INTERNATIONAL ECONOMIC REVIEW
Vol. 38, No. 1, February 1997

EXACT NONPARAMETRIC TESTS OF ORTHOGONALITY AND RANDOM WALK IN THE PRESENCE OF A DRIFT PARAMETER*

By Bryan Campbell and Jean-Marie Dufour[1]

Concordia University, Canada
Université de Montréal, Canada

In this paper, finite-sample nonparametric tests of conditional independence and random walk are extended to allow for an unknown drift parameter. The tests proposed are based on simultaneous inference methods and remain exact in the presence of general forms of feedback, nonnormality and heteroskedasticity. Further, in two simulation studies we confirm that the nonparametric procedures are reliable, and find that they display power comparable or superior to that of conventional tests.

1. INTRODUCTION

In certain modelling situations where the explanatory variables are not strictly exogenous, it is necessary to rely on asymptotic theory to justify inference based on standard regression procedures. The question of the reliability of these procedures in small samples naturally arises. A noteworthy example where there is feedback from disturbances that are contemporaneously uncorrelated with the regressors but which affect their future values has been studied by Mankiw and Shapiro (1986), Banerjee and Dolado (1987, 1988), Galbraith, Dolado and Banerjee (1987), and Banerjee, Dolado and Galbraith (1990). Here standard regression-based procedures reject much too often even in large samples. Another important example involving feedback is the random walk model.

Our work is inspired by results from classical finite-sample nonparametric statistics that show that the only tests about a median or a mean which are valid under sufficiently general distributional assumptions, allowing nonnormal, possibly heteroskedastic, independent observations, are based on sign statistics (see Lehmann

* Manuscript received May 1993; revised September 1995.

[1] This work was supported by the Social Sciences and Humanities Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, and the Government of Québec (Fonds FCAR). The authors thank Jean-Pierre Florens, Eric Renault, Jim Stock, Alain Trognon, Victoria Zinde-Walsh, two anonymous referees as well as seminar participants at the Université Libre de Bruxelles, the University of Bristol, the 1993 Meetings of the Canadian Economics Association, the 1994 Econometric Society European Meeting in Maastricht, the University of Amsterdam, Freie Universität Berlin, ENSAE (Paris) and the Institut of d'Economie Industrielle (Université de Toulouse) for helpful comments and discussion. All correspondence should be addressed to the authors at the Centre de recherche et développement en économique (C.R.D.E.), Université de Montréal, C.P.6128, Succursale A, Montreal (Québec), Canada H3C 3J7.

and Stein 1949, or Pratt and Gibbons 1981 for a more accessible discussion). Following these characterizations, we introduced in Campbell and Dufour (1991, 1995) sign and signed rank tests which were shown to be exact for a wide class of models, also allowing the presence of general forms of feedback as well as nonnormality and heteroskedasticity. Simulation results indicated that their power is comparable or superior (often by a wide margin) to that of the usual t-tests, using either asymptotic or size-corrected critical values for the Mankiw-Shapiro model and the Dickey-Fuller critical values for the random walk model. The methods were applied to the evaluation of Federal budget projections in Campbell and Ghysels (1995). These distribution-free tests, on the other hand, are only applicable when the median of the dependent variable is zero under the null hypothesis.

In this paper, we extend this nonparametric approach to cover a much wider class of applications of orthogonality tests where there is an unknown intercept or drift parameter. Often, as for example in the case of the expectations theory of the term structure of interest rates, an implicit forecast error can be associated with a model which, under the further assumption of efficiency or rationality, is hypothesized to be orthogonal to past information once a centering parameter is accounted for. In the term structure example, this parameter is interpreted as a liquidity premium; see Shiller et al. (1983), Fama (1984), Mankiw and Summers (1984), Mankiw and Miron (1986), Kugler (1990), Taylor (1992), Engsted (1993), and the surveys of Melino (1988) and Shiller (1990). Similarly, it is often of interest to allow for the presence of a drift in a random walk model. Standard regression procedures in such situations simply include an intercept term in the equation to be estimated. By contrast, more involved analysis is required to obtain distribution-free methods when the null hypothesis allows for an unknown intercept or drift as nuisance parameter. The purpose of this paper is to extend earlier results to cover such cases. Our approach is based on extending to a nonparametric context the simultaneous inference approach used in Dufour (1990) for a parametric regression model with Gaussian AR(1) disturbances. Here this work is accomplished by combining an exact nonparametric confidence set for the drift parameter, which can be obtained by "inverting" sign or signed rank tests, with "conditional" nonparametric tests linked to each point in the confidence set. The approach then yields finite-sample generalized bounds tests. For a review of earlier work on distribution-free methods in time series, the reader may consult Dufour, Lepage and Zeidan (1982) and the excellent recent survey by Hallin and Puri (1991).

Section 2 of the paper describes the general stochastic framework, which includes the special case of type of feedback found in the Mankiw-Shapiro and random walk models, and also allows for an intercept (or drift) parameter. In the first step, we assume provisionally that this nuisance parameter is known. In this context, we introduce the appropriate nonparametric statistics and derive their finite-sample distributions under the null hypothesis of conditional independence given the past. Then, in Section 3, we drop the assumption that the intercept parameter is known. For this case, we propose a three-stage testing procedure and prove a general result giving probability bounds for the procedure under the null hypothesis. In Section 4, we use Monte Carlo methods to compare a number of variants of the bounds procedures and investigate the power of the proposed nonparametric tests for simple linear regressions of the Mankiw-Shapiro (1986) type and for random walk models, both with intercept term and for various distributional assumptions (normal and nonnormal disturbances, with or without heteroskedasticity). The results confirm that the bounds nonparametric tests have the correct level while conventional asymptotic tests can easily reject much too frequently, and show that the power of the nonparametric procedures are at least comparable (and dominate often by a wide margin in the presence of outliers) to that of size-corrected conventional tests. In Section 5, we apply our methods to test the expectations theory of the term structure of interest rates using Canadian data on three and six-month rates. The nonparametric results are also contrasted with those found by the standard regression-based approach. We find that the usual results which reject efficiency of the implicit forecast may be spurious. Section 6 concludes.

2. FRAMEWORK

As in Campbell and Dufour (1995), we work within the framework of a general model involving the random variables $Y_1, \ldots, Y_n, X_0, \ldots, X_{n-1}$, and the corresponding information vectors defined by $I_t = (X_0, X_1, \ldots, X_t, Y_1, \ldots, Y_t)'$, where $t = 0, \ldots, n-1$, with the convention that $I_0 = (X_0)$. Our goal is to introduce tests of the independence of Y_t from I_{t-1}, which are exact under very weak assumptions concerning the distribution of Y_t and the relationship between Y_t and X_t. For one group of tests, we simply assume that Y_t has median b_0; for the other, we make the stronger assumption that the distribution of Y_t is symmetric about b_0. No additional assumption other than the independence of Y_t with respect to the past (represented in what follows by I_{t-1}) governs the relationship between Y_t and X_t. More formally, we assume that $Y_1, \ldots, Y_n$ and $X_0, \ldots, X_{n-1}$ have continuous distributions such that:

$$(1) \qquad Y_t \text{ is independent of } I_{t-1}, \text{ for each } t = 1, \ldots, n;$$

$$(2) \qquad P[Y_t > b_0] = P[Y_t < b_0], \text{ for } t = 1, \ldots, n.$$

These assumptions leave open the possibility of feedback from Y_t to current and future values of the X-variable, without specifying the form of feedback or any other property of the X-process; in addition, the variables Y_t need not be normal nor identically distributed. In what follows we shall also consider the stronger assumption:

$$(3) \qquad Y_1, \ldots, Y_n \text{ have continuous distributions symmetric about } b_0.$$

Clearly, the latter assumption implies (2), but the converse is not true.

What distinguishes these assumptions from those in our previous work is the presence of the parameter b_0, the median of the variables Y_t, $t = 1, \ldots, n$. To obtain methods applicable when b_0 is unknown, we need first to consider the case where this nuisance parameter is known. In so far as b_0 is known, the techniques of Campbell and Dufour (1995) can readily be modified to yield exact nonparametric tests as follows. The basic building blocks of these statistics are the simple products $Z_t(b) = (Y_t - b)X_{t-1}$, $t = 1, \ldots, n$, where b will be taken to be b_0 when the median is known as in this section, or an estimate when it is unknown as in the next section of the paper. Let $u(z) = 1$, if $z \geq 0$, and $u(z) = 0$ for $z < 0$. We first introduce an analogue of the t-statistic given by the sign statistic

$$(4) \qquad S_g(b) = \sum_{t=1}^{n} u[(Y_t - b)g_{t-1}],$$

where $g_t = g_t(I_t)$, $t = 0, \ldots, n-1$, is a sequence of measurable functions of the information vector I_t. The functions $g_t(\cdot)$ allow one to consider various (possibly nonlinear) transformations of the data, provided g_t depends only on past and current values of X_τ and Y_τ $(\tau \leq t)$. The role of such transformations is important in applications, as will be seen in Section 5; in practice, simple forms of $g_t(\cdot)$ such as the computation of restricted medians may be preferred. This point is elaborated in Campbell and Dufour (1995).

