Staff Papers Series

Staff Paper P91-7

February 1990

Testing the Significance of Deviations from Rational Behavior

Yacov Tsur

Department of Agricultural and Applied Economics

University of Minnesota
Institute of Agriculture, Forestry and Home Economics
St. Paul, Minnesota 55108

Staff Paper P91-7 February 1990

TESTING THE SIGNIFICANCE OF DEVIATIONS FROM RATIONAL BEHAVIOR

Yacov Tsur

Staff Papers are published by the Department of Agricultural and Applied Economics without formal review.

Testing the Significance of Deviations from Rational Behavior

Yacov Tsur*

Abstract

We propose procedures for testing statistically the significance of violations of nonparametric tests of optimization axioms when observed behavior is measured with error. The tests are robust against parametric specification of the error distribution, thus are nonparametric in both the statistical and economic senses, and are readily implemented numerically. An illustration with demand data is presented.

*Department of Agricultural and Applied Economics, University of Minnesota, 1994 Buford Avenue, St. Paul, MN 55108.

The helpful comments of Ted Graham-Tomasi and Lung-Fei Lee are gratefully acknowledged.

Testing the Significance of Deviations from Rational Behavior

Yacov Tsur

1. Introduction

Since their development by Afriat (1967, 1973), Hanoch and Rothchild (1972) and Varian (1982, 1984), among others, nonparametric analyses of consumer demand and producer input/output decisions continue to find useful applications in a wide variety of areas. A crucial step in the analysis entails testing whether an observed pattern of behavior is rational, in that it is the outcome of consumers maximizing preferences or producers maximizing profit (and minimizing cost). The term "nonparametric" signifies, in this context, that no parametric structure is a priori imposed on preferences or on the production technology. If the data pass the test, so that observed behavior is rational, useful information on preferences or production technologies can be recovered. A violation of the test indicates non-optimal behavior of decision makers or that structural changes in preferences or production technologies are present (or a combination of the two).

Data on observed decisions, however, generally are measured with errors. A violation of an optimization axiom by the data, thus, raises the question of how large the deviation is from optimal behavior and whether it is plausible that the true (unobserved) behavior is rational. Clearly, a measure of the significance of the deviation is required in order to answer this question.

Numerous authors have proposed statistical procedures which provide precise meaning to the adjective "significant" (see Varian, 1985, and the references he cites). These procedures are based on a particular parametric specification of the error distribution, e.g., normality. The assumption of normality (or any other specification) is bothersome since it is inconsistent, in spirit, with nonparametric analyses; as argued by Hanoch and Rothchild (1972, footnote, p. 264): "It does not seem sensible to make no assumption about the production process and then blithely impose a particular specification on the error process."

Noting this limitation, Epstein and Yatchew (1985) develop a framework of nonparametric (in the statistical sense) hypothesis testing and describe how it can be applied to test violation of optimization axioms. In particular, they show that the test proposed by Varian (1985) can be given an asymptotically nonparametric interpretation. Empirical implementation of this test, however, is computationally quite involved and rapidly looses tractability as the sample size increases; thus, applications are limited. Tsur (1989), in an attempt to mitigate the tractability problem, proposed a test which is simple and fast, and therefore can be applied with large data sets. Tsur's test, however, maintains an assumption of normally distributed errors, and hence suffers from the deficiency mentioned above.

In this work we develop a framework for testing the significance of violation of optimization axioms when the data contain measurement errors. The tests are robust against the specification of the error distribution and possess desirable computational properties. The analysis is similar in approach to that of Epstein and Yatchew (1985) and builds on ideas developed in Tsur (1989).

The hypothesis testing framework is developed in the context of consumer demand decisions. It begins, in Section 2, with a description of the data process and a summary of the relevant revealed preference concepts. The test procedure is developed in Section 3, which also presents comparison with a parametric test that maintains normally distributed errors. Section 4 discusses implementation issues. The production case is covered in Section 5, where hypothesis testing of deviation from profit maximization and cost minimization is described. A numerical illustration, presented in Section 6, applies the procedure to test deviations from optimal consumption decisions.

Concluding comments are offered in the closing section.

2. The data process and revealed preference concepts

Let X^i be k-vectors of observed quantities demanded at prices P^i, $i=1,2,..n$. The observations contain errors; the corresponding true, unobserved quantities and prices are denoted by X^{*i} and P^{*i}. Let $C(i,j) = P^i \cdot X^j$ represent the expenditure of consuming X^j at prices P^j, and $C^*(i,j) = P^{*i} \cdot X^{*j}$ be the corresponding true (unobserved) expenditure. The observed and true expenditures are related according to $C(i,j) = E(i,j)C^*(i,j)$, where the $E(i,j)$'s represent measurement errors associated with the expenditure data. Let $c_j = \log C(j,j)$, $c \equiv (c_1,c_2,..,c_n)$, $c_j^* = \log C^*(j,j)$, $c^* \equiv (c_1^*,c_2^*,..,c_n^*)$ and $\varepsilon_j = \log E(j,j)$. Then, for $j=i$:

$$c_j = c_j^* + \varepsilon_j, \quad j=1,2,..,n. \tag{1}$$

The test developed below involves (moments of) the errors on actual expenditures, i.e., the ε_j's. Without imposing a particular parametric form on their distribution, we require:

Assumption 1: $\varepsilon_1,\varepsilon_2,..,\varepsilon_n$ are iid with a zero mean and a finite fourth moment.

A zero mean is not essential; if $E(\varepsilon_j) \neq 0$ then Eq. (1) can be redefined as $\tilde{c}_j = \tilde{c}_j^* + \tilde{\varepsilon}_j$ with $\tilde{c}_j = c_j - \bar{c}$, $\tilde{c}_j^* = c_j^* - \bar{c}^*$ and $\tilde{\varepsilon}_j = \varepsilon_j - \bar{\varepsilon}_j$ such that $E(\tilde{\varepsilon}_j) = 0$, where the bar indicates sample mean.

