E C O N O M I C S B U L L E T I N

Toward a unified approach to testing for weak separability

Barry Jones
State University of New York at Binghamton

Thomas Elger
Lund University

David Edgerton
Lund University

Donald Dutkowsky
Syracuse University

Abstract

In this paper we propose a unified framework for testing weak separability. We present a new three−step procedure, which is a joint test of necessary and sufficient conditions that takes account of possible measurement error and incomplete adjustment. We illustrate the operational capability of the procedure with an empirical example. Our procedure works well in medium sized samples, but at the present time may not be practical for datasets with extremely large sample sizes. As computing technology continues to advance, however, high−powered methods like the one we propose should supplant testing approaches that were originally designed to circumvent computational limitations.

The authors wish to thank the Jan Wallander and Tom Hedelius Foundation for providing financial support for this and related research projects (Grants J98/14 and J03/19). Thomas Elger has also received support from Nottingham Trent University, where part of the research for this paper was carried out. Barry Jones was visiting Syracuse University when this paper was written. We thank Craig Lawrence for helpful advice regarding FFSQP, and James Swofford for providing us with data
Citation: Jones, Barry, Thomas Elger, David Edgerton, and Donald Dutkowsky, (2005) "Toward a unified approach to testing for weak separability." *Economics Bulletin*, Vol. 3, No. 20 pp. 1−7
Submitted: March 21, 2005. **Accepted:** May 1, 2005.
URL: http://www.economicsbulletin.com/2005/volume3/EB−05C10003A.pdf

1. Introduction

Weak separability is a fundamental condition for the existence of economic aggregates and the use of two-stage budgeting. It is therefore appropriate that this property should be tested in empirical work concerning aggregation and demand analysis. Two approaches have been used in practice: statistical tests within a parametrically formulated demand system, or deterministic checks using revealed preferences. The problem with the first approach is that it can crucially depend on the choice of specific functional form. This paper therefore concentrates on ways to improve the second approach, usually referred to as the nonparametric approach.

The nonparametric test used almost exclusively in practice is the sequential procedure based on Varian (1982, 1983) implemented in the computer program NONPAR, see for example Swofford and Whitney (1987), Manser and McDonald (1988), and Rickertsen (1998). The procedure suffers from a number of deficiencies, however. (i) Necessary and sufficient conditions are tested sequentially and not jointly. (ii) The procedure does not account for stochastic variation in the data, such as measurement error. (iii) When applying the procedure to time series data, it assumes that all adjustment of expenditure to optimal levels takes place within one period. For a variety of reasons, such as habit formation and adjustment costs, this may not be the case. These deficiencies imply that the sequential procedure is biased towards rejection of weak separability. Barnett and Choi (1989) showed in a Monte Carlo study that the rejection rate of the NONPAR test greatly exceeded the nominal significance level.

There have been a number of papers that have addressed these deficiencies separately. For example, Varian (1985) treats measurement error while Swofford and Whitney (1994) discuss incomplete adjustment. In both cases the proposed solutions strained the computational limitations of the time, and have not been implemented in empirical applications.

In this paper, we propose an approach that synthesizes the above developments into a unified framework for testing weak separability. We present a new three-step procedure, which is a joint test of necessary and sufficient conditions that takes account of the possibilities of both measurement error and incomplete adjustment. To demonstrate the operational capability of the procedure, we illustrate it with an empirical example.

2. A New Procedure for Testing Weak Separability

Let $x=(x_1,...,x_n)$ denote a vector of quantities for n goods, $p=(p_1,...,p_n)$ the corresponding vector of prices, and total expenditure $Y=px$. Let also $x=(y,z)$ and $p=(r,v)$ denote partitions of the quantity and price vectors into two groups, $\{y=(y_1,...,y_m);r=(r_1,...,r_m)\}$ and $\{z=(z_1,...,z_{n-m});v=(v_1,...,v_{n-m})\}$. Varian (1982), building on Afriat (1967), showed that a data set is consistent with the utility maximization hypothesis if and only if it is consistent with the generalized axiom of revealed preference (GARP), and gave a procedure for testing GARP.

2.1 The New Procedure

The procedure consists of three steps for testing the null that the y goods are weakly separable from the z goods, possibly with incomplete adjustment and measurement error.

I) Check if the data set $\{(r^i, y^i): i = 1,...,T\}$ satisfies deterministic GARP. If so then proceed directly to step III, otherwise go to step II.

II) Calculate the minimal perturbation of the quantity data that will satisfy GARP. If σ_*^2, the bound on the variance of the measurement error given in section 2.2, is smaller than our prior belief for the true variance, then use the measurement error corrected quantities $\hat{y}$ in step III. Otherwise reject the null.