Under the further assumption that each Y_t has a continuous symmetric distribution, that is, under (3), it is natural to use ranks as well. We will consider here aligned signed rank statistics with general form:

$$(5) \qquad SR_g(b) = \sum_{t=1}^{n} u[(Y_t - b)g_{t-1}]R_t^+(b)$$

where $R_t^+(b)$ in $SR_g(b)$ is the rank of $|Y_t - b|$, i.e. $R_t^+(b) = \Sigma_{j=1}^n u(|Y_t - b| - |Y_j - b|)$ the rank of $|Y_t - b|$ when $|Y_1 - b|, \ldots, |Y_n - b|$ are put in ascending order.

Consider first the case where the median b_0 of the variables Y_t, $t = 1, \ldots, n$, is known. The finite-sample distributions of $S_g(b_0)$ and $SR_g(b_0)$ under general conditions is given by the following proposition. By contrast with the usual definitions of Wilcoxon-type statistics, where the absolute ranks would be based on the products $(Y_t - b)g_{t-1}$, it should be noted that in the definition of the statistics $SR_g(b_0)$ the absolute ranks are defined with respect to $|Y_1 - b_0|, \ldots, |Y_n - b_0|$, which are mutually independent according to (1).

PROPOSITION 1. *Let* $Y = (Y_1, \ldots, Y_n)'$ *and* $X = (X_0, \ldots, X_{n-1})'$ *be two* $n \times 1$ *random vectors which satisfy Assumptions* (1) *and* (2). *Suppose further that* $P[Y_t - b_0 = 0] = 0$ *for* $t = 1, \ldots, n$, *and let* $g_t = g_t(I_t)$, $t = 0, \ldots, n-1$, *be a sequence of measurable functions of* I_t *such that* $P[g_t = 0] = 0$ *for* $t = 0, \ldots, n-1$.

(a) *Then the sign statistic* $S_g(b_0)$ *defined by* (4) *follows a* $\mathrm{Bi}(n, 0.5)$ *distribution, that is,* $P[S_g(b_0) = x] = \binom{n}{x}(1/2)^n$ *for* $x = 0, 1, \ldots, n$, *where* $\binom{n}{x} = n!/[x!(n-x)!]$.

(b) *If Assumption* (3) *also holds, then the signed rank statistic* $SR_g(b_0)$ *defined by* (5) *is distributed like* $W_n = \Sigma_{t=1}^n tB_t$, *where* $B_1, \ldots, B_n$ *are independent Bernoulli variables such that* $P[B_t = 0] = P[B_t = 1] = 1/2$, $t = 1, \ldots, n$.

These distributional results hold under very general conditions. It is important to keep in mind that sign statistics are the only statistics which can produce valid tests for hypothesis about a median under sufficiently general distributional assumptions; more precisely, any test with level α when $Y_1, \dots, Y_n$ are independent with distributions symmetric about a common median b_0 must have level α conditional on the vector of absolute values $|Y - b_0| = (|Y_1 - b_0|, \dots, |Y_n - b_0|)$, that is, must be a sign test (see Lehmann and Stein 1949, Pratt and Gibbons 1981, pp. 233–234, or Dufour and Hallin 1991). In this framework the nature of the distribution of each Y_t is left open; there are no assumptions concerning the existence of moments; heteroskedasticity of unknown form is permitted; the nature of the feedback mechanism between Y_t and current and future values of X_{t+s} ($s \geq 0$) is not specified. As long as Y_t has median b_0 and is independent of the past, the sign statistic $S_g(b_0)$ follows a binomial distribution $\mathrm{Bi}(n, 0.5)$. The Wilcoxon variate W_n has been extensively tabled (see, for example, Wilcoxon, Katti and Wilcox 1970), and the normal approximation with $E(W_n) = n(n+1)/4$ and $\mathrm{Var}(W_n) = n(n+1)(2n+1)/24$ works well even for small values of n (for further discussion, see Lehmann 1975). The powers of the tests $S_g(b_0)$ and $SR_g(b_0)$, with $g_t = X_t$ and $b_0 = 0$, relative to standard regression-based tests have been investigated by simulation in Campbell and Dufour (1995) for two models with feedback. The nonparametric tests displayed remarkable power, generally outperforming the t-statistic applied with correct critical values in the presence of nonnormal disturbances and/or heteroskedasticity and having comparable power with homoskedastic normal disturbances. We now need to deal with the fact that the centering parameter b_0 is generally unknown.

3. ORTHOGONALITY TESTS WITH UNKNOWN DRIFT PARAMETER

In this section we adapt a general procedure introduced in Dufour (1990) for a parametric model to the nonparametric setup described in the previous section, in order to obtain exact tests of the hypothesis that a variable is independent of past information in the presence of the unknown nuisance parameter b_0. A straightforward response to the problem of the unknown median in the spirit of the previous section is to estimate the parameter using the sample median $\tilde{b}_0$ of the observations Y_t, $t = 1, \dots, n$, and consider the statistics $S_g(\tilde{b}_0)$ and $SR_g(\tilde{b}_0)$. These aligned sign and signed rank statistics are of independent interest and their power performance will be considered in the simulation exercises conducted in the next section of this paper. However, we do not have a finite-sample theory for these statistics in the general framework studied here; and indeed it appears quite doubtful that such a theory is even possible for such statistics.

To obtain provably valid finite-sample procedures, we shall adopt a three-stage approach: First, we find an exact confidence set for the nuisance parameter b_0 which is valid at least under the null hypothesis. Second, corresponding to each value b in the confidence set, we construct a nonparametric test based on the methods of the previous section. Third, the latter are combined with the confidence set for b_0 using Bonferroni's inequality to obtain valid nonparametric tests at the desired level α.

Let $J(\alpha_1)$ be a confidence set for b_0 with level $1 - \alpha_1$ (where $\alpha_1 < \alpha$), which is valid either on the assumption that Y_t has median b_0 for $t = 1, \dots, n$ or that Y_t is symmetric about b_0 for each t. Different approaches to the construction of $J(\alpha_1)$

based on counting procedures will be discussed below. On any approach, we have $P[b_0 \in J(\alpha_1)] \geq 1 - \alpha_1$.

For any $b \in J(\alpha_1)$, we now consider the aligned sign and signed rank statistics $S_g(b)$ and $SR_g(b)$. Under different hypotheses, Proposition 1 established the exact distribution of $S_g(b_0)$ and $SR_g(b_0)$. For any $0 \leq \alpha \leq 1$, let $\bar{S}_g(\alpha)$ and $\overline{SR}_g(\alpha)$ be the critical values of the corresponding right one-sided tests with nominal level α, i.e. $\bar{S}_g(\alpha)$ and $\overline{SR}_g(\alpha)$ are the smallest points (in the extended real numbers $\bar{\mathbb{R}}$) such that

$$P\left[S_g(b_0) > \bar{S}_g(\alpha)\right] \leq \alpha, P\left[SR_g(b_0) > \overline{SR}_g(\alpha)\right] \leq \alpha. \tag{6}$$

Since $S_g(b_0)$ and $SR_g(b_0)$ have discrete distributions, it may not be possible to make the tail areas in (6) equal to α. The following proposition establishes probability bounds for the events that $S_g(b)$ is significant (or nonsignificant) at an appropriate level for all $b \in J(\alpha_1)$ for both one-sided and two-sided tests, and similarly for $SR_g(b)$.

PROPOSITION 2. *Let $Y = (Y_1, \ldots, Y_n)'$ and $X = (X_0, \ldots, X_{n-1})'$ be two $n \times 1$ random vectors satisfying the assumptions* (1) *and* (2) *with $P[Y_t = 0] = 0$ for $t = 1, \ldots, n$, and let $g_t = g_t(I_t)$, $t = 0, \ldots, n-1$, be a sequence of measurable functions of I_t such that $P[g_t = 0] = 0$ for $t = 0, \ldots, n-1$. Let also $S_g(b)$, $SR_g(b)$, $\bar{S}_g(\cdot)$ and $\overline{SR}_g(\cdot)$ be defined as in* (4), (5) *and* (6), *let $\tilde{S}_g(\delta) = n - \bar{S}_g(1-\delta)$ and $\widetilde{SR}(\delta) = (n(n+1)/2) - \overline{SR}_g(1-\delta)$ for any $0 \leq \delta \leq 1$, and choose α, α_1, α_2, and α_3 in the interval* $[0,1]$ *such that $0 \leq \alpha_2 \leq \alpha - \alpha_1 \leq \alpha + \alpha_1 \leq \alpha_3 \leq 1$.*

(a) *If $J(\alpha_1)$ is a confidence set for b_0 such that $P[b_0 \in J(\alpha_1)] \geq 1 - \alpha_1$, then*

$$P\left[S_g(b) > \bar{S}_g(\alpha_2), \forall b \in J(\alpha_1)\right] \leq \alpha_1 + \alpha_2 \leq \alpha, \tag{7a}$$

$$P\left[M - S_g(b) > \bar{S}_g(\alpha_2), \forall b \in J(\alpha_1)\right] \leq \alpha_1 + \alpha_2, \tag{7b}$$

(7c)

$$P\left[\max\{S_g(b), M - S_g(b)\} > \bar{S}_g(\alpha_2/2), \forall b \in J(\alpha_1)\right] \leq \alpha_1 + \alpha_2,$$

$$P\left[S_g(b) < \tilde{S}_g(\alpha_3), \forall b \in J(\alpha_1)\right] \leq 1 - (\alpha_3 - \alpha_1) \leq 1 - \alpha, \tag{7d}$$

$$p\left[M - S_g(b) < \tilde{S}_g(\alpha_3), \forall b \in J(\alpha_1)\right] \leq 1 - (\alpha_3 - \alpha_1), \tag{7e}$$

(7f)

$$P\left[\max\{S_g(b), M - S_g(b)\} < \tilde{S}_g(\alpha_3/2), \forall b \in J(\alpha_1)\right] \leq 1 - (\alpha_3 - \alpha_1),$$

where $M = n$.

