Working Paper No. 426

How to Extend a Model of Probabilistic Choice from Binary Choices to Choices among More Than Two Alternatives

Pavlo R. Blavatskyy

September 2009

How to Extend a Model of Probabilistic Choice from Binary Choices to Choices among More Than Two Alternatives

Pavlo R. Blavatskyy[*]

University of Zurich (IEW)
Winterthurerstrasse 30
CH-8006 Zurich
Switzerland

Phone: +41(0)446343586
Fax: +41(0)446344978
e-mail: pavlo.blavatskyy@iew.uzh.ch

July 2009

Abstract: This note presents an algorithm that extends a binary choice model to choice among multiple alternatives. Both neoclassical microeconomic theory and Luce choice model are consistent with the proposed algorithm. The algorithm is compatible with several empirical findings (asymmetric dominance and attraction effects) that cannot be explained within standard models.

Keywords: probabilistic choice, binary choice, multiple alternatives

JEL classification codes: C44, D11, D70, D81

How to Extend a Model of Probabilistic Choice from Binary Choices to Choices among More Than Two Alternatives

Many popular models of probabilistic choice such as the Fechner model of random errors (Fechner, 1860; Hey and Orme, 1994; Blavatskyy, 2008) are developed for binary choice problems because binary choice tasks are frequently used in empirical (especially experimental) research. However, many applications in the theory of individual demand, operations research, social choice, political economy and etc. involve multiple alternatives and require an extension of binary choice models to choices among multiple alternatives. For example, given marketing data on how likely an average consumer is to choose product A over product B, product B over product C and product A over product C, a retail firm may be interested to know the probability that product A is purchased when all three products A, B and C are offered on the market.

This note presents an algorithm that extends any binary choice model to choice among more than two alternatives. This algorithm is closely related to an example in Luce and Suppes (1965, pp. 351-352) who show how a ranking over three alternatives can be generated from binary choice probabilities. Köhler (2007) develops a procedural model for solving complex decision problems (such as choice among multiple alternatives) by decomposing a complex problem into a sequence of binary choices.

Let $S=\{A_1,\ldots,A_n\}$ be a finite set of $n \geq 2$ alternatives. Let $P:S \times S \rightarrow [0,1]$ be a binary choice probability function (known as fuzzy preference relation, *e.g.* Zimmerman et al. (1984)). $P(A_i,A_j)$ is the probability that a decision maker chooses alternative $A_i \in S$ over alternative $A_j \in S$ in a binary choice. For any two alternatives $A_i,A_j \in S$ such that $A_i \neq A_j$, probability $P(A_i,A_j)$ is observable from a relative frequency with which a decision maker chooses A_i when asked to choose repeatedly between A_i and A_j. Probability $P(A_i,A_i)$ cannot be observed from actual choices and it is not defined for any $A_i \in S$.

This note presents an algorithm how to extend a binary choice model P to choice among $n \geq 2$ alternatives. According to this algorithm a decision maker acts as follows:

1) Select one alternative $A_i \in S$ at random. Label it as alternative B ("the best").

2) Select alternative $A_j \in S$, $A_j \neq B$, at random. Choose between A_j and B. If A_j is chosen over B, which happens with probability $P(A_j, B)$, then label A_j as a new alternative B.

3) Repeat step 2) *ad infinitum*.

At every stage $t > 2$ of this algorithm a decision maker chooses between the alternative that was chosen in the last comparison and another randomly selected element of the choice set S. Note that a new element is selected with replacement, *i.e.* this element can be one of the non-chosen alternatives from previous iterations. Köhler (2007) analyses the same algorithm when selection is done without replacement. The proposed algorithm is not literally an infinite chain of actual pairwise choices. Rather, it is a mental process so that a decision maker is not endowed with the preferred alternative in each choice (and we can abstract from endowment effects).