III) Perform a (possibly modified) Swofford-Whitney test using $\{(r^i, y^i): i = 1,...,T\}$ or $\{(r^i, \hat{y}^i): i = 1,...,T\}$ depending on whether we come from step I or II. This will lead us to either accept weak separability with complete adjustment, accept weak separability with incomplete adjustment or reject the null, see section 2.3.

This procedure is not a complete solution to the problem, since steps II and III are performed sequentially and the possibility of measurement error in the *z* goods is not taken into account. It is, however, a significant generalization and improvement on the Varian procedure.[1]

2.2 Measurement Error Adjusted GARP Tests

Varian (1985) provided a test of optimizing behavior with measurement error in the quantity data. When applied to a GARP test of the *y*-goods, the null hypothesis is that the true quantities, $\xi^i = (\xi_1^i,...,\xi_m^i)$, $i = 1,...,T$, satisfy GARP. The true data is assumed to be related to the observed data by

$$\xi_k^i = y_k^i(1+\varepsilon_k^i), \tag{1}$$

where the measurement errors ε_k^i are i.i.d. with mean zero and variance σ^2.

Varian (1985) suggested constructing the minimal perturbation of the data that satisfies GARP, $\hat{y}^i = (\hat{y}_1^i,...,\hat{y}_m^i)$, $i = 1,...,T$. That is $\hat{y}$ minimizes the sum of squared proportional perturbations,

$$R(\zeta) = \sum_{i=1}^{T}\sum_{k=1}^{m}\left(\frac{\zeta_k^i - y_k^i}{y_k^i}\right)^2 \tag{2}$$

subject to the constraint that $\{(r^i, \zeta^i): i = 1,...,T\}$ satisfies the Afriat inequalities

$$V^i, \mu^i > 0 \quad \forall\, i. \tag{3}$$

$$V^i \le V^j + \mu^j r^j(\zeta^i - \zeta^j) \quad \forall\, i, j \tag{4}$$

and where the Afriat indices V^i and μ^i are treated as nuisance parameters. We refer to $\hat{y} = \arg\min(R(\zeta))$ as the *measurement error corrected* quantity data.

---

[1] We could also test whether the dataset $\{(p^i, x^i): i = 1,...,T\}$ satisfies GARP, with *y* possibly replaced by $\hat{y}$. This is a necessary condition for the null of weak separability with complete adjustment, allowing for possible measurement error. It need not hold under weak separability with incomplete adjustment, however.

If the measurement errors were somehow observed and normally distributed, the test statistic $R(\xi)/\sigma^2 = \sum\sum(\varepsilon_k^i)^2/\sigma^2$ would be χ^2_{Tm}. The null of weak separability would be rejected if the test statistic exceeded C_α, the critical value at the α significance level of the chi squared distribution. Since the true data is not observed, Varian (1985) proposed rejecting the null if $\hat{R}/\sigma^2 \geq C_\alpha$, where $\hat{R} = R(\hat{y})$ denotes the minimized objective function (2). This test procedure is conservative, since by definition of the minimal perturbation, $\hat{R} \leq \sum\sum(\varepsilon_k^i)^2$ under the null.

There are several limitations with this approach. To begin with, the variance of the measurement error is unknown. Following Varian we can calculate $\sigma_*^2 = \hat{R}/C_\alpha$, which can be interpreted as the largest measurement error variance for which we would reject the null. This value can be compared to our prior belief concerning the error variance, σ^2_{prior}. If σ_*^2 is smaller than σ^2_{prior}, then we might not wish to reject the null even if deterministic GARP does not hold. Varian (1985, p. 449) suggested using a parametric fit of the data to arrive at an estimate of the error variance, which could be used in place of σ^2_{prior}. Another limitation is that the computational burden of minimizing (2) is considerable. As far as we are aware, the authors of this paper are the first to implement it empirically for GARP.[2]

