Projektbereich A
Discussion Paper No. A–528

A Note on the Power of Revealed Preference Tests with Afriat Inefficiency

by

Reinhard Sippel

July 1996

Author's address:

Wirtschaftstheoretische Abteilung II
Rheinische Friedrich-Wilhelms-Universität Bonn
Adenauerallee 24–42
D–53113 Bonn
Fon +49–228–73 92 45
Fax +49–228–73 79 40
E-mail: rsippel@econ2.uni-bonn.de

A Note on the Power of Revealed Preference Tests with Afriat Inefficiency

by

Reinhard Sippel*

Abstract

We show that revealed preference tests, which allow a small degree of inefficiency in consumers' choices, have no power against the alternative of purely random demand.

JEL-Classification: C12, C15, D12

Keywords: revealed preference tests, Afriat efficiency index, nearly optimizing behaviour, power analysis

1 Introduction

Revealed preference tests along the lines of Varian (1982) have become increasingly popular in empirical demand analysis because they are easily implemented and free of misspecification errors. Since no parameters are estimated they are also known as "nonparametric tests" of consumer behaviour, though this notion is somewhat misleading. Indeed, in contrast to "nonparametric tests" in the usual sense of statistical hypothesis testing, revealed preference tests are deterministic in nature. A given set of data satisfies the revealed preference inequalities, or it does not. In the latter case, the null hypothesis of utility maximizing behaviour is plainly rejected, since no allowance is made for possible stochastic influences on the data generating process.

This has been recognized as a disadvantage of the revealed preference approach. Consequently, a number of attempts have been made to incorporate measurement error or "nearly optimizing" behaviour into the analysis. One approach, which seems attractive both in terms of its computational simplicity and its intuitive plausibility, is to introduce Afriat inefficiency (cf. Afriat 1973, 1990; Varian 1990, 1993). Unfortunately, as will be demonstrated in this note, allowing Afriat inefficiency leads to

*Thanks are due to Jim Cox for sending me his compilation of the Battalio et al. (1973) data. Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged.

a severe loss of power of the test.[1] The null hypothesis of optimizing behaviour up to an inefficiency of, say, 5%, will typically not be rejected by purely random data. Note that it has been recognized earlier (cf. Varian, 1982, p. 965; Bronars, 1987; Russell, 1992) that even the standard revealed preference test (i.e. with full efficiency) might lack power. It is therefore definitely not advisable to further reduce the power by allowing Afriat inefficiency.

The next section presents the concept of Afriat (in-)efficiency and some of its applications in the analysis of consumer demand. In section 3 we extend Bronars' (1987) simulation procedure in order to assess the power-reducing consequences of introducing inefficiency in revealed preference tests. Section 4 concludes.

2 Revealed preference analysis and Afriat inefficiency

The revealed preference approach to demand analysis is only briefly sketched here; the reader is referred to, e.g., Varian (1982) or Varian (1993) for details.

Let us assume, we have T observations of price vectors p^t and demand vectors x^t, $t = 1, \ldots, T$.[2] We define the (direct) revealed preference relation R^D by $x^t R^D x^s \iff p^t x^t \geq p^t x^s$, and let R be the transitive closure of R^D. There is a well-behaved utility function that rationalizes the demand data if and only if for all $s, t = 1, \ldots, T$, the Generalized Axiom of Revealed Preference (GARP) holds:

$$x^t R x^s \Rightarrow p^s x^s \leq p^s x^t$$

If for some pair of observations s and t GARP does not hold, the utility maximization hypothesis must be rejected, no matter how close to optimizing behaviour the data might be. Since perfect rationality is perhaps too strong a requirement, and the loss associated with a suboptimal demand might be insignificant in economic terms (Varian, 1990), Afriat (1973) introduced a measure of efficiency into revealed preference analysis. The rationale behind this measure, a number between 0 and 1 called the Afriat efficiency index and denoted by e, is that the consumer is allowed to waste a fraction $1 - e$ of his expenditure. This is achieved by relaxing the preference relation to $x^t R^D(e) x^s \iff e \cdot p^t x^t \geq p^t x^s$, and accordingly defining a weaker consistency axiom, GARP(e):[3]

$$x^t R(e) x^s \Rightarrow e \cdot p^s x^s \leq p^s x^t$$

[1] The fact that Afriat inefficiency reduces the power of the test has been noted earlier by Famulari (1995). (I am grateful to Aurelio Mattei for pointing out this paper to me.) However, she fails to give a quantification of the effect. In the present paper, we show how serious the problem really is.

[2] p^t and x^t are vectors in $\mathbb{R}^\ell$, where ℓ is the number of commodities.

[3] Again, $R(e)$ denotes the transitive closure of $R^D(e)$.

