DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A NOTE ON THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES IN PROBABILISTIC CHOICE MODELS

Paul A. Bjorn and Quang H. Vuong
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 576

November 1984

A NOTE ON THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES IN PROBABILISTIC CHOICE MODELS

Paul A. Bjorn and Quang H. Vuong
California Institute of Technology

ABSTRACT

The purpose of this note is to show that there is no necessary relationship between the independence of irrelevant alternatives (IIA) property and stochastic independence of the errors in probabilistic choice models.

A NOTE ON THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES IN PROBABILISTIC CHOICE MODELS

Paul A. Bjorn and Quang H. Vuong
California Institute of Technology

In response to the growing interest among economists of studying economic decisions involving choice over a finite number of alternatives, the literature on probabilistic choice models has rapidly evolved over the last decades (see e.g., Maddala (1983), Manski and McFadden (1981)). By far the most widely used model specification is that of the multinomial logit model (McFadden (1974, 1981)), primarily because of the computational ease in calculating the underlying choice probabilities. While it is well known that an important drawback of the multinomial logit model is the independence of irrelevant alternatives (IIA) property, conventional wisdom seems to suggest that IIA arises because of stochastic independence. The purpose of this short note is to show that there is no necessary relationship between these two concepts. In addition, we show that the more commonly used definition of IIA and a lesser-known definition are indeed equivalent.

We use the following notation. Let $t \in \{1,...,T\}$ be the set of individuals in the population and let Ξ be the set of all alternatives. Following McFadden (1974, 1981), individuals are random utility maximizers, randon utility given by $\tilde{U}_{it} = \bar{U}_{it} + \varepsilon_{it}$ for all $t \in \{1,...,T\}$ and for all $i \in C \subset \Xi$. Then the probability that individual t chooses alternative i from choice set C is given by

$P_{it}^{C} = \Pr(\tilde{U}_{it} > \tilde{U}_{kt},\ \forall k \in C\backslash i) = \Pr(\varepsilon_{it} - \varepsilon_{kt} > \bar{U}_{kt} - \bar{U}_{it},\ \forall k \in C\backslash i)$. For ease of notation, we shall drop the subscript t throughout. We then have two definitions of the IIA property, Definition 1 being the more common.

Definition 1: IIA_1. Let C' and C be any subsets of Ξ and i and j be alternatives such that $\{i,j\} \subset C^1 \subset C \subset \Xi$. Then

$$\frac{P_i^C}{P_j^C} = \frac{P_i^{C'}}{P_j^{C'}} .$$

Definition 1 states that the ratio of the probabilities of choosing alternative i over alternative j is independent of the offered choice set. In some situations, however, the choice set cannot be altered by either adding or suppressing some alternatives. The following definition is then relevant: it states that the above ratio does not depend on the characteristics of other alternatives.

Definition 2: IIA_2. For any C and any $\{i,j\} \subset C$,

$\frac{P_i^C}{P_j^C}$ does not depend on $\bar{U}_k$ for any $k \in C\backslash\{i,j\}$.

PROPOSITION 1 IIA_1 is equivalent to IIA_2.

PROOF:

IIA_1 implies IIA_2. By definition of IIA_1,

$$\frac{P_i^C}{P_j^C} = \frac{P_i^{\{i,j\}}}{P_j^{\{i,j\}}} = \frac{\Pr(\varepsilon_i - \varepsilon_j > \bar{U}_j - \bar{U}_i)}{\Pr(\varepsilon_j - \varepsilon_i) > \bar{U}_i - \bar{U}_j)} .$$

But this last quantity does not depend on $\bar{U}_k$ for any $k \in C\backslash\{i,j\}$.