The revealed preferred relation is represented by R. A taxonomy of revealed preference concepts can be found in Richter (1966, p. 638); here we use the "narrow sense" definition, as described in Varian (1982, p. 947), which under nonsatiation coincides with the "wide sense" concept. The equivalent concept for the true structure is indicated by R^*; thus $X^{*i}R^*X^{*j}$ is interpreted as "X^{*i} is preferred to X^{*j}" according to the unobserved quantities and prices (the term "revealed" is dropped since X^{*j} and P^{*i} are unobserved). The starred variables P^{*j}, X^{*j}, or $C^*(i,j)$, are referred to as the true structure.

The data $\{X^j, P^j,\ j=1,2,..,n\}$ satisfy the Generalized Axiom of Revealed Preference, or GARP (Varian, 1982, p. 947), if $X^i R X^j$ implies $C(j,j) \le C(j,i)$ for all i, j. The true structure satisfies the GARP if $X^{*i} R^* X^{*j}$ implies $C^*(j,j) \le C^*(j,i)$ for all i, j. The data $\{X^j, P^j, e^j,\ j=1,2,..,n\}$ satisfy the GARPe if $X^i R_e X^j$ implies $e^j P^j \cdot X^j \le P^j \cdot X^i$, where R_e^o is defined as $X^j R_e^o X$ if and only if $e^j P^j X^j \ge P^j X$, R_e is the transitive closure of R_e^o (i.e., the smallest transitive relation containing R_e^o) and the e^j are n scalars satisfying $0 \le e^j \le 1$ (see Varian, 1990, pp. 130-131).

Let $C_e(j,i) = e^j C(j,i)$ if j=i and $C_e(j,i) = C(j,i)$ if j≠i. The set $\{X^j, P^j, e^j,\ j=1,2,...,n\}$, or simply $C_e(j,i)$, is denoted the perturbed structure and $(e^1, e^2, ..., e^n)$ is the perturbation vector. The perturbed structure generated by the perturbation vector $e^j = \exp(-\varepsilon_j)$, j=1,2,..,n, will be denoted the perturbed-true structure (this corresponds to the expenditure matrix whose diagonal and off-diagonal elements equal the true and observed expenditures, respectively).

The perturbation indices e^j, j=1,2,..,n, proposed by Varian (1990), present an extension of Afriat's (1973, p. 463) single efficiency index and have a counterpart in the production context-- the γ_j defined by Hanoch and Rothchild (1972, p. 263). These indices are used to produce a goodness-of-fit measure of how well the data fit a particular optimization axiom (in addition to Varian [1990], see Chalfant and Alston [1988] p. 406, for a demand example, and Chavas and Cox [1990] p. 455, in the production context). Here we also follow this practice, by generating a goodness-of-fit index based on the perturbation indices (represented below by $\hat{\rho}_n$), but carry the analysis a few steps further by developing a measuring device, in terms of a statistical test, to evaluate the magnitude (significance) of the lack-of-fit (violation) of the data with an optimization axiom.

3. Testing the Significance of GARP Violation

Our test is motivated by a simple idea. We perturb the data until they satisfy the GARP and then assess the plausibility that the perturbed structure is the true one which generated the data. Put differently, we observe the set of all structures that satisfy the GARP and ask whether it is plausible that the observed data were generated by a member of this set. The formulation of the test entails (i) defining a notion of distance between the observed structure and a perturbed one, (ii) determining the minimum distance between the observed structure and a perturbed structure that satisfies the GARP, and (iii) providing statistical meaning to the plausibility that the data were generated by this particular structure.

We begin with the notion of distance between the observed structure and a perturbed structure. Any vector $v \in R^n$ generates the expenditure matrix given by $C(i,j)$ if $i \neq j$ and $\exp(v_j)$ if $i=j$; this is equivalent to the perturbed structure generated by the perturbation vector $e^j = \exp(v_j - c_j)$, $j=1,2,..,n$. In particular, $v = c$ and $v = c^*$ correspond to the observed structure and the perturbed-true structure, respectively. The terms v-vector and v-structure will be used interchangeably.

The distance between the observed structure and a perturbed v-structure is represented by $d_n(v,c) = E\{\sum_{i=1}^{n}(c_i - v_i)^2\}/n$ and is estimated by $\hat{d}_n(v,c) = \sum_{i=1}^{n}(c_i - v_i)^2\}/n$, where $E\{\cdot\}$ is the expectation operator. In particular, $d_n(c^*,c) = E\{\sum_{i=1}^{n}\varepsilon_i^2\}/n \equiv \sigma^2$ is the distance between the observed structure and the perturbed-true structure and $\hat{d}_n(c^*,c) = \sum_{i=1}^{n}\varepsilon_i^2/n \equiv s_n^2$ is its (unobserved) estimate.

Next, we define the set

$$\Gamma_n(\rho) = \{v \in R^n: d_n(v,c) \leq \rho\},$$

containing all the perturbed structures which are at most ρ away from the observed structure; the corresponding estimate is given by

$$\hat{\Gamma}_n(\rho) = \{v \in R^n: \hat{d}_n(v,c) \le \rho\}.$$

A vector (or structure) v satisfies the GARPe if its associated perturbation vector, $e^j = \exp(v_j - c_j)$, $j=1,2,..,n$, satisfies the GARPe. We shall say that $\Gamma_n(\rho)$, or $\hat{\Gamma}_n(\rho)$, satisfies the GARPe if it contains a structure that satisfies the GARPe. The minimum distance between the observed structure and a structure that satisfies the GARPe can now be defined as:

$$\rho_n = \text{Min}\{\rho: \Gamma_n(\rho) \text{ satisfies GARPe}\}, \tag{2}$$

with the estimate

$$\hat{\rho}_n = \text{Min}\{\rho: \hat{\Gamma}_n(\rho) \text{ satisfies GARPe}\}. \tag{$\hat{2}$}$$