GARP(e) is violated only if choices are very inefficient, in the sense that a considerable fraction of the expenditure could have been saved by choosing another bundle which has been revealed preferred to the one actually chosen. For a given data set, we can determine a *critical cost efficiency* e^* such that GARP(e) is satisfied for all $e \leq e^*$ but is violated for all $e > e^*$.[4] If e^* is close to 1, say 95%, choices are commonly said to be very efficient, and the remaining 5% are readily explained by measurement error or "nearly optimizing" behaviour.

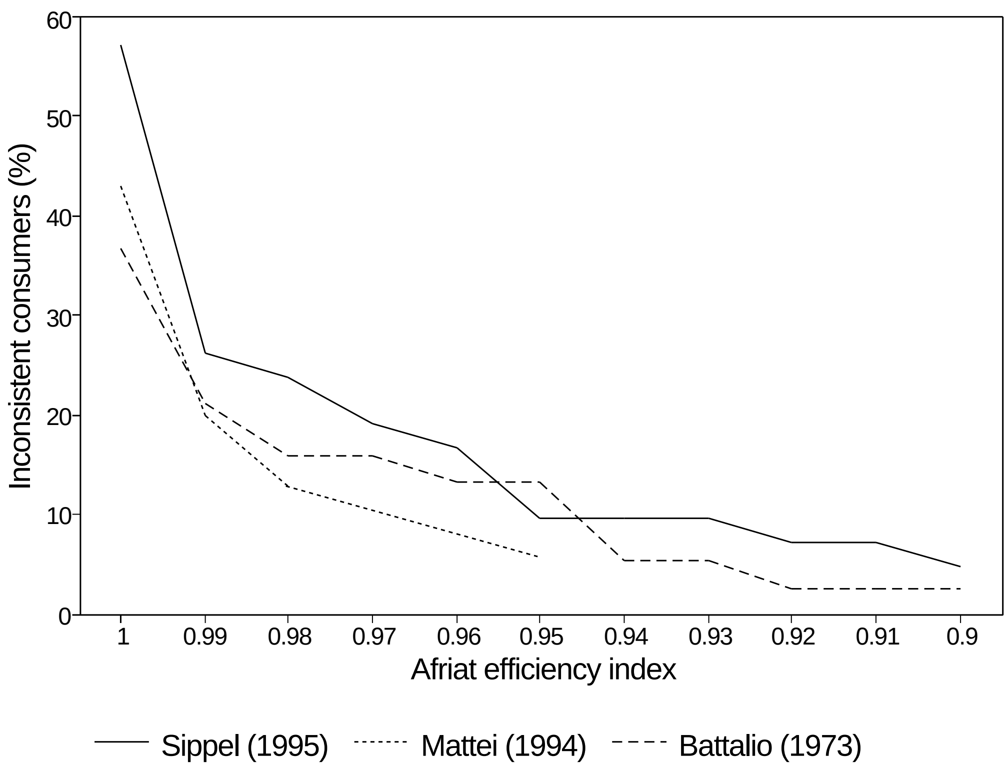

Fig. 1:
Percentage of inconsistent consumers depending on the Afriat efficiency index.

Fig. 1 summarizes the results from a number of empirical studies. It can be seen that most consumers' choices are efficient at least at the 90% level. For example, Mattei (1994) finds that with an Afriat efficiency index of 95% only 5.9% of the households remain inconsistent. Varian (1993) analyses the token economy data of

[4]Varian (1990) extends this idea by defining a different e^t for each observation $t = 1, \ldots, T$. For the purpose of the present paper it suffices to use the single index e^* since we are not interested in which particular observation causes the revealed preference violation. Rather, the point of the paper is to show that GARP(e^*) is typically *not* violated by random data for surprisingly high values of e^*. If, say, GARP(0.95) is satisfied, then also is GARP(e^t) as defined by Varian (1990, 1993), where $e^t = 0.95 \forall t$, and it will typically hold even for some $e^t > 0.95$.

Battalio et al. (1973). Since "98% of the choices were at least 95% efficient" (p. 8), he finds the subject behaviour "very close to satisfying GARP" (p. 9). Cox (1995) extends the original Battalio et al. (1973) data by taking into account labour supply and savings decisions. He reaches the same conclusion (p. 21): "(D)ata for 37 out of the 38 subjects pass a nonparametric 10% test for consistency with the utility hypothesis. Therefore, the choices (...) are highly consistent with all of the implications of the utility hypothesis." In these conclusions, it is apparently overlooked that at an efficiency of 90% the utility maximization hypothesis has no more implications left that distinguish it from most alternative hypotheses including purely random behaviour.

