ISSN 1032-3813
ISBN 0 7326 0762 0

MONASH UNIVERSITY

ANCORA IMPARO

AUSTRALIA

COMBINING CHOICE SET PARTITION TESTS FOR THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES PROPERTY: SIZE PROPERTIES IN THE FOUR ALTERNATIVES SETTING

Robert D Brooks
Tim R.L. Fry
Mark N. Harris

Working Paper 1/95
February 1995

DEPARTMENT OF ECONOMETRICS

COMBINING CHOICE SET PARTITION TESTS FOR THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES PROPERTY: SIZE PROPERTIES IN THE FOUR ALTERNATIVES SETTING

Robert D. Brooks*

Tim R.L. Fry**

Mark N. Harris**

* Department of Economics and Finance, Royal Melbourne Institute of Technology, GPO Box 2476V, Melbourne 3001 Victoria, Australia

** Department of Econometrics, Monash University, Clayton 3168 Victoria, Australia.

JEL Classification: C25

ABSTRACT

This paper conducts a Monte Carlo analysis of the size properties of combining choice set partition tests of the independence of irrelevant alternatives property in the Logit model in the four alternatives setting. Most of the tests have poor size properties. The exception is a version of the test proposed by Small and Hsiao (1985).

Correspondence to : Robert Brooks, Department of Economics and Finance, Royal Melbourne Institute of Technology, GPO Box 2476V, Melbourne Victoria 3001, Australia. Phone: 61-3-6603587, Fax: 61-3-6390584, E-mail: ROBB@OAK.BF.RMIT.EDU.AU

1. INTRODUCTION

The testing of many hypotheses in economics now requires the study of individual unit data and the application of qualitative response models. The simplest, and hence most popular, qualitative response model is the Logit model. While use of the Logit model provides many benefits in terms of computational tractability, it also imposes the cost of the independence of irrelevant alternatives (IIA) property. The IIA property implies that the choice between any two alternatives in the choice set depends solely upon the characteristics of the alternatives being compared, and not upon the characteristics (or indeed the existence) of any other alternatives in the choice set (see Debreu (1960)). This property may be unreasonable in modelling some choice problems and suggests a strong need to test the adequacy of the IIA property.

A large variety of tests exist for testing IIA. A popular class of tests involves partitioning the choice set into subsets and comparing parameter values and likelihood functions from the full choice set with those obtained from restricted subsets of the choice set. This class of tests has the advantage that no specific departure from the IIA property and the Logit model is required to be specified.

In this paper we analyse the size properties of choice set partition tests in the four alternatives setting. At present no studies have been published on the properties of choice set partition tests in the four alternative setting. In the simpler three alternatives setting Fry and Harris (1993, 1994) have found that the size and power properties of choice set partition tests are frequently dependent upon the chosen partition of the choice set. This sensitivity causes a problem for applied researchers in which partition to choose, a matter on which theory usually provides little guidance. In the three alternatives setting Brooks, Fry and Harris (1994) have analysed the size and power properties of combining all of the available partitions. In this paper we extend these results to size properties in the four alternatives setting.

The plan of this paper is as follows. In section two we briefly discuss the Logit model and the existing choice set partition tests. In section three we present results for the size properties of our combined tests. Section four contains some concluding remarks.

2. LOGIT MODELS AND TESTS FOR IIA

Following the previous literature (see, for example, Fry, Brooks, Comley and Zhang (1993)) we assume a random utility maximization model with (indirect) utility function given by:

$$U_{ij} = V_{ij}(\mathbf{Z}_{ij}, \mathbf{X}_i) + \epsilon_{ij}, \qquad i = 1, \ldots, n; j = 1, \ldots, J., \tag{1}$$

where U_{ij} is the utility individual i derives from choosing alternative j, which comprises of two components, V_{ij} and ϵ_{ij}. V_{ij} is a deterministic component which depends upon characteristics of the individual $\mathbf{X}_i$, and variables which vary across both individuals and alternatives, $\mathbf{Z}_{ij}$. ϵ_{ij} is a random component which represents unobservable factors. If we assume that the ϵ_{ij} are independent and identically distributed as Extreme Value, a Logit model arises, with selection probabilities given by:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{k=1}^{J} \exp(V_{ik})} \qquad i = 1, \ldots, n; j = 1, \ldots, J.$$

The IIA property of the Logit model states that the odds of choosing alternative j over alternative k $(k \neq j), P_{ij}/P_{ik}$, are independent of all other alternatives and of the number of alternatives in the choice set, that is

$$\frac{P_{ij}}{P_{ik}} = \frac{\exp(V_{ij})}{\exp(V_{ik})}.$$

If the Logit model provides an appropriate representation of discrete choice behaviour, considerable advantages are gained in model specification, estimation and forecasting. Thus if such models are to be used, it is important to test for the IIA property.