Observe that $\rho' \le \rho''$ if and only if $\Gamma_n(\rho') \subseteq \Gamma_n(\rho'')$ [resp. $\hat{\Gamma}_n(\rho') \subseteq \hat{\Gamma}_n(\rho'')$], thus $\Gamma_n(\rho)$ [resp. $\hat{\Gamma}_n(\rho)$] satisfies the GARPe for all $\rho \ge \rho_n$ [resp. $\rho \ge \hat{\rho}_n$]. Note further that, when the GARP is violated by the data, $\hat{\Gamma}_n(0)$ does not satisfy the GARPe (since it entails $e^j = 1$, all j) whereas $\hat{\Gamma}_n(\infty)$ vacuously satisfies the GARPe (since $v_j = -\infty$ implies $e^j = 0$).

Equipped with a notion of distance between structures, particularly the minimum distance between the data and a perturbed structure that satisfies the GARPe, we proceed to provide statistical content to the statement "it is plausible that the true structure, from which the data were generated, satisfies the GARP".

If the satisfaction of the GARPe requires that at least one of the e^j vanishes, then we interpret this as evidence that the true structure could not possibly satisfy the GARP. Attention is therefore limited to cases where this unfavorable event does not occur and we require:

Assumption 2: There exists $\delta > 0$ such that, for all n, the GARPe can be satisfied by a perturbation vector $e^j \ge \delta$, $j=1,2,..,n$.

This assumption implies that ρ_n is bounded from above for all n and we can define

$$\rho_0 = \lim\sup \rho_n. \tag{3}$$

Now, $c^* \in \Gamma_n(\rho)$ for all $\rho \geq \sigma^2$ and $c^* \notin \Gamma_n(\rho)$ for all $\rho < \sigma^2$. If $\rho_n > \sigma^2$, then c^* could not possibly satisfy the GARP. On the other hand, if $\rho_n \leq \sigma^2$ then c^* belongs to a set that satisfies the GARP and may well satisfy the GARP itself. This leads to specifying the null hypothesis, which maintains that c^* satisfies the GARP, as:

$$H_0: \quad \rho_0 \leq \sigma^2.$$

If H_0 is rejected, then the conclusion that c^* could not possibly satisfy the GARP is, subject to the qualification of a statistical test, correct.

The following result provides a test-criterion for H_0. Let $z(\alpha)$ be the $1-\alpha$ positive quantile of the standard normal distribution, $\sigma^2 = \mathrm{Var}(\varepsilon_j)$ and $\tau^2 = \mathrm{Var}(\varepsilon_j^2)$.

Proposition 1: Under Assumption 1 and H_0,

$$\lim_{n\to\infty} \Pr\{\sqrt{n}(\hat{\rho}_n - \sigma^2)/\tau \leq z(\alpha)\} \geq 1-\alpha.$$

The proof relies on

Lemma 1: $(\hat{\rho}_n - \rho_n) \xrightarrow{p} 0$.

Proof: By the Law of Large Numbers, $\hat{d}_n(v,c) - d_n(v,c) \xrightarrow{p} 0$, i.e., $\hat{d}_n(v,c) = d_n(v,c) + o_p(1)$, where $A_n = o_p(B_n)$ if $A_n/B_n \xrightarrow{p} 0$ as $n \to \infty$. Let $v' \in R^n$ be the particular vector corresponding to ρ_n, i.e., v' satisfies the GARPe, $v' \in \Gamma_n(\rho_n)$ and $d_n(v',c) = \rho_n$. Then, $v' \in \hat{\Gamma}_n(\rho_n + o_p(1))$, implying that

$$\hat{\rho}_n \leq \rho_n + o_p(1).$$

Let $v'' \in R^n$ be the particular vector corresponding to $\hat{\rho}_n$, i.e., v'' satisfies the GARPe, $v'' \in \hat{\Gamma}_n(\hat{\rho}_n)$ and $\hat{d}_n(v'',c) = \hat{\rho}_n$. Then $v'' \in \Gamma_n(\hat{\rho}_n - o_p(1))$, implying that

$$\rho_n \leq \hat{\rho}_n - o_p(1).$$

Together, the two inequalities imply $\hat{\rho}_n = \rho_n + o_p(1)$, as asserted.■

It follows immediately from the lemma, using (3), that:

Corollary: $\mathrm{Prob}\{\hat{\rho}_n \leq \rho_0\} \to 1$ as $n \to \infty$.

Proof of proposition 1: By standard application of central limit theory, $\sqrt{n}(s_n^2 - \sigma^2)/\tau \xrightarrow{d} N(0,1)$, where it is recalled that $s_n^2 \equiv \sum_{i=1}^n \varepsilon_i^2/n$. Thus

$$\Pr\{\sqrt{n}(s_n^2 - \sigma^2)/\tau \le z(\alpha)\} \longrightarrow 1-\alpha.$$

Now, $s_n^2 \xrightarrow{p} \sigma^2$, $\Pr\{\hat{\rho}_n \le \rho_o\} \to 1$ and Ho imply $\Pr\{\hat{\rho}_n \le s_n^2\} \to 1$. Thus, under Ho:

$$\Pr\{\sqrt{n}(\hat{\rho}_n - \sigma^2)/\tau \le \sqrt{n}(s_n^2 - \sigma^2)/\tau\} \longrightarrow 1.$$

