there may not be one dominant basis for setting a reference point. For example, in establishing the reference point for the quality of used cars, one might look at all the cars (and thus the 'average car' is a standard reference car, that is, there is a single reference point across all cars); a class of cars (then small cars have their own reference point against which they are judged, without being compared to luxury vehicles), "price-quality tiers" (Lemon & Nowlis, 2002); the brand (thus, there is an 'average Toyota', an average

'Honda', etc. following the well researched brand effects (e.g. Betts & Taran, 2004); the specific model of the car (average Honda Civic); the age of the car (Betts & Taran 2003; Betts & Taran 2006). Selection of particular basis for reference point may be particularly driven by existing external categorization schemes (such as, Consumer Reports' classes) offered by the information providers as well as the stores/brand owners, which may influence the formation of people's reference points.

Asymmetry

A major point of Prospect Theory is the additional disutility of losses relative to gains. Researchers have used absolute and relative loss premiums to test this loss aversion (Schmidt & Traub, 2002). We propose that such disutility will show in the difference of slope between the curves above and below the reference point (See Figure 1). To model such a shift in slope past the reference point, first we introduce a new term, *I*, an indicator variable, which is equal to 0 below the reference point and 1 above the reference point (or vice versa). Then we multiply our transformed curve variable by this indicator variable. For simplicity, we call the result of this multiplication the *slope* term. Any shift in slope past the reference point will show in the regression coefficient for this slope term. Conversely, a statistically significant regression coefficient for the slope term indicates the difference in utility of gain and loss of equal amount. In addition to its statistical significance, it also has to be of the right sign: negative in our coding scheme, or positive if 1 is ascribed to points below the reference point.

Introduction of such slope term may result in some degree of multicollinearity, since for all the points below the reference point it is equal to 0 and all the points above the reference point it is equal exactly to the curve term, which makes the curve and the slope terms correlated. The presence of both terms is desired to test the predictions of prospect theory. If the multicollinearity concern becomes overbearing, ameliorative measures, for example, ridge regression (Neter, Kutner, Nachtsheim, & Wasserman, 1996) may be applied.

Hypothesis Testing

A set of hypotheses can be formulated with regards to model predictions as depicted in Figure 1. Generic hypotheses are approximately as follows (insert your own variables X and Y):

Hypothesis 1: X will have a decreasingly positive effect on Y with X above the reference point and an increasingly positive effect on Y with X below the reference point.

Hypothesis 2: The increase in the relationship between X and Y when X is below the reference point is greater than the decrease in the same relationship when X is above the reference point.

Hypothesis 1 is tested using the curve term, and hypothesis 2 is tested using the slope term. To make sure the addition of the curve and the slope terms into the model improve it on statistically significant level, partial F-test may be performed (Neter et. al., 1996).

CONCLUSIONS

We have suggested six easy steps to fit a model based on the prospect theory to field/observational data provided it contains a valid measure of people's utilities. In a nutshell, the data is recoded into deviation from the selected reference point, transformed to reflect the curve and another term is added to model the shift of the slope. Careful consideration needs to be given to the selection of reference point, fitting the appropriate curve, and selecting data that can yield meaningful results.

It can be a useful tool for researchers investigating people's preferences in every area where the prospect theory applies, and allows using observational data, including secondary sources.

It can be also a useful tool for teaching advanced level social or management research classes, as well as topics related to human decision making and choices in senior undergraduate or graduate courses in management, marketing, consumer behavior, sales, etc, especially if students are expected to carry out a research project.

The implication for practitioners rests primarily on careful consideration of the interplay between the reference point and the rest of the model: it is to your advantage to influence your constituency (consumers, employees, etc) to form their reference points at exactly the level that you can provide. Anything less than you can provide will be coded as losses and affect your constituents more (disutility of loss); anything more than what you can provide will be coded as gains and affect them less.