A popular class of tests for testing the IIA property is choice set partition tests. The idea behind choice set partition tests is simple: if the IIA property is valid, the model's structure (maximised likelihood functions and parameters) is unchanged when choice is analyzed on the full choice set, or conditional on a restricted subset of the full choice set. That is, when one or more of the alternatives is removed from the choice set. Therefore, such tests are based upon whether the realization of the estimated parameters, or maximized log-likelihoods, from the full choice set and a proper subset of the full choice set, are significantly different. The choice set partition tests we consider are the Hausman-McFadden (HM) test (Hausman and McFadden (1984)), the McFadden-Train-Tye (MTT) test (McFadden, Train and Tye (1981)), the Horowitz (H) test (Horowitz (1981)) and the Small-Hsiao (SH) test (Small and Hsiao (1985)). The first of these, the HM test, is a Hausman specification test, and the other three are all variants of a likelihood ratio test. The MTT test is biased asymptotically towards accepting the null hypothesis, while the H test is biased asymptotically towards rejecting the null hypothesis. The SH test being a combination of the MTT and H tests is free of any asymptotic bias.

For each of the choice set partition tests one faces the problem of how to partition the choice set. For instance, in testing for IIA in the Logit model with four different alternatives there are ten different versions of each of the above tests. This is because the choice set can be partitioned by deleting either the four sets of a single alternative, or the six pairs of two alternatives. In the three alternatives setting, Brooks, Fry and Harris (1994) consider all possible partitions, pool the information and then base the test on a combination of the information. Three natural, and simple, possibilities in this regard are the maximum, median and minimum. Rejecting the null hypothesis for the maximum corresponds to the case where only one single test rejects. Rejecting the null hypothesis for the median corresponds to the case where more than half of the individual tests reject. Rejecting the null for the minimum corresponds to the case where all of the individual tests reject. We therefore now consider the size properties of conducting the tests using these combinations. For notation purposes we denote the different combinations by using as subscripts *max* for the maximum, *med* for the median and *min* for the minimum.

3. SIZE PROPERTIES

In this section of the paper we consider the size properties of combining the tests. We do this by Monte Carlo methods using an experimental design similar to that of Fry and Harris (1993, 1994) and Brooks, Fry and Harris (1994). The model used has utility functions of the form,

$$U_{ij} = x_i'\beta_j + \epsilon_{ij},$$

where x_i is a vector of regressors including a constant and two regressors generated as independent drawings from the standard normal distribution, ϵ_{ij} are independent, identically distributed Extreme Value random disturbances and $\beta_1 = (\gamma, 0.5, -0.5)'$, $\beta_2 = (\gamma, -0.5, 0.5)'$, $\beta_3 = (\gamma, 0.25, -0.25)'$ and $\beta_4 = (0,0,0)'$ with γ being Euler's constant.

The results on size properties are presented in table 1. Empirical sizes are presented at three different nominal significance levels (1%, 5%, 10%) and for three different sample sizes ($n = 250, 500, 1000$). As a benchmark we compare our results to the results obtained in the three alternatives setting by Brooks, Fry and Harris (1994). For those tests where the empirical size is not significantly different from the nominal size the results are given in bold. An examination of the results in table 1 reveals the following conclusions. First, a number of the tests are badly undersized, with sizes well below the nominal level. All of the versions of the MTT test, the HM_{min} test, the SH_{min} test and the HM_{med} test are very undersized, never rejecting the null hypothesis. The results for all versions of MTT test and the HM_{min} test are identical to the three alternative results. For the SH_{min} and HM_{med} tests their performance has deteriorated relative to their performance in the three alternatives setting. The H_{min} test is also undersized although its performance is not as poor. For instance, at all sample sizes it has an empirical size between 1% and 2% when the nominal significance level is 10 %. This is a drop in performance when compared to results in the three alternative case. The HM_{max} test, SH_{max} test, H_{max} and H_{med} are all very oversized. The worst results are obtained for the H_{max} test. These results are comparable to the results in the three alternatives setting.