Let A denote the event $\{\sqrt{n}(s_n^2-\sigma^2)/\tau \le z(\alpha)\}$, B denote the event $\{\sqrt{n}(\hat{\rho}_n-\sigma^2)/\tau \le \sqrt{n}(s_n^2-\sigma^2)/\tau\}$ and C denote the event $\{\sqrt{n}(\hat{\rho}_n-\sigma^2)/\tau \le z(\alpha)\}$. Then, we have $\Pr\{A\} \to 1-\alpha$ and, under Ho, $\Pr\{B\} \to 1$, implying that $\Pr\{A \cup B\} \to 1$; thus $\Pr\{A \cap B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{A \cup B\} \to 1-\alpha$. Since $C \supseteq A \cap B$, $\Pr\{C\} \ge \Pr\{A \cap B\}$ for all n. Taking limits on both sides gives

$$\lim_{n\to\infty} \Pr\{\sqrt{n}(\hat{\rho}_n-\sigma^2)/\tau \le z(\alpha)\} \equiv \lim_{n\to\infty} \Pr\{C\} \ge \lim_{n\to\infty} \Pr\{A \cap B\} = 1-\alpha,$$

as asserted. ■

Remarks: **(i)** If $\rho_o = \sigma^2$ and $\hat{\rho}_n = s_n^2 + o_p(1/\sqrt{n})$, an event which is permitted under Ho, then $\Pr\{\sqrt{n}(\hat{\rho}_n-\sigma^2)/\tau \le z(\alpha)\} \to 1-\alpha$ and Proposition 1 holds with equality. **(ii)** If the ε_j's are normal, then ns_n^2/σ^2 is distributed as $\chi^2_{(n)}$ and Proposition 1 becomes: Under Ho, $\lim_{n\to\infty} \Pr\{n\hat{\rho}_n/\sigma^2 \le \chi^2_{(n)}(\alpha)\} \ge 1-\alpha$, where $\chi^2_{(n)}(\alpha)$ is the $1-\alpha$ right quantile of $\chi^2_{(n)}$.

According to Proposition 1, a test which rejects Ho whenever $\hat{\rho}_n \ge \sigma^2 + z(\alpha)\tau/\sqrt{n}$, has a significance level (for large enough n) no greater than α, and actually attains the size α under the conditions of Remark (i). Thus, the test criterion

$$\textit{reject Ho if} \quad \sigma^2 \le \hat{\rho}_n/(1 + z(\alpha)\sqrt{\theta}/\sqrt{n})$$

has a significance level no greater than α for all distributions F in the set

$$\mathfrak{F}(\theta) = \{F\colon\ \tau^2/\sigma^4 \le \theta\}.$$

The set $\mathfrak{F}(\theta)$ satisfies $\mathfrak{F}(\theta') \subseteq \mathfrak{F}(\theta'') \Leftrightarrow \theta' \le \theta''$; thus, for example, $\mathfrak{F}(\theta)$ contains the normal distribution for all $\theta \ge 2$ and the uniform distribution for all $\theta \ge 4/5$. If τ^2/σ^4 is known, then θ is set equal to this value. Lacking such knowledge, the value of θ is set equal to the least upper bound of τ^2/σ^4.

To evaluate the index $\hat{\rho}_n$ we utilize the (well known) fact that the data $C(j,i)$ satisfy the GARP if and only if there exist positive scalars u_j and λ_j, $j=1,2,..,n$, satisfying Afriat's inequalities

$$u_j \le u_i + \lambda_j(C(j,i) - C(j,j)), \quad i,j=1,2,\dots,n.$$

It is readily verifiable that the perturbed data $C_e(j,i)$ satisfy the GARP$_e$ if and only if there exist $u_j, \lambda_j > 0$ and $0 < e^j \le 1$, $j=1,2,..,n$, satisfying

$$u_j \le u_i + \lambda_j(C(j,i) - e^jC(j,j)), \quad i,j=1,2,\dots,n.$$

An operational definition of $\hat{\rho}_n$ can now be given as:

$$\hat{\rho}_n = \text{MIN} \sum_{j=1}^{n} (\log e^j)^2/n \qquad (4)$$

subject to: $u_j \le u_i + \lambda_j[C(j,i) - e^jC(j,j)], \quad j,i = 1,2,..,n;$

$u_j, \lambda_j, e^j > 0, \; e^j \le 1, \; j = 1,2,\dots,n.$

Given $\hat{\rho}_n$ and θ, implementing the test requires information on σ^2. The actual expenditure data can provide some rough bound on σ^2, since (cf. Eq. (1)) $\sigma^2 \le \sigma_c^2$, where $\sigma_c^2 = E(c_j - \bar{c})^2$ can be estimated by $\hat{\sigma}_c^2 = \sum_{j=1}^{n}(c_j - \bar{c})^2/n$ with $\bar{c} = \sum_{j=1}^{n} c_j/n$. But this bound is useful only if $\hat{\rho}_n \ge \hat{\sigma}_c^2$. If $\hat{\rho}_n < \hat{\sigma}_c^2$, then one needs to resort to extraneous information on σ^2; such information can be obtained, for instance, by studying the process by which the expenditure data were collected (a process conducted by human beings) and calibrating its accuracy. The need to rely on extraneous information of the error variance appears to be a recurrent property of hypothesis testing in the context of nonparametric analyses (see Varian [1985] and Epstein and Yatchew [1985]).

It is illuminating to compare the distribution-free test with the test that maintains normal errors. Following Remark (ii), with normally distributed errors, H₀ is rejected if $\sigma^2 \le \hat{\rho}_n/[\chi^2_{(n)}(\alpha)/n]$. Now, using the approximation $\chi^2_{(n)}(\alpha) = \frac{1}{2}\left(z(\alpha) + \sqrt{2n-1}\right)^2$ (see, e.g., Lindgren [1976, p. 195]), it is easy to verify that $\chi^2_{(n)}(\alpha)/n = 1 + z(\alpha)\sqrt{2}/\sqrt{n} + O(1/n)$, where $A_n = O(B_n)$ if A_n/B_n is bounded for all n. It follows that the normal test can be approximated, to the order O(1/n), by the test

$$\textit{reject } H_0 \textit{ if } \sigma^2 \le \hat{\rho}_n/(1 + z(\alpha)\sqrt{2}/\sqrt{n}),$$

which is equivalent to the distribution-free test with $\theta = 2$. But $\theta = 2$ is the least θ-value for which $\mathfrak{F}(\theta)$ still contains the normal distribution. It appears therefore that the distribution-free test utilizes efficiently the limited available information. If $\theta > 2$, the normal test is sharper, in that its rejection region is larger than that of the distribution-free test. The normal test, however, could be misleading if the moment ratio τ^2/σ^4 exceeds 2. As expected, the two tests coincide in the limit of large n.