(b) *If the additional Assumption* (3) *holds and $K(\alpha_1)$ is a confidence set for b_0 such that $P[b_0 \in K(\alpha_1)] \geq 1 - \alpha_1$, then the inequalities* (7a) *to* (7f) *also hold with $S_g(b)$ replaced by $SR_g(b), \bar{S}_g(\cdot)$ by $\overline{SR}_g(\cdot), \tilde{S}_g(\cdot)$ by $\widetilde{SR}_g(\cdot), J(\alpha_1)$ by $K(\alpha_1)$, and M by $M' = n(n+1)/2$.*

Under the maintained hypothesis (1) and (2), or (1) through (3), the probability bounds established by the proposition suggest the following bounds test for the hypothesis that Y_t is orthogonal to past information I_{t-1}, for $t = 1, \ldots, n$. Using the notations adopted in Proposition 2, define

(8a) $$Q_L(S_g) = \text{Inf}\{S_g(b)\colon b \in J(\alpha_1)\}, \quad Q_L(SR_g) = \text{Inf}\{SR_g(b)\colon b \in K(\alpha_1)\},$$

(8b) $$Q_U(S_g) = \text{Sup}\{S_g(b)\colon b \in J(\alpha_1)\}, \quad Q_U(SR_g) = \text{Sup}\{SR_g(b)\colon b \in K(\alpha_1)\}.$$

From Proposition 2(a), it is clear that

(8c) $$P\big[Q_L(S_g) > \bar{S}_g(\alpha_2)\big] \le \alpha, P\big[Q_U(S_g) < \tilde{S}_g(\alpha_3)\big] \le 1 - \alpha,$$

where it is easy to see that the conjunction of the events $Q_L(S_g) > \bar{S}_g(\alpha_2)$ and $Q_U(S_g) < \tilde{S}_g(\alpha_3)$ has probability zero, and similarly for $Q_L(SR_g)$ and $Q_U(SR_g)$. Consequently, a reasonable right one-sided test would reject the hypothesis of conditional independence if $Q_L(S_g) > \bar{S}_g(\alpha_2)$ (alternatively, if $Q_L(SR_g) > \overline{SR}_g(\alpha_2)$), and would accept it if $Q_U(S_g) < \tilde{S}_g(\alpha_3)$ [alt., $Q_U(SR_g) < \widetilde{SR}_g(\alpha_3)$]; otherwise, we consider the test inconclusive. For example, for $\alpha = 0.05$ and $\alpha_1 = 0.025$, the null is rejected if $S_g(b)$ is significant at level 0.025 [$S_g(b) > S_g(0.025)$] for each b in a 97.5% confidence interval for b_0, and accepted if $S_g(b)$ is never significant at level 0.075 over the confidence interval. According to the proposition, the probability of a Type I error is bounded from above by α, whereas the probability of accepting the true hypothesis according to this procedure is bounded from above by $1 - \alpha$. It is clear that one should normally set $\alpha_2 = \alpha - \alpha_1$ and $\alpha_3 = \alpha + \alpha_1$.

To obtain a left one-sided test of the model described by the assumptions of Proposition 2, one can proceed in exactly the same way with $S_g(b)$ replaced by $M - S_g(b) = n - S_g(b)$, and $SR_g(b)$ by $M' - SR_g(b)$; e.g., the rejection region of the sign test is $\text{Inf}\{M - S_g(b)\colon b \in J(\alpha_1)\} > \bar{S}_g(\alpha_2)$ and the acceptance region $\text{Sup}\{M - S_g(b)\colon b \in J(\alpha_1)\} < \tilde{S}_g(\alpha_3)$. Finally, we obtain a two-sided sign test with level α by considering

$$QB_L(S_g) = \text{Inf}\{\max\{S_g(b), M - S_g(b)\}\colon b \in J(\alpha_1)\},$$

$$QB_U(S_g) = \text{Sup}\{\max\{S_g(b), M - S_g(b)\}\colon b \in J(\alpha_1)\},$$

and then taking $QB_L(S_g) > \bar{S}_g(\alpha_2/2)$, and $QB_U(S_g) < \tilde{S}_g(\alpha_3/2)$ as the rejection and acceptance regions, respectively. The procedures are similar for the Wilcoxon-type tests.

It remains to discuss the construction of the confidence set $J(\alpha_1)$ for b_0, which should be valid at least under the null hypothesis. If Y_t is assumed to have median b_0, the order statistics $Y_{(1)}, \ldots, Y_{(n)}$ of the random sample $Y_1, \ldots, Y_n$ can be used to construct a confidence interval for b_0. Let B be a binomial random variable with number of trials n and probability of success equal to 0.5. Choose k the largest integer such that $P[B \le k] \le \alpha/2$. Then $[Y_{(k+1)}, Y_{(n-k)}]$ is a confidence interval for

b_0 with level $1-\alpha$ (see Hettmansperger 1984, pp. 12–15 for details). On the other hand, if the distributions of the Y_t's are symmetric, one can obtain a (tighter) confidence interval for b_0 by considering the $n(n+1)/2$ Walsh averages defined by $(Y_i+Y_j)/2$, $1 \le i \le j \le n$ (Hettmansperger 1984, pp. 38–41 for details). One difficulty with using Walsh averages, particularly in simulations, is the large number of averages that must be computed and then ordered. For $n=200$, there are some 20000 Walsh averages to be ordered. In what follows, we only use the method based on the binomial distribution to derive the confidence interval for b_0 even though the underlying distributions may be symmetric. An example of the construction of a nonparametric confidence interval is given in Section 5.

To address the issue of the power of the procedure proposed above, it is instructive to consider the following linear model:

$$Y_t = \beta_0 + \beta_1 X_{t-1} + e_t,\ t = 1,\dots,n \tag{9}$$

where e_t has the same properties as Y_t in (1) and (2) [or (1) through (3)] with median 0. Suppose that we wish to test the null hypothesis that $\beta_1 = 0$ against the alternative that $\beta_1 \ne 0$. If β_1 is in fact zero, then Y_t satisfies (1) and (2) [or (1) through (3)], and the bounds testing procedure will have the properties described above; in particular, the probability of rejecting the null will be at most as large as α. Now suppose that β_1 is not equal to zero and let $m_t = m(X_t)$ be the median of X_t. To continue the illustration, if we assume that $m(X_t)$ is constant, i.e. $m_t = m(X)$ for all t, then $J(\alpha_1)$ is a confidence set for $b_0 = \beta_0 + \beta_1 m(X)$ instead of β_0. When $g_t = X_t$, it follows that the basic building block of the nonparametric statistics introduced in the previous section can be rewritten:

$$\begin{aligned} Z_t(b_0) &= (Y_t - b_0)X_{t-1} = (\beta_0 + \beta_1 X_{t-1} + e_t - b_0)X_{t-1} \\ &= \beta_1[X_{t-1} - m(X)]X_{t-1} + e_t X_{t-1},\ t = 1,\dots,n. \end{aligned}$$

If we assume that X_{t-1} and e_t have symmetric distributions, it is easy to see that $Z_t(b_0)$ will have median 0 even if $\beta_1 \ne 0$, since e_t is independent of X_{t-1} by assumption. Accordingly, a sign statistic based on $Z_t(b_0)$ will have virtually no power to detect $\beta_1 \ne 0$, no matter the size of β_1.

This general problem, suggested by the previous illustration, can be resolved by altering the definition of $Z_t(b_0)$. Let us replace X_{t-1} by $X_{t-1} - m(X)$ in $Z_t(b_0)$:

$$\begin{aligned} Z_t(b_0) &= (Y_t - b_0)[X_{t-1} - m(X)] \\ &= \beta_1[X_{t-1} - m(X)]^2 + e_t[X_{t-1} - m(X)]. \end{aligned}$$

We see now that the median of $Z_t(b_0)$ is clearly shifted toward the right or left depending on whether $\beta_1 > 0$ or $\beta_1 < 0$. In practice, of course, we will need to replace $m(X)$ by an estimator $\hat{m}_{t-1}$. Further, in order to have $g_t = g_t(I_t)$, $\hat{m}_t$ should only depend on observations up to time t, e.g., $\hat{m}_t = \text{med}(X_0, X_1, \dots, X_t)$ the sample median of $X_0, \ldots, X_t$. This suggests replacing $g_t = X_t$ by

$$g_t = [X_t - \hat{m}_t], t = 0, \ldots, n-1, \tag{10}$$

where $\hat{m}_t$ is an estimate of $m(X_t)$ that is a function of I_t. Of course, if X_t is nonstationary, other centering functions $\hat{m}_t$ may be more appropriate. There is no need here to assume that the median of X_t is constant.