The best performing test is the SH_{med} test, which is the only test that produces empirical sizes which are not significantly different from the nominal level. This occurs for all significance levels when $n = 250$ and for the 1% significance level when $n = 500$. For the other cases at $n = 500$ and 1000 the test is undersized. At the 5% significance level the empirical size is 3%, while at the 10% significance level the empirical size is 7% to 8%. Overall the performance of the SH_{med} test is vastly superior to all other tests.

4. CONCLUSIONS

We have studied the size properties of combining choice set partition tests for the IIA property in the situation where the choice set contains four alternatives. In general, most of the tests have very poor size properties. The exception is the SH_{med} test which is vastly superior to all of the other tests. Therefore, if a test is required in such a situation we would suggest that researchers consider using the SH_{med} test. However, the four alternative case has not been considered before and ideally more work should be done before any strong conclusions are drawn. In particular, evidence on the power of these IIA tests should be considered. Unfortunately, such a study would be difficult to undertake as not only does the researcher have to consider the issues addressed here, but also whether the tests should be "size corrected", and more importantly, the choice of alternative, non-IIA model to use. This last issue is particularly difficult.

Table 1

Size Performance Of The Combined Tests

	n	=	250	n	=	500	n	=	1000
	1%	5%	10%	1%	5%	10%	1%	5%	10%
H_{max}	.55	.79	.87	.52	.73	.85	.44	.72	.82
H_{med}	.18	.42	.57	.17	.39	.55	.13	.32	.47
H_{min}	.00	.00	.01	.00	.00	.01	.00	.00	.01
HM_{max}	.20	.26	.31	.22	.31	.37	.26	.34	.41
HM_{med}	.00	.00	.00	.00	.00	.00	.00	.00	.00
HM_{min}	.00	.00	.00	.00	.00	.00	.00	.00	.00
MTT_{max}	.00	.00	.00	.00	.00	.00	.00	.00	.00
MTT_{med}	.00	.00	.00	.00	.00	.00	.00	.00	.00
MTT_{min}	.00	.00	.00	.00	.00	.00	.00	.00	.00
SH_{max}	.09	.28	.48	.07	.27	.44	.06	.21	.38
SH_{med}	.02	.06	.11	.01	.03	.08	.00	.03	.07
SH_{min}	.00	.00	.00	.00	.00	.00	.00	.00	.00

REFERENCES

Brooks, R.D., Fry, T.R.L. and M.N. Harris (1994), The Size and Power Properties of Combining Choice Set Partition Tests for the IIA Property in the Logit Model, mimeo, RMIT.

Debreu, G. (1960), Review of R.D. Luce Individual Choice Behaviour, American Economic Review 50, 186-188.

Fry, T.R.L., Brooks, R.D., Comley, B.R. and J. Zhang (1993), Economic Motivations for Limited Dependent and Qualitative Variable Models, Economic Record 69, 193-205.

Fry, T.R.L. and M.N. Harris (1993), A Monte Carlo Study of Tests for the Independence of Irrelevant Alternatives Property, Working Paper 8/93, Department of Econometrics, Monash University.

Fry, T.R.L. and M.N. Harris (1994), Testing for Independence of Irrelevant Alternatives: Some Empirical Results, Working Paper 2/94, Department of Econometrics, Monash University.

Hausman, J.A. and D. McFadden (1984), Specification Tests for the Multinomial Logit Model, Econometrica 52, 1219-1240.

Horowitz, J. (1981), Identification and Diagnosis of Specification Errors in the Multinomial Logit Model, Transportation Research 15B, 345-360.

McFadden, D., Train, K. and W. Tye (1981), An Application of Diagnostic Tests for the Independence of Irrelevant Alternatives Property of the Multinomial Logit Model, Transportation Research Record 637, 39-46.

Small, K.A. and C. Hsiao (1985), Multinomial Logit Specification Tests, International Economic Review 16, 471-486.