This completes our formulation of the test procedure. Implementing the test requires calculating $\hat{\rho}_n$, which, as it stands now, entails solving the nonlinear problem (4). This task can become quite formidable, since the number of constraints raises rapidly with the number of observations n (like n^2). Fortunately, there are effective ways to make this task more manageable, and in some cases to avoid it altogether, as is discussed in the next section.

4. Implementation

The "curse" of the nonlinear problem (4) lies in the number of constraints, which increases like n^2. It would therefore be useful if these constraint were linear (on this point see Brooke *et al.* [1988, p. 158]). Fortunately, this can easily be achieved by a proper redefinition of the variables. In particular, by defining $g_j = \lambda_j e^j$, Problem (4) becomes:

$$\hat{\rho}_n = \text{MIN} \sum_{j=1}^{n} \left(\log(g_j/\lambda_j)\right)^2 /n \tag{5}$$

subject to: $u_j \le u_i + \lambda_j C(j,i) - g_j C(j,j), \quad j,i = 1,2,\ldots,n;$

$g_j - \lambda_j \le 0, \; j = 1,2,\ldots,n;$

$u_j, \lambda_j, g_j > 0, \; j = 1,2,\ldots,n.$

In this form, experience shows (see next section), that situations with n=100 are readily handled on a micro (Vax) type computer.

While the linearized form (5) reduces drastically computation requirements relative to (4), these are still quite substantial (and may even be prohibitive) for large n (say, n ≥ 150). It is therefore important to note that in many cases the nonlinear programming task, of solving (5), can be avoided altogether. To see this, suppose there exists a distance measure $\tilde{\rho}_n$ which is close to $\hat{\rho}_n$ but lies above it, and is easy to calculate. Suppose further that a test that uses $\tilde{\rho}_n$ instead of $\hat{\rho}_n$ cannot reject H_0. Then, since $\tilde{\rho}_n \geq \hat{\rho}_n$, H_0 cannot be rejected with $\hat{\rho}_n$ either and there is no need for $\hat{\rho}_n$.

Indeed, an estimator $\tilde{\rho}_n$, which performs extremely well (in the sense of being very close to $\hat{\rho}_n$), is attainable by the algorithm proposed in Tsur (1989). (For the sake of completeness we repeat the algorithm in the appendix.) This procedure can take at most n iterations, is easy to implement numerically (a case with 150 observations took seconds on a pc) and, as the application in the next section reveals, gives estimates which are very close to $\hat{\rho}_n$.

If, alas, $\tilde{\rho}_n$ rejects H_0, then $\hat{\rho}_n$ is required to verify this result (because $\tilde{\rho}_n \geq \hat{\rho}_n$, it is possible that $\tilde{\rho}_n$ rejects H_0 whereas $\hat{\rho}_n$ does not). In this case the output of the $\tilde{\rho}_n$-algorithm is still useful, as it can serve to provide a "good" initial feasible solution from which the minimization of (5) departs (by a good solution we mean a solution close to the true minimum). A good initial point is important since in general the objective in (5) possesses multiple (local) minima and the global minimum is more likely to be reached if the initial point lies in its close vicinity. Of course, it is always possible to experiment with many initial values and to choose the least of all convergent points, but such an approach increases computation significantly. (On how to use the output of the $\tilde{\rho}_n$-algorithm to calculate initial values for Problem (5), see discussion in the appendix.)

5. Production decisions

In the production context the observed data consist of netput vectors $z_j = (y_j, -x_j)$ and their associated prices $q_j = (p_j, w_j)$, $j=1,2,..,n$. Here y_j and p_j are scalars representing quantity and price of output, and x_j and w_j are k-vectors of input quantities and prices. The profit data $R(j,i) = q_j \cdot z_i$ contain measurement errors. The true structure, denoted by $R^*(j,i)$, is related to the observed structure according to $R(j,i) = \Omega(j,i)R^*(j,i)$, where $\Omega(j,i)$ represents measurement errors associated with the profit data. Letting $r_j = \log R(j,j)$, $r_j^* = \log R^*(j,j)$ and $\omega_j = \log \Omega(j,j)$, yields (cf. Eq. (1))

$$r_j = r_j^* + \omega_j, \quad j=1,2,..,n,$$

where the errors $\omega_1, \omega_2, .., \omega_n$ satisfy Assumption 1.

The data satisfy the Weak Axiom of Profit Maximization, or WAPM, (Varian, 1984. p. 584) if, and only if, $R(j,j) \geq R(j,i)$ for all $j,i = 1,2,..,n$. The true structure satisfies the WAPM if, and only if, $R^*(j,j) \geq R^*(j,i)$ for all $j,i = 1,2,..,n$. We introduce the perturbation scalars $e_j \geq 1$, $j=1,2,..,n$, and the associated perturbed data (z_j, q_j, e_j), $j=1,2,..,n$, and say that the perturbed data satisfy the WAPMe if, and only if, $e_jR(j,j) \geq R(j,i)$ for all $j,i = 1,2,..,n$.