REFERENCES

1. Anderson, K.E. (1997). An Application of Cumulative Prospect Theory to the Decisions of Taxpayers under Risk. *The Journal of the American Taxation Association*, 19(1), pp. 122-123.
2. Betts, S.C. & Taran, Z. (2006). The halo effect of age on durable goods prices: age, reliability and the used car market. *Review of Business Research*, forthcoming.
3. Betts, S.C. & Taran, Z. (2003). Prospect theory and perceptions of quality: non-linear effects of quality comparisons on price in the used car market. *Academy of Marketing Studies Journal*, 7(1), 1-10.
4. Betts, S.C. & Taran, Z. (2004). The 'brand halo' effect on durable goods prices: brand reliability and the used car market. *Academy of Marketing Studies Journal*, 8(1), 7-18.
5. Bolton, R.N. & Lemon, K.N. (1999). A dynamic model of customers' usage of services: Usage as an antecedent and consequence of satisfaction. *Journal of Marketing Research*. 36(2), 171-187.
6. Donkers, B., Melenberg, B., & Van Soest, A. (2001). Estimating risk attitudes using lotteries: A large sample approach. *Journal of Risk and Uncertainty*. 22(2); p. 165 -195.
7. Erdem, T., Mayhew, G., & Sun, B. (2001). Understanding reference-price shoppers: A within- and cross-category analysis. *Journal of Marketing Research*, 38(4), 445-457.
8. Friedman, M. & Savage, L. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, 56, 279-304.
9. Ghiglino, C. & Tvede, M. (2000). Optimal policy in OG models. *Journal of Economic Theory*, 90(1), 62-83.
10. Helson, H. (1964). *Adaptation-level theory*. New York: Harper.
11. Hogarth, R.M. & Reder, M.W. (1986). Editors' comments: Perspectives from economics and psychology. *The Journal of Business*, 59(4), S185-S207.
12. Jagpal, H.S. (1998). *Marketing under uncertainty*. Oxford: Oxford University Press.
13. Johnson, M., Herrmann, A., & Bauer, H. (1999). The effects of price bundling on consumer evaluations of product offerings. *International Journal of Research in Marketing*, 16(2), 129-142.
14. Jullien, B. & Salanie, B (2000). Estimating preferences under risk: The case of racetrack bettors. *The Journal of Political Economy*. Chicago: 108(3), 503-531.
15. Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263-291.
16. Kmenta, J. (1971). *Elements of Econometrics* (2nd ed.). New York: Macmillan.
17. Leclerc, F., Schmitt, B., & Dube´, L. (1995). Waiting time and decision making: is time like money? *Journal of Consumer Research*, 22(1), 110-119.
18. Lemon, K. & Nowlis, S. (2002). Developing synergies between promotions and brands in different price – quality tiers. *Journal of Marketing Research*, 39(2), 171- 185.
19. List J.A. (2004). Neoclassical theory versus prospect theory: evidence from the marketplace. *Econometrica*, 72(2), 615-626.
20. Mayhew, G.E. & Winer, R.S. (1992). An empirical analysis of internal and external reference prices using scanner data. *Journal of Consumer Research*, 19(1), 62-70.
21. Mowen, J. & Mowen, M. (1991). Time and outcome valuation: implications for marketing decision making. *Journal of Marketing*, 55(4), 54-62.
22. Neilson, W. & Stowe, J. (2002). A Further Examination of Cumulative Prospect Theory Parameterizations. *Journal of Risk and Uncertainty*, 24(1), 31-46.

23. Neter, J., Kutner, M.H., Nachtsheim, C.J., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). New York: WCB/McGrawHill.
24. Niedrich, R.W., Sharma, S., & Wedell, D. (2001). Reference price and price perceptions: A comparison of alternative models. *Journal of Consumer Research*, 28(3), 339-354.
25. The Nobel Foundation (2002). The Bank of Sweden prize in economic sciences in memory of Alfred Nobel 2002. retrieved February 15, 2005 from http://nobelprize.org/economics/laureates/2002/.
26. Ong, B.S. (1994). Conceptualizing "reference quality" claims: Empirical analysis of its effects on consumer perceptions. *American Business Review*, 12(1), 86-94.
27. Salminen, P. & Wallenius, J.J. (1993). Testing prospect theory in a deterministic multiple criteria decision-making environment. *Decision Sciences*, 24(2), 279-294.
28. Schmidt, U. & Traub, S. (2002). An experimental test of loss aversion. *Journal of Risk and Uncertainty*. 25(3), 233-249.
29. Seber, G.A.F. & Wild, C.J. (1989). *Nonlinear regression*. NY: Wiley.
30. Shafir, E. (1999). Belief and decision: The continuing legacy of Amos Tversky. *Cognitive Psychology*, 38(1), 3-15.
31. Slovic, P. & Tversky, A. (1974). Who accepts savage's axiom? *Behavioral Science*, 19(4), 368-373.
32. Smith, P.L. (1979). Splines as a useful and convenient statistical tool. *American Statistician*, 33(2), 57-62.
33. Stremersch, S. & Tellis, G.J. (2002). The strategic bundling of products and prices: A new synthesis for marketing. *Journal of Marketing*, 66(1), 55-72.
34. Thaler, R. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.
35. Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297-323.
36. Tversky, A. & Kahneman, D. (1986). Rational choice and the framing of decisions. *The Journal of Business*, 59(4), S251-S278.
37. Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31-48.
38. Von Neumann, J. & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.
39. Wang, M. & Fishbeck, P.S. (2004). Incorporating framing into prospect theory modeling: a mixture-model approach. *Journal of Risk and Uncertainty*, 29(2), 181- 197.

NOTES