It is straightforward to apply the above results to test the random walk hypothesis in the presence of a drift. Here we should mention that there are tests of the random walk hypothesis based on a transformation of the data involving signs (level crossings); see Granger and Hallman (1991) and Burridge and Guerre (1995). But these statistics are more specific than ours (involving for example the number of times the series changes sign), and in contrast to the results presented in this paper only the asymptotic distributions of the test statistics are at best established. Furthermore, from the asymptotic results of Burridge and Guerre (1995), it is clear that their tests are not distribution-free in finite samples. Consider the model in the following form:

$$Y_t - Y_{t-1} = \beta_0 + \beta_1 Y_{t-1} + e_t, \qquad t = 1, \ldots, n. \tag{11}$$

The null hypothesis of a random walk is then equivalent to $\beta_1 = 0$, with $\beta_1 < 0$ under the alternative of stationarity. This null hypothesis should not be confused *strictu senso* with the unit root hypothesis. Appropriate nonparametric statistics to consider in this context are given by:

$$S_g(b) = \sum_{t=1}^{n} u[(Y_t - Y_{t-1} - b) g_{t-1}], \tag{12}$$

$$SR_g(b) = \sum_{t=1}^{n} u[(Y_t - Y_{t-1} - b) g_{t-1}] R_t^+(b), \tag{13}$$

where $R_t^+(b)$ is the rank of $|Y_t - Y_{t-1} - b|$ among $|Y_\tau - Y_{\tau-1} - b|$, $\tau = 1, \ldots, n$ and g_t is given by

$$g_t = [Y_t - \hat{m}_t(Y)], t = 0, \ldots, n-1, \tag{14}$$

with $\hat{m}_t(Y) = \text{med}(Y_0, \ldots, Y_t)$, the sample median of Y_s, $s = 0, \ldots, t$. Once a confidence interval for β_0 is determined under the null, the bounds procedures are defined precisely as before. Against the alternative of stationarity ($\beta_1 < 0$), the most appropriate test here is a left one-sided test with rejection region of the form: $S_g(b) < \bar{S}_g(1 - \alpha_2)$ for all $b \in J(\alpha_1)$ [or equivalently, $M - S_g(b) > \bar{S}_g(\alpha_2)$ for all $b \in J(\alpha_1)$]. The power of these procedures applied in the random walk context will be assessed in the next section.

4. A SIMULATION STUDY OF TWO EXAMPLES

The specifications of model (9) considered in this section correspond to those studied in Campbell and Dufour (1995) with the addition of the intercept β_0. The first example is drawn from Mankiw and Shapiro (1986). X_t is assumed to follow a stationary autoregressive process given by

$$X_t = \theta_0 + \theta_1 X_{t-1} + \epsilon_t, t = 1, \ldots, n. \tag{15}$$

where the ϵ_t are assumed to be mutually independent and each ϵ_t is independent of X_{t-j}, $j \geq 1$; the disturbances e_t and ϵ_t are also assumed to follow a bivariate normal distribution with correlation coefficient ρ. The results of the simulations presented in our previous study elaborated the basic theme of Mankiw and Shapiro, who found that the usual t-test considerably over-rejects the null hypothesis when ρ and θ_1 are close to one and asymptotic critical points are used. The simulations presented in this section contrast the power of the nonparametric bounds procedure proposed above with the t-statistic based on standard regression procedures. The organization of this Monte Carlo study follows that of our earlier work which investigated the performance of both parametric and nonparametric procedures under different data generating mechanisms, including nonnormal and heteroskedastic patterns. Since these processes are essentially those of our previous work with the addition of an intercept term, we focus here primarily on issues related to the application of the nonparametric procedures introduced in the previous section, and direct the interested reader to Campbell and Dufour (1995) for a more thorough presentation of the details of the models studied. The parameter values are $\theta_1 = 0.99$, $\rho = 0.9$ and $\beta_0 = \theta_0 = 0.0$. In this study, sample sizes $n = 100, 200$ are considered. Finally, there are 1000 replications in each experiment.

In the application of the bounds procedure, there is an evident tradeoff between the width of the confidence interval $J(\alpha_1)$ and the significance level $\alpha_2 = \alpha - \alpha_1$ of the tests based on elements of $J(\alpha_1)$. For $n = 200$, the following confidence intervals based on counting procedures associated with the binomial distribution are considered: $[Y_{(80)}, Y_{(121)}]$, $[Y_{(82)}, Y_{(119)}]$, $[Y_{(83)}, Y_{(118)}]$ and $[Y_{(85)}, Y_{(116)}]$, where $Y_{(k)}$ is the kth order statistic, corresponding, respectively, to α_1 equal to 0.4%, 0.9%, 1.3% and 2.8%. It should be noted that there is not a sizable decrease in the width of the confidence interval as its significance decreases, a reflection of the fact that the tails of the binomial distribution are relatively thin.

With α fixed at 0.05 and for sample size $n = 200$, there is a different bounds test corresponding to each of these confidence intervals $J(\alpha_1)$, where α_1 is 0.003, 0.009, 0.013 or 0.028. The construction of the statistics $S_g(b)$ and $SR_g(b)$ for each $b \in J(\alpha_1)$, with g defined as in (10), does not vary with α_1. According to the bounds procedure denoted SB [alternatively, SRB], the null is rejected if $S_g(b)$ [alt., $SR_g(b)$] is significant at level $\alpha - \alpha_1$ for each b in $J(\alpha_1)$; the null is accepted if no $S_g(b)$ [alt., $SR_g(b)$] is significant at level $\alpha + \alpha_1$; otherwise, the procedure is considered inconclusive. The results of these procedures for the Mankiw-Shapiro model in the case of normal disturbances are given in Table 1. Overall, the results suggest that it is better to take a wider confidence interval for β_0 in the first step of the bounds procedure in order to expand the critical region of the nonparametric statistics used

TABLE 1
MANKIW-SHAPIRO MODEL WITH NORMAL DISTURBANCES*: $\rho = 0.9$, $\theta_1 = 0.99$, $n = 200$.
COMPARISONS BETWEEN BOUNDS TESTS

Testing Strategy			Bounds Tests			
			SB		*SRB*	
α_1	α_2	β_1	Reject	Accept	Reject	Accept
2.8	2.2					
		0.00	0.1	67.6	0.4	63.5
		0.05	6.2	35.3	14.7	29.5
		0.07	17.2	23.1	28.5	15.7
1.3	3.7					
		0.00	0.1	69.5	0.4	62.9
		0.05	6.6	36.8	16.9	30.5
		0.07	18.3	27.1	31.9	16.5
0.9	4.1					
		0.00	0.1	67.4	0.5	62.1
		0.05	8.0	34.8	17.1	29.6
		0.07	20.1	22.4	32.0	16.4
0.4	4.6					
		0.00	0.0	71.5	0.5	60.0
		0.05	7.1	37.7	16.6	29.2
		0.07	19.0	23.8	30.2	16.3

* $\beta_0 = \theta_0 = 0.0$. Probabilities are given in percentages. A nonparametric confidence interval $J(\alpha_1)$ with level $1 - \alpha_1$ is first constructed for β_0. The null ($\beta_1 = 0$) is rejected if for all $b \in J(\alpha_1)$ the nonparametric test based on b is significant at level α_2; the null is accepted if no such test is significant at level $\alpha_3 = 0.05 + \alpha_1$; otherwise, the procedure is inconclusive. *SB* refers to the sign procedure, SRB to the Wilcoxon. The level of each strategy is bounded by 0.05. See text for further details.

in the second stage. There is a clear gain in power: when β_1 is 0.05, there is a 30% increase in power for the procedure based on the sign statistic and a 15% gain for the Wilcoxon in passing from a procedure based on the narrowest confidence interval to the the confidence interval given by $\alpha_1 = 0.009$. There does not appear to be any additional gain in power available from reducing α_1 even further. Accordingly, in the comparative studies for $n = 200$ presented in Tables 3 and 4, the results for the testing strategy represented by $\alpha_1 = 0.009$ will be pursued. A similar analysis was conducted to investigate the impact on the power of the nonparametric procedures obtained by varying α_1 when $n = 100$. The results (not reported here) also suggest that power is increased somewhat by taking a wide confidence interval for the unknown intercept parameter ($\alpha_1 = 0.007$), but that there appear to be no further gains in power associated with smaller α_1. The results of the bounds tests given in Table 2 are obtained for this α_1.

In what follows, we also study the performance of the following statistics based on the sample median $\tilde{b}_0$ of $Y_1, \ldots, Y_n$:

$$S_g(\tilde{b}_0) = \sum_{t=1}^{n} u\left[(Y_t - \tilde{b}_0)g_{t-1}\right], \tag{16}$$

$$SR_g(\tilde{b}_0) = \sum_{t=1}^{n} u\left[(Y_t - \tilde{b}_0)g_{t-1}\right]R_t^+(\tilde{b}_0), \tag{17}$$

TABLE 2
MANKIW-SHAPIRO MODEL; VARIOUS TYPES OF DISTURBANCES*: $\rho = 0.9$, $\theta_1 = 0.99$, $n = 100$.
LEVEL AND POWER COMPARISONS

	t-test			Median-Estimate Tests		Bounds Tests			
						SB		*SRB*	
β_1	Asymptotic	Size-Corrected (specific)	Size-Corrected (model)	$S(\tilde{b}_0)$	$SR(\tilde{b}_0)$	Reject	Accept	Reject	Accept
Cauchy Distribution									
0.00	13.8	5.0	3.4	2.3	4.0	0.0	65.6	0.8	67.7
0.07	40.0	29.9	26.8	64.2	67.5	29.8	7.5	36.9	10.1
0.10	64.4	48.9	43.1	74.2	76.7	37.5	4.9	44.9	6.0
$t(3)$ Distribution									
0.00	18.9	5.0	2.8	1.6	3.8	0.0	67.9	0.3	58.9
0.07	28.2	14.7	11.9	25.1	30.0	6.7	36.9	12.0	32.2
0.10	50.0	36.3	30.8	42.7	46.5	13.7	24.9	19.7	21.7
Normal Distribution									
0.00	18.4	5.0	2.8	2.6	4.2	0.2	70.5	1.0	55.6
0.07	26.2	14.0	11.0	14.9	22.3	2.2	48.6	6.5	36.2
0.10	43.5	32.4	28.1	26.0	35.5	6.3	35.7	12.2	24.8
Lognormal Distribution									
0.00	19.1	5.0	2.7	2.1	9.3	0.1	67.6	0.8	53.1
0.07	29.9	17.3	13.6	40.9	45.8	14.3	19.6	20.8	20.5
0.10	51.3	37.6	32.2	59.7	62.1	24.8	9.3	34.2	11.2

* $\beta_0 = \theta_0 = 0.0$. Probabilities are given in percentages. Empirical critical points are used in power calculations for both the size-corrected t-test: when $\beta_1 = 0$, the rejection frequency for the specific size correction is 5.0% by construction. The model-correction critical values are obtained when $\rho = \theta_1 = 0.9999$ and the disturbances are normal. The statistics $S(\tilde{b}_0)$ and $SR(\tilde{b}_0)$ are defined by (16) and (17), with g_t given by (10). The bounds tests, *SB* and *SRB*, are described in Table 1, with $\alpha_1 = 0.7\%$ and $\alpha_2 = 4.3\%$.