Any n-vector v generates a perturbed structure given by $R(j,i)$ if $j \neq i$ and $\exp(v_j)$ if $j=i$; this is equivalent to the perturbed structure generated by the perturbation vector $e_j = \exp(v_j - r_j)$. A vector (or structure) v satisfies the WAPMe if its associated perturbed structure satisfies the WAPMe. The distance between the observed structure R and a v-structure is measured by $d_n(r,v) = E\{\Sigma_{j=1}^n (r_j - v_j)^2\}/n$ and estimated by $\hat{d}_n(r,v) = \Sigma_{j=1}^n (r_j - v_j)^2/n$.

The sets

$$\Gamma_n(\rho) = \{v \in R^n: d_n(v,c) \leq \rho\} \quad \text{and} \quad \hat{\Gamma}_n(\rho) = \{v \in R^n: \hat{d}_n(v,c) \leq \rho\}$$

are said to satisfy the WAPM if they contain a v-structure that satisfies the WAPMe. Consequently, the indices

$$\rho_n = \text{Min}\{\rho\text{: } \Gamma_n(\rho) \text{ satisfies the WAPM}\}$$

and

$$\hat{\rho}_n = \text{Min}\{\rho\text{: } \hat{\Gamma}_n(\rho) \text{ satisfies the WAPM}\}$$

are defined as in Eqs. (2).

We require the existence of an upper bound $M < \infty$, such that, for all n, the WAPMe can be satisfied by perturbations $e_j \le M$, $j=1,2,..,n$. Thus, ρ_n is bounded from above and ρ_0 can be defined as in Eq. (3).

Let $h_j = \underset{i=1,2,...,n}{\text{MAX}} \left(\log R(j,i) - \log R(j,j)\right)$ and $g_j = \text{MAX}\left(0, h_j\right)$, $j=1,2,...,n$. Then, it is easy to verify that

$$\hat{\rho}_n = \textstyle\sum_{j=1}^{n} g_j^2/n.$$

With $\sigma^2 = \text{Var}(\omega_j)$, the null hypothesis, which maintains that the true structure satisfies the WAPM, remains

$$\text{Ho: } \rho_0 \le \sigma^2.$$

Following Proposition 1, Ho is rejected whenever $\hat{\rho}_n \ge \sigma^2 + z(\alpha)\tau/\sqrt{n}$, where σ^2 = $\text{Var}(\omega)$ and $\tau^2 = \text{Var}(\omega^2)$. The test criterion

$$\textit{reject Ho whenever } \sigma^2 \le \hat{\rho}_n/(1+z(\alpha)\sqrt{\theta}/\sqrt{n})$$

ensures a significance level no greater than α for all ω-distributions F in the set $\mathfrak{F}(\theta) = \{F\text{: } \tau^2/\sigma^4 \le \theta\}$.

To test for cost minimization behavior (note that profit maximization is stronger than cost minimization, as the former implies the latter but not vice versa), redefine q_j and z_j as: $q_j = w_j$ and $z_j = x_j$, $j=1,2,..,n$. Thus, $R(j,i) = q_j \cdot z_j$ represents the cost of using input x_i at prices w_j and $\Omega(j,i)$ is the measurement errors associated with the cost data. The data satisfy the Weak Axiom of Cost Minimization, or WACM (Varian, 1984, p. 582) if, and only if, $R(j,j) \le R(j,i)$ for all $y_j \le y_i$. A perturbed structure associated with the perturbation vector $e_j \le 1$, $j=1,2,..,n$, is given by $R(j,i)$ for $i \ne j$, and $e_j R(j,j)$, $i,j=1,2,..,n$. The perturbed structure satisfies the WACMe if, and only if, $e_j R(j,j) \le R(j,i)$ for all $y_j \le y_i$, $i,j=1,2,..,n$. Let $M_j = \{i\text{: } y_j \le$

$y_i\}$, and redefine h_j and g_j as: $h_j = \underset{i \in M_j}{\text{MIN}} \left(\log R(j,i) - \log R(j,j) \right)$ and $g_j = \text{MIN}\left(0, h_j\right)$, j=1,2,..,n. With obvious modifications, the procedure for testing the significance of the WACM violation proceeds along the same steps as above, using $\hat{\rho}_n = \sum_{j=1}^{n} g_j^2/n$ evaluated at the above redefined g_j's.

In the absence of price data, i.e., when only y_j and x_j, j=1,2,..,n, are available, one can proceed by seeking a set of input prices under which WACM is satisfied. Programs (8) of Hanoch and Rothchild (1972, p. 262) is designed for this purpose. The outcome of this sequence of linear programs includes the indices γ_j, j=1,2,..,n, which are equivalent to the minimal perturbations needed to ensure the existence of (positive) input prices under which the data satisfy the WACM.

5.1 *Technological change*

There are two main reasons for the violation of an optimization axiom. First, the input-output decisions may not be determined only according to profit maximization or cost minimizing considerations. Second, the production technology may vary across producers (this is particularly relevant when dealing with time series data, where technological differences are likely to prevail over time as a result of technological progress). These two causes are observationally indistinguishable, in that the data do not contain enough information to identify the cause of the violation.

If, however, the assumption of optimization behavior is maintained, then violations, if they occur, must be due to technological progress, and the nature of the violation can then be used to study the nature of the technical change process. Indeed, this idea has been utilized by Chavas and Cox (1988, 1990), who incorporated technological change into nonparametric production theory and used this approach to study technological progress processes in U.S. agriculture and in U.S. and Japanese manufacturing.

Abstracting from measurement errors, Chavas and Cox (1990, p. 455) discuss goodness-of-fit indices of nonparametric tests of cost minimization, which are based on perturbation indices similar to the above e_j's. It is unclear, however, which values of these indices constitute good fit and which values constitute lack of fit. In the present construction, $\hat{\rho}_n$ represents such a goodness-of-fit measure, with a decreasing fit indicated by $\hat{\rho}_n$ moving away from zero, and the magnitude, or significance, of $\hat{\rho}_n$ is measured relative to the error variance (σ^2) via the hypothesis testing procedure.