IIA_2 implies IIA_1. Without loss of generality, let choice set C consist of N alternatives and let i = 1, j = 2. Then

$$\frac{P_1^C}{P_2^C} = \frac{\Pr(\varepsilon_1 - \varepsilon_2 > \bar{U}_2 - \bar{U}_1, \varepsilon_1 - \varepsilon_3 > \bar{U}_3 - \bar{U}_1, \ldots, \varepsilon_1 - \varepsilon_N > \bar{U}_N - \bar{U}_1)}{\Pr(\varepsilon_2 - \varepsilon_1 > \bar{U}_1 - \bar{U}_2, \varepsilon_2 - \varepsilon_3 > \bar{U}_3 - \bar{U}_2, \ldots, \varepsilon_1 - \varepsilon_N > \bar{U}_N - \bar{U}_2)}$$
$$\equiv G(\bar{U}_1, \bar{U}_2, \ldots, \bar{U}_N)$$

But from IIA_2, we know that $P_1^C \backslash P_2^C$ does not depend on $\{\bar{U}_3, \ldots, \bar{U}_N\}$. Since we are therefore free to set $\{\bar{U}_3, \ldots, \bar{U}_N\}$ however we want, let $\bar{U}_k \to -\infty, k = (3, \ldots, N)$. But

$$\lim_{\bar{U}_k \to -\infty, k=(3,\ldots,N)} G(\bar{U}_1, \bar{U}_2, \bar{U}_3, \ldots, \bar{U}_N) = \frac{\Pr(\varepsilon_1 - \varepsilon_2 > \bar{U}_2 - \bar{U}_1)}{\Pr(\varepsilon_2 - \varepsilon_1 > \bar{U}_1 - \bar{U}_2)} = \frac{P_1^{\{1,2\}}}{P_2^{\{1,2\}}}$$

which gives us IIA_1.

Q.E.D.

Since we have shown the two definitions of IIA to be equivalent, we will use Definition 1 throughout the rest of this note. The remaining propositions establish the relationships between the independence of irrelevant alternatives property and stochastic independence of ε_i.

PROPOSITION 2 IID does not imply IIA.

PROOF: It is sufficient to provide one counterexample. Assume a three alternative model, $C' = \{a_1,a_2,a_3\} \subset \Xi$, with corresponding mean utilities given by $\{\bar{U}_1,\bar{U}_2,\bar{U}_3\}$. Let individuals have ranking over mean utilities given by $\bar{U}_1 = \bar{U}_2 + 1 = \bar{U}_3 + 2$. In addition, assume $\varepsilon_i, i = 1,2,3$, takes on values 4/3, 2/3, -2, each with probability 1/3. Therefore $P_1^{C'} = \Pr(\varepsilon_1 - \varepsilon_2 > \bar{U}_2 - \bar{U}_1,\ \varepsilon_1 - \varepsilon_3 > \bar{U}_3 - \bar{U}_1) = \Pr(\varepsilon_1 - \varepsilon_2 > -1,\ \varepsilon_1 - \varepsilon_3 > -2)$, $P_2^{C'} = \Pr(\varepsilon_2 - \varepsilon_1 > \bar{U}_1 - \bar{U}_2,\ \varepsilon_2 - \varepsilon_3 > \bar{U}_3 - \bar{U}_2) = \Pr(\varepsilon_2 - \varepsilon_1 > 1,\ \varepsilon_2 - \varepsilon_3 > -1)$, and $P_3^{C'} = \Pr(\varepsilon_3 - \varepsilon_1 > \bar{U}_1 - \bar{U}_3,\ \varepsilon_3 - \varepsilon_2 > \bar{U}_2 - \bar{U}_3) = \Pr(\varepsilon_3 - \varepsilon_1 > 2,\ \varepsilon_3 - \varepsilon_2 > 1)$. Now alternative 2 is chosen whenever $\varepsilon_1 = -2$, $\varepsilon_2 = 4/3$ or $2/3$, and ε_3 taking on any of the three values. Therefore, $P_2^{C'} = 6/27$. Alternative 3 is chosen only when $\varepsilon_1 = -2$, $\varepsilon_2 = -2$, and $\varepsilon_3 = 4/3$ or $2/3$. Therefore $P_3^{C'} = 2/27$. Thus $P_1^{C'} = 19/27$. Thus

$$\frac{P_1^{C'}}{P_2^{C'}} = \frac{19}{27} .$$

Now eliminate the third alternative such that the set of available alternatives is given by $C = \{a_1,a_2\}$. Then $P_1^C = \Pr(\varepsilon_1 - \varepsilon_2 > \bar{U}_2 - \bar{U}_1) = \Pr(\varepsilon_1 - \varepsilon_2 > -1)$ and $P_2^C = \Pr(\varepsilon_2 - \varepsilon_1 > \bar{U}_1 - \bar{U}_2) = \Pr(\varepsilon_2 - \varepsilon_1 > 1)$. Alternative 2 is chosen whenever $\varepsilon_1 = -2$ and ε_2 takes on the valves 4/3 or 2/3, giving us $P_2^C = 2/9$ and $P_1^C = 7/9$. Therefore $\frac{P_1^C}{P_2^C} = \frac{7}{2} \neq \frac{P_1^{C'}}{P_2^{C'}}$.