Let $Q(A_i|S)_t$ denote the probability that the i-th alternative $A_i \in S$ is chosen over another randomly selected alternative at the iteration $t \in \mathbb{N}$ of this algorithm. The probabilities $Q(A_i|S)_t$ are recursively defined by the following system of equations:

$$(1) \quad \begin{pmatrix} Q(A_1|S)_{t+1} \\ Q(A_2|S)_{t+1} \\ \ldots \\ Q(A_n|S)_{t+1} \end{pmatrix} = \frac{1}{n-1} \begin{pmatrix} \sum_{\substack{i=1 \\ i \neq 1}}^{n} P(A_1, A_i) & P(A_1, A_2) & \ldots & P(A_1, A_n) \\ P(A_2, A_1) & \sum_{\substack{i=1 \\ i \neq 2}}^{n} P(A_2, A_i) & \ldots & P(A_2, A_n) \\ \ldots & \ldots & \ddots & \ldots \\ P(A_n, A_1) & P(A_n, A_2) & \ldots & \sum_{\substack{i=1 \\ i \neq n}}^{n} P(A_n, A_i) \end{pmatrix} \cdot \begin{pmatrix} Q(A_1|S)_t \\ Q(A_2|S)_t \\ \ldots \\ Q(A_n|S)_t \end{pmatrix}$$

with the initial condition $Q(A_i|S)_1 = 1/n$ for all $i \in \{1, \ldots, n\}$.

The probability that an alternative A_i is chosen from the set S is defined as $Q(A_i|S) \equiv \lim_{t \to \infty} Q(A_i|S)_t$ for all $i \in \{1, \ldots, n\}$. In other words, we calculate the chance that the i-th alternative is chosen from the set S as if an individual continues the sequence of binary choices *ad infinitum*. A vector $\mathbf{Q} = (Q(A_1|S), \ldots, Q(A_n|S))'$ is an asymptotic probability distribution on S if it solves the following homogeneous matrix equation

(2) $$(\mathbf{P} - (n-1)\mathbf{I}) \cdot \mathbf{Q} = 0,$$

where $\mathbf{P}$ is the transition matrix on the right-hand side of (1) and $\mathbf{I}$ is the $n \times n$ identity matrix.

Equation (2) is a system of homogeneous linear equations. It has a trivial solution $\mathbf{Q} = \mathbf{0}$, which we may discard because all choice probabilities should sum up to one. Equation (2) also has a non-trivial solution, which can be conveniently written in a closed form using notation from graph theory.

Definition 1 An arborescence $G \equiv (S, E)$ is a directed graph, *i.e.* an ordered pair of the vertex set S and the edge set E, which has one vertex $R \in S$, called the root of G, such that for any other vertex $v \in S$, there is only one directed path from R to v.

Let G denote an arborescence with the vertex set S and let $\Gamma(S)$ be the set of all arborescences with the vertex set S. Let $R(G)$ be the root of G and let $E(G)$ be the edge set of G. Note that the elements of $E(G)$ are ordered pairs of alternatives $\{A_i, A_j\}$ such that $A_i, A_j \in S$ and $A_i \neq A_j$. Product $\prod_{\{A_i, A_j\} \in E(G)} P(A_i, A_j)$ is the probability that a decision maker chooses the initial vertex (tail) A_i over the terminal vertex (head) A_j in all edges of an arborescence G. With this notation, the vector $\mathbf{Q}$ that solves (2) can be written as

(3) $$\mathbf{Q} = \lambda \left(\sum_{G \in \Gamma(S) | R(G) = A_1} \prod_{\{A_i, A_j\} \in E(G)} P(A_i, A_j), \ldots, \sum_{G \in \Gamma(S) | R(G) = A_n} \prod_{\{A_i, A_j\} \in E(G)} P(A_i, A_j) \right)',$$

where λ is an arbitrary constant.

Since all choice probabilities should add up to one, *i.e.* $\sum_{i=1}^{n} Q(A_i|S)=1$, it follows immediately that $\lambda = 1/\sum_{G\in\Gamma(S)}\prod_{\{A_i,A_j\}\in E(G)}P(A_i,A_j)$. Thus, a decision maker chooses an alternative $A_k\in S$ from a non-singleton choice set S with a probability

(4)
$$Q(A_k|S) = \frac{\sum_{G\in\Gamma(S)|R(G)=A_k}\prod_{\{A_i,A_j\}\in E(G)}P(A_i,A_j)}{\sum_{G\in\Gamma(S)}\prod_{\{A_i,A_j\}\in E(G)}P(A_i,A_j)},$$

for all $k\in\{1,\dots,n\}$.