TABLE 3
MANKIW-SHAPIRO MODEL; VARIOUS TYPES OF DISTURBANCES*: $\rho = 0.9$, $\theta_1 = 0.99$, $n = 200$.
LEVEL AND POWER COMPARISONS

	t-test			Median-Estimate Tests		Bounds Tests			
						SB		*SRB*	
β_1	Asymptotic	Size-Corrected (specific)	Size-Corrected (model)	$S(\tilde{b}_0)$	$SR(\tilde{b}_0)$	Reject	Accept	Reject	Accept
Cauchy Distribution									
0.00	10.0	5.0	2.9	3.1	4.4	0.3	59.8	0.8	68.0
0.03	30.5	22.6	18.9	78.0	78.4	50.6	3.7	55.7	5.8
0.05	61.0	47.4	38.4	86.8	89.6	62.5	1.1	69.4	1.9
$t(3)$ Distribution									
0.00	14.3	5.0	2.1	2.7	4.6	0.0	65.9	0.6	66.5
0.03	19.7	10.2	6.2	21.7	26.8	6.9	39.3	10.4	38.8
0.05	47.4	34.5	28.3	45.5	53.5	19.6	19.0	28.2	21.2
Normal Distribution									
0.00	14.4	5.0	2.3	2.7	4.6	0.1	67.4	0.5	62.1
0.03	14.9	7.3	5.2	12.8	18.0	2.5	52.6	4.7	46.3
0.05	42.9	30.8	24.1	29.4	37.4	8.0	34.8	17.1	29.6
Lognormal Distribution									
0.00	15.1	5.0	1.7	2.8	14.1	0.4	66.7	1.5	54.4
0.03	17.5	9.4	5.2	39.1	48.1	17.0	25.6	26.2	28.3
0.05	46.3	32.8	27.5	67.6	70.7	40.4	10.4	51.1	12.0

* $\beta_0 = \theta_0 = 0.0$. Probabilities are given in percentages. See Table 2 for details.

TABLE 4
MANKIW-SHAPIRO MODEL; HETEROSCEDASTIC DISTURBANCES*: $\rho = 0.9$, $\theta_1 = 0.99$, $n = 200$.
LEVEL AND POWER COMPARISONS

			Median-Estimate Tests		Bounds Tests			
					SB		*SRB*	
β_1	*t*-test	*wm*-test	$S(\tilde{b}_0)$	$SR(\tilde{b}_0)$	Reject	Accept	Reject	Accept
Break at $t = 100$								
0.00	18.5	9.6	3.0	3.5	0.0	67.7	0.5	76.6
	5.0	5.0						
	3.9	3.4						
	0.0	0.0						
0.05	48.0	39.1	31.9	42.4	9.8	30.9	22.2	33.4
	31.6	33.4						
	30.6	30.0						
	0.0	11.5						
Breaks at $t = 75, 150$								
0.00	18.7	8.3	3.9	4.1	0.1	66.0	0.9	78.4
	5.0	5.0						
	5.0	3.3						
	0.0	0.0						
0.05	47.7	36.4	30.4	40.9	10.5	32.9	23.1	36.1
	31.3	31.5						
	31.3	27.6						
	0.0	8.7						
Linear								
0.00	17.1	10.0	3.5	4.1	0.0	70.4	0.8	82.3
	5.0	5.0						
	3.9	1.4						
	0.0	0.1						
0.05	46.1	38.2	32.1	45.7	8.7	34.5	21.7	35.6
	32.9	29.3						
	29.8	20.7						
	0.0	9.5						
Exponential								
0.00	86.0	11.4	4.9	5.1	0.3	74.7	2.8	92.1
	5.0	5.0						
0.30	84.9	14.1	46.0	42.0	13.0	27.7	32.0	51.3
	6.7	4.9						

* $\beta_0 = \theta_0 = 0.0$. In the Break model, the variance of the disturbances jumps from 1 to 16 at $t = 100$; in the two-break model, the variance jumps first by 16 then by 64 at the indicated points; in the linear (alt., exponential model), the variance grows linearly (alt., exponentially) with time. The median-estimate tests are given by (16) and (17); the *wm* test is described in the text. For break and linear heteroscedasticity models, the entries under the asymptotic percentage rejections for the *t*-test (alt., *wm*-test where indicated) represent rejections according to different empirical critical values for each statistic determined by: (i) specific model; (ii) two-break model (linear model with $\rho = \theta_1 = 0.9999$); (iii) exponential model. For exponential heteroscedasticity, only asymptotic and specific percentage rejections are reported.

where g_t is the usual centering function given by (10) and R_t^+ defined in (5). These are simply aligned sign and signed rank statistics, which give rise to what are termed median-estimate tests in the account that follows, based on a reasonable point estimate of β_0. In what follows, critical points associated with the $\mathrm{Bi}(n, 0.5)$ distribution are used for (16), while those for (17) come from the distribution of the Wilcoxon variate W_n. We do not have analytical results for the distribution of these statistics, but is easy to see that the conditional distributions of $S_g(\tilde{b}_0)$ and $SR_g(\tilde{b}_0)$ given $|Y - b_0| \equiv (|Y_1 - b_0|, \ldots, |Y_n - b_0|)'$ depend on $|Y - b_0|$, (e.g., by considering the special case where $g_t = 1$). Furthermore, for the general nonparametric hypothesis (1) and (2) [or (1) through (3)], it is well known that any valid test of such an hypothesis should be a "randomization text," that is, a test whose level is the same conditional on any value of the (unknown) vector $|Y - b_0|$ (see Lehmann and Stein 1949, and Pratt and Gibbons 1981, pp. 216–222). Critical values for such tests are determined by assigning random signs to the absolute values $|Y_t - b_0|$, $t = 1, \ldots, n$ (according to a uniform distribution) and will generally depend on $|Y - b_0|$. This characterization does not apply to the median-estimate tests considered above.

In Tables 2 and 3, the power of the t-test applied with both asymptotic and size-corrected critical values is compared with median-estimate tests and nonparametric bounds procedures for sample size $n = 100$ (with $\alpha = 0.05$ and $\alpha_1 = 0.007$) for various types of disturbances. Two types of size correction are considered. In the first (specific size-correction), we use the empirical critical values obtained when $\beta_1 = 0$, $\rho = 0.9$ and $\theta_1 = 0.99$. In the second (model size-correction), we use the larger critical values associated with the specification $\rho = \theta_1 = 0.9999$ with normal disturbances to emphasize the point that, ultimately, the correct analysis of power must be relative to all potential specifications of the model, compatible with the null hypothesis (1) and (2) [or (1) through (3)]. Even with these corrections, the power comparisons are biased in favor of the parametric tests because the (unknown) correct critical values should be greater than the ones used. Each of these size corrections, moreover, remains specific to the particular distributions considered and so none yields a truly distribution-free test. The "size-corrected tests" should not be viewed as alternative tests (because they are not feasible in practice, especially under the general assumptions (1) and (2)), but as theoretical benchmarks to which truly distribution-free tests may be compared. In particular, we would like to see whether the distribution-free procedures have power not too far below these benchmarks.

First, as expected, it is clear from the results in Tables 2 and 3 that the asymptotic t-tests do not have the stated level. Interestingly, the level distortion is especially strong for the normal and lognormal distributions. Second, it is quite striking that the bounds procedure using Wilcoxon statistics in the second stage outperforms the model size-corrected t-test in the case of Cauchy disturbances and is comparable in power for alternatives close to the null when the disturbances are $t(3)$. Moreover, the bounds procedure based on the sign statistic is comparable in power to the model size-corrected t-test for Cauchy and lognormal disturbances. A further interesting result is that the median-estimate tests do not over-reject under the null, except in the case of asymmetric lognormal disturbances where the Wilcoxon statistic should not be applied; here the sign-based test appears to have empirical level bounded by 5%. Both these tests, moreover, outperform the parametric tests in having comparable (better, in the case of the Wilcoxon variant) power to the size-corrected t-test in the case of normal disturbances, while outperforming by a wide margin the size-corrected t-tests for both the fat-tailed disturbances. When the sample size is increased to $n = 200$, the relative performance of the two bounds procedures improves considerably. Both bounds tests are considerably more powerful that the specific size-corrected t-test when the disturbances are Cauchy (as does the sign-based procedure under lognormal disturbances) and are comparable to the model size-corrected t-test under $t(3)$ disturbances. Even when the disturbances are normal, the Wilcoxon-based bounds procedures performs respectably compared to the size-corrected t-test. As in the previous table, the median-estimate tests dominate the size-corrected parametric test no matter the type of disturbance with nominal size bounded by 5% in all the appropriate circumstances.