Q.E.D.

PROPOSITION 3 IIA does not imply IID.

PROOF: Again, it is sufficient to provide one counterexample. Consider a three alternative model $C = \{a_1,a_2,a_3\}$ with ranking over mean utilities given by $\bar{U}_1 = \bar{U}_2 + 1 = \bar{U}_3 + 2$. Let $\varepsilon_i, i = 1,2,3$, take on the valves -1.5 and 1.5 with the following probabilities

	ε_2: -1.5		1.5	
ε_1 \ ε_3	-1.5	1.5	1.5	1.5
-1.5	1/2	0	0	0
1.5	0	0	0	1/2

It can be readily checked that $\frac{P_i^C}{P_j^C} = \frac{P_i^{\{i,j\}}}{P_j^{\{i,j\}}}$ so that IIA holds. But then we have, for example,

$$\Pr(\varepsilon_1 = 1.5, \varepsilon_2 = 1.5, \varepsilon_3 = 1.5)$$
$$= 1/2 \neq \Pr(\varepsilon_1 = 1.5) \cdot \Pr(\varepsilon_2 = 1.5) \cdot \Pr(\varepsilon_3 = 1.5) = (1/2)^3.$$

which shows that the ε's are not independent.

Q.E.D.

Although PROPOSITIONS 2 and 3 have shown that there is no necessary relationship between the IIA property and stochastic independence of the error terms, ε_i, we now show that IIA must hold if the error terms are perfectly dependent. By perfect dependence, we mean that if ε_k takes on a certain valve, the remaining error terms for the other alternatives must also take on that value, i.e. $\varepsilon_i = \varepsilon_j$ for any (i,j).

PROPOSITION 4 Perfect dependence among the error terms implies IIA.

PROOF: Consider an N alternative model such that $C = \{a_1, a_2, \ldots, a_N\}$ with ranking over mean utilities given by $\bar{U}_1 > \ldots > \bar{U}_i > \bar{U}_j > \ldots > \bar{U}_N$.

Then $P_i^C = \Pr(\bar{U}_i + \varepsilon_i > \bar{U}_1 + \varepsilon_1, \ldots, \bar{U}_i + \varepsilon_i > \bar{U}_j + \varepsilon_j, \ldots, \bar{U}_i + \varepsilon_i > \bar{U}_N + \varepsilon_N)$

But since $\varepsilon_1 = \ldots = \varepsilon_i = \varepsilon_j = \ldots = \varepsilon_n$, we have

$$P_i^C = \Pr(\bar{U}_i > \bar{U}_1, \ldots, \bar{U}_i > \bar{U}_j, \ldots, \bar{U}_i > \bar{U}_N)$$
$$= \Pr(\bar{U}_i > \bar{U}_j) = \Pr(\bar{U}_i + \varepsilon_i > \bar{U}_j + \varepsilon_j) = P_i^{\{i,j\}} .$$

Similarly, we have that $P_j^C = P_j^{\{i,j\}}$. Therefore we get $\dfrac{P_i^C}{P_j^C} = \dfrac{P_i^{\{i,j\}}}{P_j^{\{i,j\}}}$.

Q.E.D.

REFERENCES

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior." In Frontiers of Econometrics, edited by P. Zarembra. New York: Academic Press, 1974.

__________. "Econometric Models of Probabilistic Choice." In Structural Analysis of Discrete Data with Econometric Applications, edited by C. Manski and D. McFadden. Cambridge: MIT Press, 1981.

Maddala, G. S. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press, 1983.

Manski, C. and McFadden, D. Structural Analysis of Discrete Data with Econometric Applications. Cambridge: MIT Press, 1981.