3 The power of Afriat inefficient revealed preference tests

It is immediately obvious from the definition of $R^D(e)$ that the number of preferences revealed by the data decreases monotonically with the inefficiency allowed.[5] But revealed preference theory only has implications if there are any preferences revealed. Therefore, one might intuitively feel a little uncomfortable with a procedure that deprives the theory of its single empirical content. On the other hand, some of the preferences "revealed" by the data might actually not be true preferences. For one thing, the consumer might not be so sure about whether to prefer one or the other bundle. Additionally, in the presence of measurement error regarding the price vector, some bundle "revealed worse" than another bundle might in fact not have been available at all. From this perspective, loosing some preferences by introducing a slight inefficiency does not seem to be a matter to worry about.

However, simulations run in order to estimate the power of these tests show that the effect gives cause for concern. The need for simulations stems from the fact that for revealed preference tests no simple formula exists for the probability of rejecting the null given that it is not true. Bronars (1987) suggested the following Monte-Carlo technique instead: Create purely random demand data as a rather extreme alternative to the utility maximization hypothesis. See whether this random demand satisfies GARP. Repeat this procedure sufficiently often (Bronars ran 200 simulations) and take as an approximation of the true power the percentage of times GARP was violated. Bronars applied this technique to U.S. aggregate data and found the power of the test to be surprisingly high.[6]

We extended this procedure in that we applied it to three different data sets, and

[5] At $e = 0$, no preferences at all are revealed, and GARP(0) is hence vacuously satisfied.

[6] This is in contrast to what one would expect (see also Varian, 1982, p. 965) since the number of budget set intersections in aggregate data is typically quite small. Without any budget set intersection GARP can never be violated. Note that Bronars (1987) increased the number of budget set intersections by using per capita rather than the original aggregate data.

for each data set, we estimated the power for 11 different values of the efficiency index e, namely 1, 0.99, ..., 0.91, 0.9. The three data sets used were our own (Sippel, 1995) experimental data, the Battalio et al. (1973) token economy data with 5 commodities as analysed by Cox (1995), and U.S. aggregate consumption data found in Varian (1990), Table 3. For each data set, we ran 1000 simulations, each time creating random budget shares with an algorithm equivalent to Bronars' algorithm 2 and applying these to the respective price and expenditure data.[7] Fig. 2 shows our estimations of the power (the number of times GARP was violated divided by 1000) plotted against the 11 different values of e.

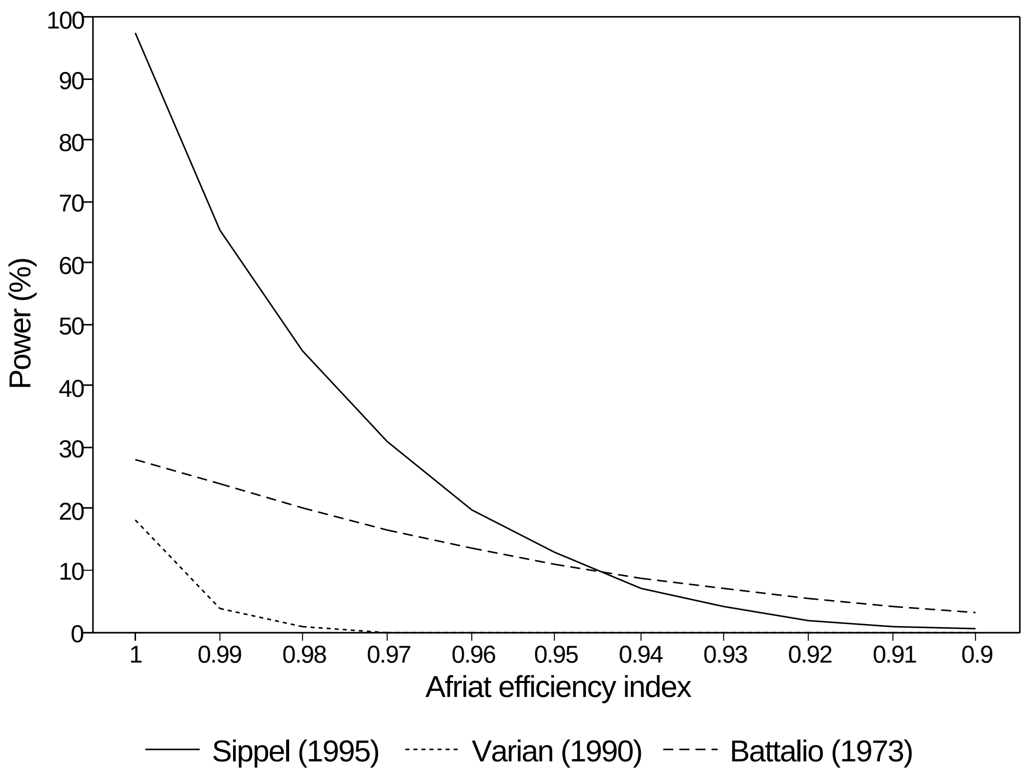

Fig. 2:
Power estimates for revealed preference tests with Afriat inefficiency.