Four general types of heteroskedasticity are studied in Table 4. In the first, the variance of the underlying normal disturbances jumps from 1 to 16 halfway though the sample; in the two-break model the variance jumps first from 1 to 16 at $t = 75$ and then to 64 at $t = 150$. In the third variant, the variability of the disturbances grows linearly through the sample (i.e. e_t is a $N(0, 1)$ variable multiplied by t), while in the last variability grows exponentially (e_t is a $N(0, 1)$ variable multiplied by $\exp(t/2)$). Along with the testing procedures presented in Tables 2 and 3, we consider in this context an attempt taken from MacKinnon and White (1985), to correct in a general manner for heteroskedasticity through the preliminary estimation of a heteroskedastic-consistent covariance matrix which is then used in a GLS estimation of the model coefficients. A consistent quasi-t statistics (denoted by wm) can be computed and its performance is compared here with the other statistics. We consider three types of size-correction in investigating the power of the parametric tests in the cases of break heteroskedasticity. The first applies the empirical critical points associated with each specification studied; the second applies the largest critical points associated with specifications involving break or linear heteroskedasticity in either of the specifications of the Mankiw-Shapiro model considered in Tables 2 and 3; in the third, the empirical critical points determined in the case of exponential heteroskedasticity are applied, because they are the largest of all those considered. Of course, the largest critical value is by definition the one closest to the (unknown) critical value that would be appropriate in making the test truly robust to heteroskedasticity of unknown form.

The results of Table 4 repeat the previous themes. The asymptotic tests are unreliable. The power of the bounds procedure based on the Wilcoxon statistic is comparable to the size-corrected parametric statistics corrected according to the first and second procedures described above. Moreover, the bounds procedure based on the sign test is at least comparable in power to the wm-test corrected to account for all possibilities of heteroskedasticity, while the Wilcoxon-based bounds procedure is superior. It should be emphasized that if the t-test were to be corrected in a similar manner it would have zero power. The Wilcoxon version of the median-estimate test is superior in power to the wm-test corrected for the specific model considered, while the sign version is comparable in power. Finally, it should be noted that in all the experiments considered here the empirical level of the median-estimate tests does not exceed the nominal level.

In comparing the properties of the nonparametric procedures introduced in this paper with standard parametric tests based on the t-statistic, in contexts with different data generating processes such as those considered in the previous Tables, it is tempting to suggest that the parametric procedures should be modified to account for the specific data generating process and that a fairer comparison would match the nonparametric bounds procedure against the modified parametric test. The new parametric test could even have better power against such a parametric alternative in this context, but would of course provide no grounds for rejecting a wider nonparametric hypothesis. The approach here is to consider procedures that are provably valid in finite samples for a general nonparametric hypothesis. The power comparisons in the simulations of this section serve to illustrate that such general procedures do have reasonable power relative to tests that are valid only for narrower hypotheses.

We now turn to a simulation study of the random walk model given by (11). The parametric tests considered in what follows are based on $n\tilde{\beta}_1$, and the t-statistic, both defined using the OLS estimate of $\beta_t = \theta - 1$ in (11). Since these statistics are sensitive to the value of the intercept, it is usual practice to consider tests based on $n\hat{\beta}_1$, and the t-statistic now defined using $\hat{\beta}_1$, the OLS estimate of β_1 in the presence of a trend term. Critical points for the various parametric tests have been determined by simulation; see Fuller (1976, pp. 371, 373) for the relevant tables. As indicated in the previous section, the nonparametric bounds procedure are based on statistics given by (12) and (13). We will also consider median-based tests given by

$$S_g(\tilde{b}_0) = \sum_{t=1}^{n} u\left[(Y_t - Y_{t-1} - \tilde{b}_0)g_{t-1}\right], \tag{18}$$

$$SR_g(\tilde{b}_0) = \sum_{t=1}^{n} u\left[(Y_t - Y_{t-1} - \tilde{b}_0)g_{t-1}\right]R_t^+(\tilde{b}_0), \tag{19}$$

where $\tilde{b}_0$ is the sample median of $Y_t - Y_{t-1}$, $t = 1, \ldots, n$, and g_t is the centering function given by (14). Critical points associated with the $\mathrm{Bi}(n, 0.5)$ distribution are used for (18), while those for (19) come from the distribution of the Wilcoxon variate W_n.

To assess the relative merits of the six parametric and nonparametric tests of the random walk hypothesis, we follow the same pattern of Monte Carlo simulation used in the analysis of the Mankiw-Shapiro specification. The intercept is $\beta_0 = 2.0$ in all experiments with the point of departure $Y_0 = 0.0$ under the null and $Y_0 = \beta_0/(1-\theta)$ under the alternative. The results are presented in Tables 5, 6 and 7. With regard to the issue of the appropriate bounds strategy to pursue, the results of Table 5 confirm in this setting the wisdom that it is best to choose a wide confidence interval for β_0, so that the significance level for the second stage of the bounds procedure is not too small. Accordingly, the testing strategy represented by $\alpha_1 = 0.009$ is considered in the following Tables.

The results reported in Table 6 concerning the relative power of the different tests under various types of disturbances when $n = 100$ are noteworthy. First, the empirical size of the t-test based on the model without trend is considerably less than 0.05; it should be recalled that the Dickey-Fuller critical values are derived from simulations with $\beta_0 = 0$ and when circumstances indicate a non-zero drift, it is usual practice to estimate a model with a trend term. The nonparametric signed-rank bounds procedure is comparable in power to the t-test based on the trend model

TABLE 5
RANDOM WALK WITH DRIFT; NORMAL DISTURBANCES*:
COMPARISONS BETWEEN BOUNDS TESTS $N = 200$

Testing Strategy			Bounds Tests			
			SB		*SRB*	
α_1	α_2	θ	Reject	Accept	Reject	Accept
2.8	2.2					
		1.00	0.0	0.0	0.0	2.3
		0.96	0.2	45.1	0.6	36.0
		0.94	3.9	32.7	11.0	23.6
1.3	3.7					
		1.00	0.0	0.0	0.0	0.3
		0.96	2.4	49.2	8.4	38.0
		0.94	4.7	36.4	17.2	26.4
0.9	4.1					
		1.00	0.0	0.0	0.0	0.3
		0.96	3.3	47.1	8.7	38.9
		0.94	6.3	35.2	18.1	27.4
0.4	4.6					
		1.00	0.0	0.0	0.0	0.0
		0.96	2.4	44.7	9.3	38.7
		0.94	5.5	32.5	18.2	27.7

* Model (11) with $\beta_0 = 2.0$ and $\beta_1 = \theta - 1$. Probabilities are given in percentages. A nonparametric confidence interval $J(\alpha_1)$ with level $1 - \alpha_1$ is first constructed for β_0. The null ($\theta = 1.0$) is rejected if for all $b \in J(\alpha_1)$ the nonparametric test based on b is significant at level α_2; the null is accepted if no such test is significant at level $\alpha_3 = 0.05 + \alpha_1$; otherwise, the procedure is inconclusive. *SB* refers to the sign procedure, *SRB* to the Wilcoxon. The level of each strategy is bounded by 0.05. See text for further details.

TABLE 6
RANDOM WALK WITH DRIFT; VARIOUS TYPES OF DISTURBANCES*:
LEVEL AND POWER COMPARISONS. $N = 100$

	t-test		Median-Estimate Tests		Bounds Tests			
					SB		*SRB*	
θ	Without trend	With trend	$S(\tilde{b}_0)$	$SR(\tilde{b}_0)$	Reject	Accept	Reject	Accept
Cauchy Distribution								
1.00	3.3	6.5	0.5	1.4	0.0	21.6	0.2	26.6
0.95	7.3	7.9	72.9	78.2	32.5	5.6	50.4	6.6
0.90	20.6	13.5	88.2	91.3	51.2	0.8	72.4	2.1
t(3) Distribution								
1.00	0.1	5.2	0.0	0.0	0.0	0.0	0.0	0.1
0.95	12.1	9.2	17.1	25.8	1.1	45.3	6.9	42.1
0.90	33.8	20.6	38.4	52.7	5.2	23.0	20.4	22.0
Normal Distribution								
1.00	0.2	5.4	0.0	0.2	0.0	0.0	0.0	0.3
0.95	12.2	8.5	9.7	18.8	0.9	57.5	3.7	46.4
0.90	31.7	20.8	21.1	37.2	1.4	39.5	10.9	33.0

* Model (11) with $\beta_0 = 2.0$ and $\beta_1 = \theta - 1$. Entries for the *t*-test and median-estimate tests are percentage rejections; the latter statistics are given by (18) and (19).