2.3 Swofford and Whitney's Test of Weak Separability with Incomplete Adjustment

Swofford and Whitney (1994) assumed that expenditure on the separable group might not adjust completely within one period, and thus be suboptimal. They showed that the necessary and sufficient conditions for weak separability of *y* from *z*, with incomplete adjustment of expenditure on the *y* goods, are that there exist numbers U^i, τ^i, V^i, μ^i and ϕ^i such that

$$U^i > 0,\ \tau^i > 0,\ V^i > 0,\ \mu^i > 0,\ \phi^i > 0 \quad \forall i, \tag{5}$$

$$V^i \leq V^j + \mu^j r^j (y^i - y^j) \quad \forall i, j, \text{ and} \tag{6}$$

$$U^i \leq U^j + \tau^j v^j (z^i - z^j) + \phi^j (V^i - V^j) \quad \forall i, j, \tag{7}$$

where $\theta = \phi\mu - \tau$ is the shadow price of the expenditure constraints on the *y* goods. If θ is negative (positive), then expenditure on *y* is greater (less) than the optimal expenditure. They suggest that the inequalities can be checked by minimizing $F = \sum(\theta^i)^2$ subject to (5) - (7). If a feasible solution to the minimization problem is found then *y* is weakly separable, with complete adjustment if $F = 0$ and with incomplete adjustment if $F > 0$. If there is no feasible solution then *y* is not weakly separable, even with incomplete adjustment. When using measurement error corrected quantities, $\hat{y}$ replaces *y* in (6).

We propose a modification of the Swofford and Whitney (1994) test, which concentrates on an alternative measure of incomplete adjustment, $\psi = \theta/\tau = \phi\mu/\tau - 1$. According to Swofford and Whitney (1994, p 244), ψ represents the increment of utility from

[2] Varian (1985), however, implemented the procedure for firm data with constraints based on the weak axiom of cost minimization (WACM). The constraints for WACM are linear, which reduces the computational burden.

spending an additional dollar on y relative to the marginal utility of an additional dollar of total expenditure. Since it is a ratio of marginal utilities, it is a cardinal measure and thus easier to interpret than θ, which is merely an ordinal measure.[3] We therefore propose that weak separability is checked by minimizing, w.r.t. U^i, τ^i, V^i, μ^i and ψ^i, the objective function

$$G = \sum_{i=1}^{T} (\psi^i)^2 , \tag{8}$$

subject to (5) - (7) with ϕ replaced by $(1+\psi)\tau / \mu$. Note that there is complete adjustment if and only if $\psi^i = 0 \;\; \forall i$. Useful measures of incomplete adjustment are therefore the maximum and average of the absolute values,

$$\psi^{max} = \max_{i=1,\dots T} |\psi^i| \quad \text{and} \quad \bar{\psi} = \frac{1}{T}\sum_{i=1}^{T} |\psi^i| . \tag{9}$$

Large values of these measures may indicate implausible levels of incomplete adjustment.[4]

3. An Empirical Illustration

We programmed our procedure in FORTRAN 95 using the commercial solver FFSQP, which is documented in Zhou *et al.* (1997). The subroutines are available from the authors upon request.[5] Although steps II and III are quite computationally burdensome, we were able to run all tests on standard PC's.

We used our procedure on the dataset given in Swofford and Whitney (1994), who tested weak separability on the following two utility structures [6]

$$U1 = U(NDUR, SER, LEIS, SD_1, SD_2, SD_3, SD_4, STD_1, STD_2, STD_3, V(OM1, OCD))$$
$$U2 = U(NDUR, SER, LEIS, STD_1, STD_2, STD_3, V(OM1, OCD, SD_1, SD_2, SD_3, SD_4))$$

They found it impossible to analyze the complete sample of 62 observations, which we call S, so they divided it into two overlapping periods of 40 observations; S1 and S2. They found no feasible solution in either sub-sample for *U1*, but found weak separability with incomplete adjustment for both sub-samples for *U2*.

[3] Equations (10a) and (10b) in Swofford and Whitney (1994), show that ψ can also be interpreted as the proportional deviation of the marginal rate of substitution from the relative price ratio when comparing a y good and a z good.

[4] The objective function is basically an artifact used for finding a solution to (5) - (7), and the choice of function is thus somewhat arbitrary. $\bar{\psi}$ can of course be calculated even if we minimize the original objective function, but in this case we might well calculate a higher degree of incomplete adjustment than necessary.

[5] AEM Design made FFSQP available to us (as academics) free of charge, but we cannot distribute it to third parties. We tried to use the IMSL non-linear programming subroutines, but they were unable to handle problems with large numbers of constraints. Similar problems with these IMSL subroutines have been reported in the computer science literature.