Equation (4) allows us to generalize any model of binary choice to choice among $n\geq 2$ alternatives. This generalization has several intuitive properties. For example, if a decision maker always (never) chooses one alternative over every other element of a choice set in a direct binary choice, then he also always (never) chooses this alternative from the overall choice set. If a decision maker chooses with probabilities 50%-50% between any two alternatives from a given choice set, then each alternative is selected with equal probability ($1/n$) from this choice set.

Example 1 (Neoclassical microeconomic theory) In standard microeconomic theory there exists an utility function $u:S\rightarrow\mathbb{R}$ such that

(5)
$$P(A_i,A_j) = 0.5 + 0.5\frac{u(A_i)-u(A_j)}{|u(A_i)-u(A_j)|},$$

for any $A_i,A_j\in S$ such that $A_i\neq A_j$. Formula (5) simply states that alternative A_i is chosen over alternative A_j with probability one if $u(A_i)>u(A_j)$ and with probability zero if $u(A_i)<u(A_j)$. The binary choice probability $P(A_i,A_j)$ is not defined if $u(A_i)=u(A_j)$.[1] If $u(.)$ is an injective (one-to-one) function, we can plug in formula (5) into equation (4) and then we obtain the following result

[1] This indeterminacy plays a crucial role for the existence of a mixed-strategy equilibrium in non-cooperative game theory.

(6)
$$Q(A_k|S)=\begin{cases}1, & u(A_k)=\max_{A_i\in S} u(A_i)\\ 0, & u(A_k)\neq\max_{A_i\in S} u(A_i)\end{cases},$$

for all $k\in\{1,\dots,n\}$. Equation (6) simply states that a decision maker always picks the alternative with the highest utility.

Example 2 (Luce or strict utility model) In Luce or strict utility model (Luce, 1959), there exists a function $v:S\rightarrow\mathbb{R}_+$, which is unique up to a proportionality factor, such that

(7)
$$P(A_i,A_j)=\frac{v(A_i)}{v(A_i)+v(A_j)},$$

for all $A_i,A_j\in S$. If we plug in formula (7) into equation (4) we obtain the following result

(8)
$$Q(A_k|S)=\frac{v(A_k)}{\sum_{A_i\in S} v(A_i)},$$

for all $k\in\{1,\dots,n\}$. Luce (1959) originally derived (8) from a famous Luce choice axiom.

Example 3 (Intransitive choice) Let $n=3$ and let $P(A_1,A_2)=P(A_2,A_3)=P(A_3,A_1)$ so that the decision maker violates weak stochastic transitivity. If we plug these binary choice probabilities into equation (4) we obtain $Q(A_k|S)=1/3$ for all $k\in\{1,2,3\}$.

Example 4 (Asymmetric dominance effect) Asymmetric dominance effect is an empirical observation that a choice probability $Q(A_1|\{A_1,A_2,A_3\})$ is often greater than a binary choice probability $P(A_1,A_2)\neq 1$ if alternative A_3 is dominated by alternative A_1 but not by alternative A_2 so that $P(A_3,A_1)=0$ but $P(A_3,A_2)\neq 0$ (e.g. Huber et al., 1982; Herne, 1999; Dhar and Simonson, 2003). According to equation (4), if $P(A_3,A_1)=0$ then

(9)
$$Q\big(A_1|\{A_1,A_2,A_3\}\big)=P(A_1,A_2)+\frac{P(A_3,A_2)P(A_2,A_1)^2}{2+P(A_3,A_2)P(A_2,A_1)}.$$

The right-hand-side of equation (9) is always strictly greater than $P(A_1,A_2)$ if $P(A_3,A_2)\neq0$ and $P(A_1,A_2)\neq1$. Thus, the proposed algorithm actually predicts asymmetric dominance effect, which cannot be explained within standard models of probabilistic choice such as random utility model or Luce choice model (Theorem 41 in Luce and Suppes, 1965).