Note the surprisingly fast loss of power for seemingly small inefficiencies. Note, in particular, that at the "prominent" efficiency of 95% almost 90% of the random data pass the consistency test. At $e = 90\%$, less than 3% of the random demand is still inconsistent. This is true for all data sets, including the experimental ones with considerable price variation and, hence, a large number of budget set intersections.[8]

[7]Since expenditure was different from subject to subject in the Battalio et al. (1973) experiment, we calculated the power for each subject separately and then took the average over all 38 subjects.

[8]The low power of the Varian (1990) aggregate annual data set could be expected since there is

Note also how similar the violation rates are, for $e \leq 0.95$, between the true demand data of Fig. 1 and the artificial random demand data of Fig. 2. In view of this one might jump to the conclusion that irrational randomizers are doing almost just as well as conscious human optimizers. Such a conclusion is, of course, completely unwarranted; yet, it shows how misleading the notion of "efficiency" is in the present context.

4 Conclusion

Consumer demand data usually fail to satisfy the strict revealed preference axioms. However, most of the violations disappear if we require less than full efficiency e because "tests of the data are more relaxed for smaller values of e" (Afriat 1987, p. 174). We have shown that this relaxation is so profound that purely random demand typically also passes a test for, say, 95% "efficiency." Since "(o)ne should be wary of tests that have low power against the rather naive alternative of random behavior" (Bronars, 1987, p. 697), this particular approach to accounting for "nearly optimizing" behaviour in revealed preference analysis should thus better be avoided.

It follows, then, that truly statistical tests such as the ones developed by Epstein and Yatchew (1985) and Varian (1985) should be employed in nonparametric demand analysis.[9] While these tests may be computationally more burdensome, the additional resources required seem to be well invested.

only a single budget set intersection (between years 1948 and 1949) in the data covering 41 years from 1947 to 1987.

[9]Note, on the other hand, that Tsur's (1989) statistical test is also based on the Afriat efficiency index and, hence, can be expected to suffer from similar power problems as described above.

References

Afriat, S. N. (1973), On a System of Inequalities in Demand Analysis: An Extension of the Classical Method, *International Economic Review* **14**, 460–472.

______ **(1987),** *Logic of Choice and Economic Theory*, Oxford: Clarendon Press.

Battalio, R. C., Kagel, J. C., Winkler, R. C., Fisher, E. B., Basmann, R. L. and Krasner, L. (1973), A Test of Consumer Demand Theory Using Observations of Individual Consumer Purchases, *Western Economic Journal* **11**, 411–428.

Bronars, S. G. (1987), The Power of Nonparametric Tests of Preference Maximization, *Econometrica* **55**, 693–698.

Cox, J. C. (1995), On Testing the Utility Hypothesis, mimeo, University of Arizona, Tucson.

Epstein, L. G., and A. J. Yatchew (1985), Non-Parametric Hypothesis Testing Procedures and Applications to Demand Analysis, *Journal of Econometrics* **30**, 149–169.

Famulari, M. (1995), A Household-based, Nonparametric Test of Demand Theory, *Review of Economics and Statistics* **77**, 372–382.

Mattei, A. (1994), La cohérence des choix des consommateurs suisses, *Revue suisse d'Economie politique et de Statistique* **130**, 3–20. English version: A Non-Parametric Test of Consumer Behavior Using Panel Data, Université de Lausanne, deep, Discussion Paper no. 9112.

Russell, R. R. (1992), Remarks on the Power of Non-Parametric Tests of Consumer-Theory Hypotheses, in *Aggregation, Consumption and Trade* (ed. L. Phlips and L. D. Taylor), Dordrecht: Kluwer Academic Publishers, 121–135.

Sippel, R. (1995), An Experiment on the Pure Theory of Consumer's Behaviour, *SFB 303, Universität Bonn, Discussion Paper No. B–274.*

Tsur, Y. (1989), On Testing for Revealed Preference Conditions, *Economics Letters* **31**, 359–362.

Varian, H. R. (1982), The Nonparametric Approach to Demand Analysis, *Econometrica* **50**, 945–973.

______ **(1985),** Non-Parametric Analysis of Optimizing Behavior with Measurement Error, *Journal of Econometrics* **30**, 445–458.

______ **(1990),** Goodness-of-Fit in Optimizing Models, *Journal of Econometrics* **46**, 125–140.

______ **(1993),** Goodness-of-Fit for Revealed Preference Tests, mimeo, University of Michigan.