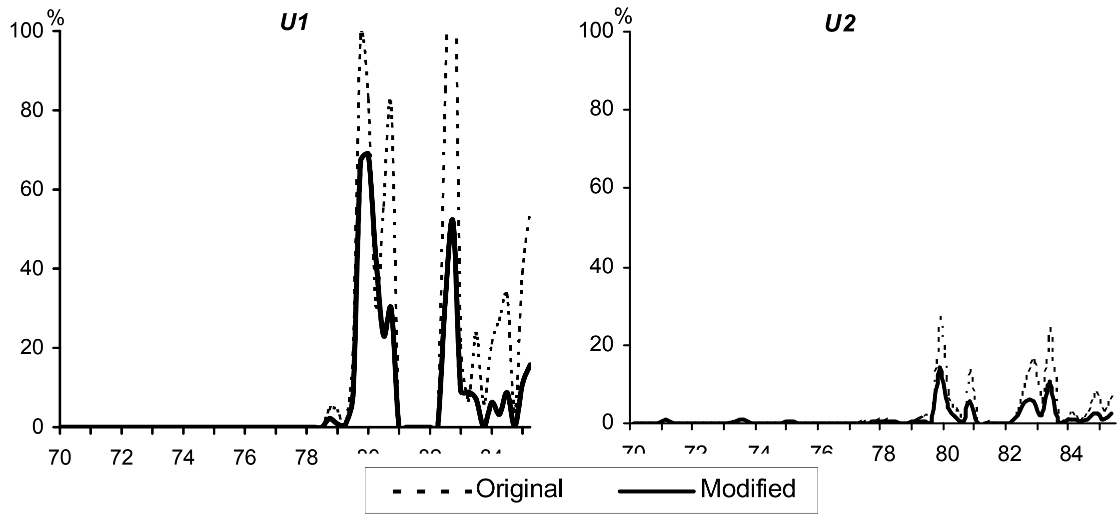
[6] The notation stands for non-durables (NDUR), services (SER), leisure (LEIS), currency and demand deposits (OM1), other checkable deposits (OCD), four types of savings deposits (SD_i) and three types of small denomination time deposits (STD_i).

When we applied our three-step procedure, the data satisfied GARP in step I in all cases, and we could therefore proceed directly to step III without any measurement error correction. We ran step III using both the original Swofford-Whitney test and our modification of the test for comparison. Feasible solutions were obtained in all cases, indicating that neither structure can be rejected in any sample with incomplete adjustment.[7]

Table I: Measures of Incomplete Adjustment for Structures U1 and U2

Structure, Sample	Swofford-Whitney Test		Modified Test	
	Average (%)	Maximum (%)	Average (%)	Maximum (%)
U1,S1	0.000	0.000	0.013	0.169
U2,S1	0.294	2.334	0.004	0.044
U1,S2	22.270	152.430	7.141	85.167
U2,S2	0.639	9.234	0.670	8.585
U1,S	13.290	149.266	6.405	68.912
U2,S	2.466	27.442	1.104	14.174

Figure 1: $|\psi^i|$ from the Swofford-Whitney Test for U1 and U2 in the Whole Sample



In Table I we report the average and maximum amounts of incomplete adjustment, $\bar{\psi}$ and ψ^{max}, as percentages. The time profile $|\psi^i|$ can also help point to particular events that are associated with violations of weak separability with complete adjustment, and an example is given in Figure 1. The incomplete adjustment from the modified test is very small for both structures in S1. The maximum amount of incomplete adjustment seems excessive for *U1* in S2 (85% for the modified test). It is much lower for *U2* but is still over 8% for S2 and over

[7] Following Swofford and Whitney (1994), we determined if the solution was feasible by looking at the sum of constraint violations. These were of the order 10-15 or less in all cases. We obtained almost identical results using several reasonable methods to obtain starting values. The results presented in Table I mainly use the LP algorithm from Fleissig and Whitney (2003) to obtain start values for V and μ and Divisia indices to obtain start values for U and τ.

14% using the full sample. Relative to the modified test, the Swofford-Whitney test tends to overstate the degree of incomplete adjustment, but tends to converge more quickly.

In theory, complete adjustment is accepted if the objective function in Step III can be minimized to exactly zero. In practice, due to the convergence properties of any numerical optimization procedure, this condition can only be approximately satisfied. A practical solution is to check the dataset $\{((v^i, 1/\mu^i), (z^i, V^i)) : i = 1, ..., T\}$ for GARP, using the indices derived from the optimization in step III. If it passes, then weak separability with complete adjustment is accepted (Varian, 1983). Using this check, both *U1* and *U2* hold with complete adjustment in S1. Neither holds with complete adjustment for either S2 or S. If a dataset fails the check, we calculate the number of violations. Based upon the Swofford-Whitney test, the number of violations is 7 for U1 (S2 and S) and 2 for U2 (S2 and S).

Our findings for *U2* are very similar to Swofford and Whitney (1994), but are qualitatively different for *U1*. These differences are most likely explained by our use of an advanced commercial solver.