If the violation of optimization behavior is due to technical changes, as is maintained by the nonparametric productivity analysis, then the significance of the violation indicates the significant of the technical change process. Thus, for example, a rejection of the hypothesis that the true structure does not violate WACM, can be interpreted as evidence that the technical change process is not merely due to data measurement errors, but rather a persistent and significant process. The technology coefficients can then be evaluated using the procedures suggested by Chavas and Cox (1988, 1990). If, however, the violation is insignificant, then some caution ought to be exercised when drawing conclusions from nonparametric analysis of technical change.

6. **Application**

The data consist of monthly consumption and prices of four major meat types (the data were collected in Spain and are available upon request). Different samples, corresponding to sub-periods of lengths n = 20, 50, 75, 100 and 150 months (n stands for the number of months in the sample), are considered. All samples exhibit some violations of the GARP, thus $\tilde{\rho}_n$ and $\hat{\rho}_n$ are calculated (except for $\hat{\rho}_{150}$) and a test of the significance of the violations is performed.

The index $\tilde{\rho}_n$ is produced by a computer code realization (Fortran) of the $\tilde{\rho}_n$-algorithm described in the appendix. The output of this routine is then used to calculate feasible values for u_j, λ_j and g_j, j=1.2,..,n, i.e., values satisfying the Afriat inequalities in (5). The nonlinear minimization task to produce $\hat{\rho}_n$ is performed by GAMS (Brooke *et al.* [1988]) installed on a Vax 6000-510 machine. The results are presented in Table 1.

Table 1

In the first three cases (n = 20, 50 and 75), $\tilde{\rho}_n$ is slightly above $\hat{\rho}_n$, as it should be, and the difference between the two is extremely small. In the fourth case, that with n = 100, $\tilde{\rho}_n$ is slightly below $\hat{\rho}_n$, which, by the definition of $\hat{\rho}_n$, is impossible. Clearly, the nonlinear minimization routine picked a local minimum which does not coincide with the global minimum. In fact, other runs of GAMS with different, arbitrary initial points yielded other local minima which were all greater than the value of $\hat{\rho}_n$ reported in Table 1. For the n=150 sample, only $\tilde{\rho}_n$ is reported; the computations required to calculate $\hat{\rho}_n$ exceeded the capacity of the computer.

These results seem to illuminate the importance of the $\tilde{\rho}_n$-algorithm, both (i) in producing good initial points for the nonlinear programming routine and (ii) in providing an alternative for $\hat{\rho}_n$ when the nonlinear problem (5) is unmanageable.

We investigate now the significance of the GARP violation of the n = 150 sample. The estimated variance of the expenditure sample $\{c_j = \log C(j,j), j=1,2,..,n\}$ is $\hat{\sigma}_c^2 = 0.376$ and $\tilde{\rho}_{150} = 1.011\times10^{-5}$ (see Table 1). With $\alpha = 5\%$, $\tilde{\rho}_n/(1+z(5\%)\sqrt{\theta}/\sqrt{n}) = 8.5\times10^{-6}$ or 4.3×10^{-6} as $\theta = 2$ or 100, respectively.

Because $\hat{\sigma}_c^2$ exceeds $\tilde{\rho}_n/(1+z(5\%)\sqrt{\theta}/\sqrt{n})$, a clear-cut rejection of H_0 is impossible (see discussion in Section 3). The magnitudes of these parameters, however, convey some information. Suppose $\sigma^2 \le \tilde{\rho}_n/(1+z(5\%)\sqrt{\theta}/\sqrt{n})$, so that $\sigma^2/\hat{\sigma}_c^2 \le [\tilde{\rho}_n/(1+z(5\%)\sqrt{\theta}/\sqrt{n})]/\hat{\sigma}_c^2 \cong 10^{-5}$. Eq. (1) implies, in this case, that the measurement errors account for no more than 0.001 percent of the variation in c_j, the rest being contributed by the variation in c_j^*. This entails an extreme level of accuracy of measurement, which is not typical for data collected in an uncontrolled experiment. It appears likely therefore that $\sigma^2 \geq \tilde{\rho}_n/(1+z(5\%)\sqrt{\theta}/\sqrt{n})$, in which case Ho cannot be rejected. But, since $\tilde{\rho}_n \geq \hat{\rho}_n$, a test based on $\hat{\rho}_n$ could not have rejected Ho either. Thus, based on $\tilde{\rho}_n$ and without having to calculate $\hat{\rho}_n$, we conclude that the violation of the GARP by the data is not severe enough to render the satisfaction of the GARP by the true structure unlikely.

By way of comparison, suppose normal errors. Following Remark (ii) gives $n\tilde{\rho}_n/\chi_n^2(5\%) = 150\times10.11\times10^{-6}/179.3 = 8.46\times10^{-6}$ as the 5 percent upper critical level for σ^2. As one would expect, this level is almost identical with the critical level of the distribution-free test with θ=2 (see discussion in Section 3). Also, the normal test is sharper in that its rejection region is larger: if σ^2 lies between 8.46×10^{-6} and 4.3×10^{-6} then the normal test would reject Ho, leading to the conclusion that the true structure could not possibly satisfy the GARP, whereas the distribution-free test with θ=100 would not reject Ho. This is reasonable because the nonparametric test accommodates an entire family of distributions, of which the normal distribution is but one member. If the normal assumption is wrong, however, the parametric test could be misleading.

7. Concluding Comments

This works develops a framework for testing the significance of deviation from optimal consumption and production decisions when observations are measured with errors. The tests are free of a parametric specification of the error distribution and are easy to implement numerically.

The distribution-free property is accomplished up to an independence requirement and some moment bounds. The iid requirement can be relaxed, a task left for future research. Placing bounds on the moments of the error distribution appears to be unavoidable and occurs in other, related tests (see Varian, 1985, and Epstein and Yatchew, 1985). This is so because nonparametric analysis does away with the parametric structure, usually maintained in econometric models, that allows one to estimate moments of the error distribution from observed data.