TABLE 7
RANDOM WALK WITH DRIFT; VARIOUS TYPES OF DISTURBANCES*:
Level and Power Comparisons. $n = 200$

	t-test		Median-Estimate Tests		Bounds Tests			
					SB		*SRB*	
θ	Without trend	With trend	$S(\tilde{b}_0)$	$SR(\tilde{b}_0)$	Reject	Accept	Reject	Accept
Cauchy Distribution								
1.00	4.3	5.6	0.1	2.5	0.1	43.8	0.6	30.5
0.98	6.2	7.2	81.5	84.9	54.1	1.8	64.1	3.6
0.96	13.1	10.8	93.1	94.8	76.1	0.0	81.6	0.8
$t(3)$ Distribution								
1.00	0.1	5.0	0.0	0.1	0.0	0.0	0.0	0.7
0.96	20.0	13.4	32.0	45.3	6.8	24.6	18.5	24.8
0.94	42.2	25.2	49.1	66.4	17.7	14.4	35.5	13.5
Normal Distribution								
1.00	0.2	3.9	0.0	0.1	0.0	0.0	0.0	0.3
0.96	21.8	12.5	15.7	29.5	3.3	47.1	8.7	38.9
0.94	44.3	24.5	24.8	45.0	6.3	35.2	18.1	27.7

* Model (11) with $\beta_0 = 2.0$ and $\beta_1 = \theta - 1$. Entries for the t-test and median-estimate tests are percentage rejections; the latter statistics are given by (18) and (19).

and is strikingly superior when the disturbances are Cauchy, as is the sign-based bounds procedure in this case. The Wilcoxon bounds procedure performs respectably even when the disturbances are normal. Moreover, in all cases considered, the median-estimate test based on signed ranks outperforms the t-test based on a regression without drift with size bounded by the nominal level of the test. The strong power performance is repeated in Table 7 when $n = 200$. It should be noted that the median-estimate sign test has power comparable to the t-test based on a regression with trend even when the disturbances are normal.

5. AN APPLICATION

According to the strict theory of the term structure of interest rates, differences in yields at different maturities of, say, government bonds can be explained in a straightforward way by agents' expectations concerning future interest rates: long-term rates can be analyzed as the expected return from a series of shorter rates plus a constant risk or liquidity premium. Mankiw and Summers (1984) tested the expectations theory at the short end of the term structure with the additional assumption that expectations are formed rationally with strikingly negative results. Such findings, as more general analyses over the full term structure such as Shiller et al. (1983), are based on parametric statistical inference which may not be valid as the relevant regression disturbances are generally not normally distributed. Our goal in this section is to illustrate the nonparametric approach in this context. In striking contrast to the usual literature, we find for Canadian data that the expectations theory cannot be rejected when more correct nonparametric procedures are used.

The strict form of the expectations theory states that the relation between the return on three month and six month bonds is given by

(20) $$r_t^{(6)} = \theta + 0.5 r_t^{(3)} + 0.5 \hat{r}_{t+1}^{(3)},$$

where $r_t^{(3)}$ and $r_t^{(6)}$ are the yields of three- and six-month bonds sold at time t and $\hat{r}_{t+1}^{(3)}$ is the market forecast at time t of $r_{t+1}^{(3)}$. It follows according to the theory that the implied forecast error may be written as

(21) $$r_{t+1}^{(3)} - \hat{r}_{t+1}^{(3)} = \left(r_{t+1}^{(3)} - 2r_t^{(6)} + r_t^{(3)}\right) + 2\theta.$$

If we assume further that expectations are formed efficiently, then the implied forecast error must be independent of all information available to the market at time t, in particular the spread $r_t^{(6)} - r_t^{(3)}$. Further, this also implies that the forecast errors observed at the monthly frequencies should be serially independent at lags greater than two. These implications can be tested by considering the regression

(22) $$\left(r_{t+1}^{(3)} - 2r_t^{(6)} + r_t^{(3)}\right) = \delta_0 + \delta_1\left(r_t^{(6)} - r_t^{(3)}\right) + \epsilon_{t+1},$$

where $\delta_0 = -2\theta$ and the ϵ_t are serially independent at lags greater than two. We wish to test the null hypothesis that $\delta_1 = 0$.

The first section of Table 8 presents the results of OLS estimation of (22) based on monthly three- and six-month Canadian government bonds from 1969 to 1989.

TABLE 8
TERM STRUCTURE OF INTEREST RATES:
PARAMETRIC AND NONPARAMETRIC EFFICIENCY RESULTS*

	Data Set		
	S_1	S_2	S_3
OLS Estimates			
δ_0 (t-test)	−0.22 (−1.66)	−0.09 (−0.75)	−0.01 (−0.04)
δ_1	−0.30 (−0.66)	−1.01 (−2.46)	−1.52 (−3.56)
Residual Normality (χ_2^2)	46.41	113.51	250.24
Nonparametric Analysis Confidence Interval (99%)	(−0.38, −0.03)	(−0.46, −0.05)	(−0.38, −0.05)
Sign Tests			
Median-Estimate	0.36	0.58	0.20
$Q_L(S)$	0.71	0.85	0.99
$Q_U(S)$	0.14	0.46	0.20
Wilcoxon Tests			
Median-Estimate	0.57	0.53	0.87
$Q_L(SR)$	0.70	0.58	0.98
$Q_U(SR)$	0.44	0.51	0.72

* Monthly data taken from Statistics Canada on three- and six-month Canadian Government bonds from 1960 to 1989 is divided into three sub-samples, S_1, S_2 and S_3. The regression equation is given by (22); the Jarque-Bera (1987) normality test is applied to the residuals. The Median-Estimate tests are given by (16) and (17). According to the nonparametric bounds procedure, the null is rejected if $Q_L \leq 0.04$; the null is accepted if $Q_U \geq 0.06$.

Since the model applied to interest rates of these maturities describes the relationship at three-month intervals, the monthly data are treated as three sub-samples of observations, S_1, S_2 and S_3, taken at three-month intervals. The regression results for this particular data set confirm the general findings of Mankiw and Summers (1984). The null hypothesis is rejected for two of the three samples with δ_1 in all cases less than 0. It should be noticed, however, that a standard test of disturbance normality rejects the normality assumption in all three subsamples.

The second section of Table 8 presents the results of the nonparametric tests. Following the approach described in the previous sections, we first construct a confidence interval for the intercept -2θ, where θ is the constant liquidity premium. The confidence intervals are roughly identical for the three samples. Aligned sign and signed rank statistics are constructed based on different point estimates of the intercept taken throughout the confidence interval. The test based on the mid-point corresponds to the median-estimate test. For each sample, the maximum and minimum values of these statistics are found. The associated p-values are given in Table 8. In all samples, not only is the null not rejected, since $Q_L(S)$ and $Q_L(SR)$ are greater than 0.04; but the null is accepted as well in all samples, since $Q_U(S)$ and $Q_U(SR)$ are not less than 0.09 (for illustrative purposes we are taking $\alpha = 0.05$). The median-estimate tests are not significant as well.

The contrast between the parametric and nonparametric results is striking. Where the parametric results pointed to a rejection of the expectations theory of the term structure, the nonparametric analysis confirms the theory. It should be emphasized that there is accompanying evidence (normality tests) that the parametric inference is not appropriate here, while the nonparametric procedures are valid for such small samples under the framework of the model.

6. CONCLUSION

The testing procedures presented in this paper have been developed in response to a specific challenge. In many situations which arise naturally in testing fundamental implications of the rational expectations hypothesis, standard regression-based testing procedures reject much too often, even when the sample size is as large as 200. This paper, along with earlier work in Campbell and Dufour (1995), offers an alternative nonparametric approach which does not suffer from this parametric test defect. Nonparametric tests based on signs and signed-ranks are valid for a wide class of models involving feedback; these include as specific cases the model studied by Mankiw and Shapiro (1986), and the random walk model. Our previous results were suited to models involving no intercept or drift term. In this paper we have extended these results to cover these important cases as well. To complement the results establishing the validity of our nonparametric procedures, the results of simulation studies presented here show that the tests have good power relative to the parametric alternatives even in circumstances favorable to the usual regression tests. In cases involving outliers or heteroskedastic disturbances, nonparametric tests remain valid and can exhibit considerably greater power.

The procedures presented in this paper can be generalized in different directions and is the subject of ongoing research. For example, it is possible to relax somewhat the assumption of the independence of Y_t from past information in a manner that permits the possibility of stochastic volatility of a recursive type such as in ARCH models. Our procedures would then apply to a variety of examples of interest to researchers in empirical finance. These results will be presented in a future paper.

We do not want to over-emphasize the usefulness of the nonparametric procedures presented in this paper. The tests are best applied in situations where the null hypothesis simplifies the model, as in the null of efficiency in rational expectations models. The procedures cannot be readily applied to a more complicated testing environment. Such extensions are pursued in current research. But there is no good reason to continue to use flawed parametric regression-based tests in situations where there are valid nonparametric testing alternatives that have good power.

APPENDIX

PROOF OF PROPOSITION 1. Follows directly from Propositions 1, 2 and 3 of Campbell and Dufour (1995).

PROOF OF PROPOSITION 2. To simplify the notation, write $S = S_g$, $\bar{S} = \bar{S}_g$, $\tilde{S} = \tilde{S}_g$, $SR = SR_g$, $\overline{SR} = \overline{SR}_g$ and $\widetilde{SR} = \widetilde{SR}_g$.