In order to illustrate measurement error adjustment we use the following utility structure, (STD_3 are deposits at credit unions, whose quantities are minimal)

$$U3 = U(NDUR, SER, LEIS, OM1, OCD, STD_3, V(SD_1, SD_2, SD_3, SD_4, STD_1, STD_2)).$$

We find stage I violations of GARP for both sub-samples using this structure, and it is therefore necessary to correct the data for measurement error. The results are given in the following table.

Table II: Results for Structure U3

Sample	Stage I GARP Violations	Measurement Error σ_* (99% level)	Measures of Incomplete Adjustment			
			Swofford-Whitney Test		Modified Test	
			Average (%)	Max (%)	Average (%)	Max (%)
S1	5	0.0040	43.961	99.991	34.733	99.771
S2	2	0.0007	19.607	83.994	8.333	60.131

The data would have to be measured with a standard error of less than 0.40% in S1 and 0.07% in S2 in order to reject the null hypothesis of GARP at the very conservative 99% significance level. The figures are 0.33% and 0.06% using a more standard 5% significance level. If one believes that the true measurement error variance is larger than these figures, it would be reasonable to test for weak separability using the measurement error corrected data. We ran the test in both sub-samples to illustrate the procedure. We found feasible solutions in both cases, but the levels of incomplete adjustment in Table II seem excessive. Based upon the Swofford-Whitney test, the number of violations with complete adjustment is 424 for S1 and 178 for S2.

4. Conclusions

We propose a unified approach to testing weak separability, which synthesizes a joint test of necessary and sufficient conditions based on Swofford and Whitney (1994) with the measurement error approach proposed by Varian (1985).

The commonly used sequential procedure is based on an algorithm to construct Afriat indices for the group of goods or assets being tested. NONPAR uses one particular algorithm, while Fleissig and Whitney (2003) have recently proposed an alternative algorithm. The test based on the alternative algorithm performed well in their simulation study. However, since their "LP test" only uses sufficient conditions for weak separability, the intrinsic bias towards rejection remains. It also cannot account for incomplete adjustment or measurement error.

The test procedure proposed in this paper is more general than the LP test but is also more computationally burdensome. We were, however, able to implement it on standard PC's with the help of the commercial solver, FFSQP. Based upon computational considerations, we recommend testing for weak separability, with data either corrected or uncorrected for measurement error, using the LP-test first. If weak separability with complete adjustment is rejected with the LP-test then continue with the original Swofford and Whitney test. If the results from this test marginally reject, with either complete or incomplete adjustment, then continue with the modified test.

At the present time our procedure works well in medium sized samples, but may not be practical for datasets with extremely large sample sizes. As computing technology continues to advance, however, high-powered methods like the one we propose should supplant testing approaches that were originally designed to circumvent computational limitations.

References

Afriat, S. N. (1967) "The construction of a utility function from expenditure data" *International Economic Review* **8**, 67-77.

Barnett, W. A., and S. Choi (1989) "A Monte Carlo study of tests of blockwise weak separability" *Journal of Business and Economic Statistics* **7**, 367-377.

Fleissig, A. R, and G. A. Whitney (2003) "A new PC-based test for Varian's weak separability conditions" *Journal of Business and Economic Statistics* **21**, 133-144.

Manser, M. E., and R. J. McDonald (1988) "An analysis of substitution bias in measuring inflation, 1959-85" *Econometrica* **56**, 909-930.

Rickertsen, K. (1998) "The demand for food and beverages in Norway" *Agricultural Economics* **18,** 89-100.

Swofford J. L., and G. A. Whitney (1987) "Nonparametric tests of utility maximization and weak separability for consumption, leisure, and money" *Review of Economics and Statistics* **69**, 458-464.

Swofford J. L., and G. A. Whitney (1994) "A revealed preference test for weakly separable utility maximization with incomplete adjustment" *Journal of Econometrics* **60**, 235-249.

Varian, H. (1982) "The nonparametric approach to demand analysis" *Econometrica* **50**, 945-974.

Varian, H. (1983) "Non-parametric tests of consumer behaviour" *Review of Economic Studies* **50**, 99-110.

Varian, H. (1985) "Non-parametric analysis of optimizing behavior with measurement error" *Journal of Econometrics* **30**, 445-458.

Zhou J. L., A. L. Tits, and C. T. Lawrence (1997) "User's guide for FFSQP version 3: A Fortran code for solving optimization programs, possibly minimax, with general inequality constraints and linear equality constraints, generating feasible iterates" Technical Report

SRC-TR-92-107r5, Institute for Systems Research, University of Maryland, College Park, MD 20742. Online at: http://64.238.116.66/aemdesign/FSQPframe.htm