The computational requirements of the GARP test depend on whether $\tilde{\rho}_n$ alone can do the job, or is $\hat{\rho}_n$ also needed (see discussion in Section 4). In the former case, there are no practical limits on the number of observations (the case with n=150 took seconds on a 386-pc). The latter case is more involved: solving the nonlinear program (5) with n=100 using GAMS installed on a Vax 6000-510 computer took about two hours, and the case with 150 observations exceeded the computer's capacity.

Though no results are available on the relationship between $\tilde{\rho}_n$ and $\hat{\rho}_n$, the application here suggests that they are extremely close to each other. Thus, with large models, $\tilde{\rho}_n$ should be calculated first. If the test based on $\tilde{\rho}_n$ rejects H_0, then, and only then, should an attempt at calculating $\hat{\rho}_n$ be considered based on the computational resources available and the nature of the problem on hand. In this case the output of the $\tilde{\rho}_n$-algorithm serves as a useful starting point for the nonlinear procedure.

Because production output is observable, unlike preferences output, i.e., utility, the computations required to test rationality of production decisions are substantially lighter than those required to test rationality of consumption decisions.

Appendix: (a) *The $\tilde{\rho}_n$-algorithm*

Input: the n by n expenditure matrix C;

Output: a perturbation vector, M_j, j=1,2,..,n, satisfying the GARPe and the associated index $\tilde{\rho}_n$;

(1) set M(j)=1 and Ce(i,j)=C(i,j) for i,j=1,2,..,n;

(2) set Ce(j,j)=M(j)C(j,j), j=1,2,..,n;

(3) for i,j=1,2,..,n set $R_e^0(i,j)$ = 1 or 0 as Ce(i,i) ≥ or < Ce(i,j), respectively, and calculate its transitive closure Re [for an algorithm to calculate the transitive closure of a matrix see Varian (1982, p. 972)];

(4) set Ge = {j: Re(i,j) = 1 and Ce(j,j) > Ce(j,i) for at least one case i};

n5) if Ge = ø then go to (7), else go to (6);

(6) calculate the n by 1 vector M as

$$M_j = \begin{cases} \min_{x^i R_e x^j} \{C(j,i)/C(j,j)\} & ; j \in Ge \\ M_j & ; j \notin Ge \end{cases}$$

and go to (2);

(7) calculate

$$\tilde{\rho}_n = \sum_{j=1}^{n} [\log M_j]^2 / n$$

and stop.

(b) *Calculating initial values for Problem (5)*

Initial values of u_j, λ_j and g_j, j=1,2,..,n, associated with the output M_j, j=1,2,..,n, of the above algorithm are calculated as follows. The perturbed data $Ce(i,j) = \begin{cases} M_j C(j,i) & \text{if } j=i \\ C(j,i) & \text{if } j \neq i \end{cases}$ satisfies, by construction, the GARPe. Thus, the Afriat numbers (u_j,λ_j, j=1.2,..,n) associated with Ce(i,j) can be calculated using Algorithm 3 of Varian (1982, p. 968). The corresponding g_j's are then given by $g_j = \lambda_j M_j$, j=1,2,..,n.

References

Afriat, Sydney N., 1967, The construction of utility function from expenditure data, *International Economic Review* 8, 67-77.

Afriat, Sydney N., 1973, On a system of inequalities in demand analysis: An extension of the classical method, *International Economic review*, 14, 460-472.

Brooke, Anthony, David Kendrick and Alexander Meeraus, 1988, *GAMS: a user guide*, The Scientific Press, Redwood City.

Chalfant, James A. and Julian M. Alston, 1988, Accounting for changes in tastes, *Journal of Political Economy*, 96, 391-410.

Chavas, Jean-Paul and Thomas L. Cox, 1988, A nonparametric analysis of agricultural technology, *American Journal of Agricultural Economics*, 70, 303-310.

Chavas, Jean-Paul and Thomas L. Cox, 1990, A non-parametric analysis of productivity: The case of U.S. and Japanese manufacturing, *American Economic Review*, 80, 450-464.

Epstein, Larry G. and Adonis J. Yatchew, 1985, Non-parametric hypothesis testing procedures and applications to demand analysis, *Journal of Econometrics*, 30, 149-169.

Hanoch, Giora and Michael Rothchild, 1972, Testing the assumptions of production theory, *Journal of Political Economy* 80, 256-275.

Lindgren, Bernard W., 1976, *Statistical Theory*, Macmillan, New York.

Richter, Marcel K. 1966, Revealed preference theory, *Econometrica*, 34, 635-645.

Tsur, Yacov, 1989, On testing for revealed preference conditions, *Economics Letters*, 31, 359-362.

Varian, Hal R., 1982, The nonparametric approach to demand analysis, *Econometrica* 50, 945-972.

Varian, Hal R., 1984, The nonparametric approach to production analysis, *Econometrica* 52, 579-597.

Varian, Hal R., 1985, Nonparametric analysis of optimizing behavior with measurement error, *Journal of Econometrics* 30, 445-458.

Varian, Hal R., 1990, Goodness-of-fit in optimizing models, *Journal of Econometrics* 46, 125-140.

Table 1

Violation of the GARP by meat demand data

n = # of months in the sample[a]	# of GARP violations[b]	$\tilde{\rho}_n$ $(\times 10^{-6})$	$\hat{\rho}_n$ $(\times 10^{-6})$	$\hat{\sigma}_c^2 = \sum_{i=1}^{n}(c_i-\bar{c})^2/n$
20	4	3.92	3.5	0.0208
50	11	4.32	4.24	0.0691
75	14	9.57	8.47	0.1223
100	18	7.30	7.90	0.1916
150	33	10.11	NA	0.3764

[a]All samples begin on January 1970 and cover n consecutive months.

[b]No. of months for which at least one violation of the GARP was detected.