Example 5 (Attraction effect) Attraction affect refers to an empirical finding that is very similar to the asymmetric dominance effect. Specifically, people choose an alternative A_1 from the choice set $\{A_1,A_2,A_3\}$ more frequently than they do from the choice set $\{A_1,A_2\}$ if A_3 appears to be inferior to A_1 so that $P(A_3,A_1)$ is relatively small but not equal to zero, and at the same time A_3 is not dominated by A_2 so that $P(A_3,A_2)\neq0$ (e.g. Huber and Puto, 1983; Simonson and Tversky, 1992; Tversky and Simonson, 1993). If a choice probability $Q(A_1|\{A_1,A_2,A_3\})$ is given by equation (4), then this probability is greater than a binary choice probability $P(A_1,A_2)$ if and only if the following holds:

$$(10) \qquad P\left(A_3,A_1\right)<\frac{P\left(A_3,A_2\right)P\left(A_2,A_1\right)^2}{P\left(A_3,A_2\right)+P\left(A_1,A_2\right)P\left(A_2,A_3\right)+P\left(A_1,A_2\right)^2}.$$

The right-hand side of inequality (10) is strictly positive provided that people do not always choose A_1 over A_2 i.e. $P(A_1,A_2)\neq1$. Hence, there always exists a choice probability $P(A_3,A_1)$ sufficiently close to zero such that inequality (10) is satisfied. In other words, the proposed algorithm is compatible with the attraction effect.

Examples 1-3 show that the proposed algorithm is quite general and it can be used to extend very different models of binary choice to choice among multiple alternatives. Moreover, both neoclassical microeconomic theory (Example 1) and Luce choice model (Example 2) are consistent with the proposed algorithm. Thus, we can use the new algorithm to extend binary choice models to choice among multiple alternatives in an analogous manner to choice functions in neoclassical microeconomic theory and

Luce choice model. The proposed algorithm is also compatible with several stylized empirical facts that cannot be explained within standard models of probabilistic choice (Example 4 and 5).

References

Blavatskyy, P. (2008) "Stochastic Utility Theorem" *Journal of Mathematical Economics* **44**, 1049-1056

Dhar, R. and I. Simonson (2003) "The Effect of Forced Choice on Choice" *Journal of Marketing Research* **40**, 146-160

Fechner, G. (1860) "Elements of Psychophysics" NewYork: Holt, Rinehart and Winston

Herne, K (1999) "The Effects of Decoy Gambles on Individual Choice" *Experimental Economics* **2**, 31-40

Hey, J.D. and C. Orme (1994) Investigating generalisations of expected utility theory using experimental data, *Econometrica* **62**, 1291-1326

Huber, J., J.W. Payne, C. Puto (1982) "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis" *Journal of Consumer Research* **9**, 90-98

Huber, J. and C. Puto (1983) "Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects" *Journal of Consumer Research* **10**, 31-44

Köhler, W. (2007) "Why does context matter? Attraction effects and binary comparisons" IEW working paper 330, www.iew.uzh.ch/wp/iewwp330.pdf

Luce, R. D. (1959) "Individual choice behavior" New York: Wiley

Luce, R. D. and Suppes, P. (1965) "Preference, utility, and subjective probability" in R. D. Luce, R. R. Bush & E. Galanter (eds.), *Handbook of mathematical psychology*, Vol. III, 249–410, Wiley, New York NY

Simonson, I., A. Tversky (1992) "Choice in Context: Tradeoff Contrast and Extremeness Aversion" *Journal of Marketing Research* **29**, 281-295

Tversky, A. and I. Simonson (1993) "Context-Dependent Preferences" *Management Science* **30**, 1179-1189

Zimmerman, H.J., Gaines, B.R. and Zadeh, L. A. (1984) "Fuzzy Sets and Decision Analysis", North-Holland