(a) Let A be the event that $S(b) > \bar{S}(\alpha_2)$ for all $b \in J(\alpha_1)$. We wish to show that $P[A] \leq \alpha_1 + \alpha_2$. First, define $I = \{b\colon\ b \in J(\alpha_1) \text{ and } S(b) \leq \bar{S}(\alpha_2)\}$. Then, by standard rules of the probability calculus, it follows that

$$P[b_0 \in I] = 1 - P\left[b_0 \notin J(\alpha_1) \text{ or } S(b_0) > \bar{S}(\alpha_2)\right]$$

$$\geq 1 - P\left[b_0 \notin J(\alpha_1)\right] - P\left[S(b_0) > \bar{S}(\alpha_2)\right] \geq 1 - \alpha_1 - \alpha_2,$$

since by definition $P[b_0 \in J(\alpha_1)] \geq 1 - \alpha_1$ and $P[S(b_0) > \bar{S}(\alpha_2)] \leq \alpha_2$. Observe that $P[A] = P[B^c]$, where B is the event that $S(b) \leq \bar{S}(\alpha_2)$ for some $b \in J(\alpha_1)$. Since $B \supseteq \{S(b_0) \leq \bar{S}(\alpha_2) \text{ and } b_0 \in J(\alpha_1)\}$, we have

$$P[B] \geq P[b_0 \in I] \geq 1 - \alpha_1 - \alpha_2 \geq 1 - \alpha,$$

with the immediate consequence that $P[A] \leq \alpha_1 + \alpha_2 \leq \alpha$, so that (7a) is established.

The two inequalities (7b) and (7c) follow by using Proposition 1, which implies that $S(b_0) \sim \text{Bi}(n, 0.5)$, a symmetric distribution on the integers $\{0, 1, \ldots, n\}$, so that $M - S(b_0) = n - S(b_0) \sim \text{Bi}(n, 0.5)$. The proof of (7b) is then similar to the one of (7a) with $S(b)$ replaced by $M - S(b)$, while the proof of (7c) is obtained on replacing $S(b)$ by $\text{Max}\{S(b), M - S(b)\}$ and $\bar{S}(\alpha_2)$ by $\bar{S}(\alpha_2/2)$ in the same proof.

We now turn to (7d). Let C represent the event that $S(b) < \tilde{S}(\alpha_3)$ for all $b \in J(\alpha_1)$. We have to show that $P[C] \leq 1 - (\alpha_3 - \alpha_1)$. By the definition of $\tilde{S}(\alpha_3)$, we have

$$P\left[S(b_0) < \tilde{S}(\alpha_3)\right] = P\left[S(b_0) < n - \bar{S}(1 - \alpha_3)\right] = P\left[n - S(b_0) > \bar{S}(1 - \alpha_3)\right] \leq 1 - \alpha_3.$$

Now, as in the proof of (7a), we consider the complement of C, i.e. the event that $S(b) \geq \tilde{S}(\alpha_3)$ for some $b \in J(\alpha_1)$, and let $\tilde{I} = \{b\colon b \in J(\alpha_1) \text{ and } S(b) \geq \tilde{S}(\alpha_3)\}$. Then we have

$$P\left[b_0 \in \tilde{I}\right] \geq 1 - P\left[S(b_0) < \tilde{S}(\alpha_3)\right] - P\left[b_0 \notin J(\alpha_1)\right]$$

$$\geq 1 - (1 - \alpha_3) - \alpha_1 = \alpha_3 - \alpha_1 \geq \alpha$$

and, since the event C implies $b_0 \notin \tilde{I}$,

$$P[C] \leq P\left[b_0 \notin \tilde{I}\right] \leq 1 - (\alpha_3 - \alpha_1) \leq 1 - \alpha,$$

and (7d) is established. The inequalities (7e) and (7f) follow on observing that $n - S(b_0) \sim \text{Bi}(n, 0.5)$: the proofs of (7e) and (7f) are similar to that of (7d) with $S(b)$ replaced by $n - S(b)$ and $\text{Max}\{S(b), n - S(b)\}$ respectively.

(b) The same argument as in (a) with $S(b)$ replaced with $SR(b)$ and $\bar{S}$ with $\overline{SR}$ establishes the result.

REFERENCES

BANERJEE, A. AND J. DOLADO, "Do We Reject Rational Expectations Models Too Often? Interpreting Evidence Using Nagar Expansions," *Economics Letters* 24 (1987), 27–32.

——— AND ———, "Tests of the Life-cycle Permanent Income Hypothesis in the Presence of Random Walks: Asymptotic Theory and Small Sample Interpretation," *Oxford Economic Papers* 40 (1988), 610–633.

———, ——— AND JOHN W. GALBRAITH, "Orthogonality Tests with De-trended Data: Interpreting Monte Carlo Results Using Nagar Expansions," *Economics Letters* 32 (1990), 19–24.

BERA, A. AND C. JARQUE, "A Test for Normality of Observations and Regression Residuals," *International Statistical Review* 55 (1987), 163–172.

BURRIDGE, P. AND E. GUERRE, "The Limit Distribution of Level Crossing of a Random Walk, and a Simple Unit Root Test," University of Birmingham, mimeo, 1995.

CAMPBELL, B. AND J. M. DUFOUR, "Over-rejections in Rational Expectations Models: A Nonparametric Approach to the Mankiw-Shapiro Problem," *Economics Letters* 35 (1991), 285–290.

——— AND ———, "Exact Nonparametric Orthogonality and Random Walk Tests," *The Review of Economics and Statistics* 77 (1995), 1–16.

——— AND E. GHYSELS, "Federal Budget Projections: A Nonparametric Assessment of Bias and Efficiency," *The Review of Economics and Statistics* 77 (1995), 17–32.

DUFOUR, J. M., "Rank Tests for Serial Dependence," *Journal of Time Series Analysis* 2 (1981), 117–128.

———, "Exact Tests and Confidence Sets in Linear Regressions with Autocorrelated Errors," *Econometrica* 58 (1990), 475–495.

——— AND M. HALLIN, "Nonuniform Bounds for Nonparametric *t*-tests," *Econometric Theory* 7 (1991), 253–263.

———, Y. LEPAGE AND H. ZEIDAN, "Nonparametric Testing for Time Series: A Bibliography," *Canadian Journal of Statistics* 2 (1982), 1–38.

ENGSTED, T., "The Term Structure of Interest Rates in Denmark 1982–89: Testing the Rational Expectations/Constant Liquidity Premium Theory," *Bulletin of Economic Research* 45 (1993), 19–37.

FAMA, E. F., "The Information in the Term Structure," *Journal of Financial Economics* 13 (1984), 509–528.

FULLER, W. A., "Introduction to Statistical Time Series" (New York: Wiley and Sons, 1976).

GALBRAITH, J. W., J. DOLADO AND A. BANERJEE, "Rejections of Orthogonality in Rational Expectations Models: Further Monte Carlo Evidence for an Extended Set of Regressors, *Economics Letters* 25 (1987), 243–247.

GRANGER, C. W. J. AND J. HALLMAN, "Nonlinear Transformations of Integrated Time Series," *Journal of Time Series Analysis* 12 (1991), 207–224.

HALLIN, M. AND M. L. PURI, "Rank Tests for Time Series Analysis: A Survey," In David Brillinger et al., eds., *New Directions in Time Series Analysis*: Part I (New York: Springer Verlag, 1991, pp. 111–153).

HETTMANSPERGER, T. P., "Statistical Inference Based on Ranks" (New York: Wiley and Sons, 1984).

KUGLER, P., "The Term Structure of Euro Interest Rates and Rational Expectations," *Journal of International Money and Finance* 9 (1990), 234–244.

LEHMANN, E. L., "Nonparametrics: Statistical Methods Based on Ranks" (San Francisco: Holden Day, 1975).

——— AND C. STEIN, "On the Theory of Some Nonparametric Hypotheses," *Annals of Mathematical Statistics* 20 (1949), 28–45.

MACKINNON, J. G. AND H. WHITE, "Some Heteroskedastic-consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics* 29 (1985), 305–325.

MANKIW, N. G., The Term Structure of Interest Rates Revisited, *Brookings Papers on Economic Activity* (1986), 61–110.

——— AND J. A. MIRON, "The Changing Behavior of the Term Structure of Interest Rates," *Quarterly Journal of Economics* 101 (1986), 211–228.

——— AND M. D. SHAPIRO, "Do We Reject Too Often? Small Sample Properties of Rational Expectation Models," *Economics Letters* 20 (1986), 243–247.

——— AND L. SUMMERS, "Do Long-term Interest Rates Overreact to Short-term Interest Rates?," *Brookings Papers on Economic Activity* (1984), 223–242.

MELINO, A., "The Term Structure of Interest Rates: Evidence and Theory," *Journal of Economic Surveys* 2 (1988), 335–366.

PRATT, J. W. AND J. D. GIBBONS, Concepts of Nonparametric Theory (New York: Springer-Verlag, 1981).

SHILLER, R. J., "The Term Structure of Interest Rates," In B. Friedman and F. Hahn, eds., *Handbook of Monetary Economics*, Vol. 1 (North Holland: Elsevier Science Publishers, 1990).

———, J. Y. CAMPBELL AND K. L. SHOENHOLTZ, "Forward-rates and Future Policy: Interpreting the Term Structure of Interest Rates," *Brookings Papers on Economic Activity* (1983), 173–217.

TAYLOR, M. P., "Modeling the Yield Curve," *Economic Journal* 102 (1992), 524–537.

WILCOXON, F. S., S. K. KATTI AND R. WILCOX, "Critical Values and Probability Levels for the Wilcoxon Rank Sum Test and the Wilcoxon Signed Rank Test", In H. L. Harter and D. B. Owen, eds., *Selected Tables in Mathematical Statistics* (Providence: Institute of Mathematical Statistics, 1970).