ROBUST POWER CALCULATIONS WITH TESTS FOR SERIAL CORRELATION IN STOCK RETURNS

by

Matthew Richardson
Tom Smith

12-91

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

The contents of this paper are the sole responsibility of the author(s).

Robust Power Calculations with Tests for Serial Correlation in Stock Returns

Matthew Richardson*
Wharton School, University of Pennsylvania

Tom Smith
Fuqua School of Business, Duke University

May 1989
Latest Revision May 1991

Abstract

This paper provides an asymptotically most powerful test for a particular class of statistics which test the hypothesis of no serial correlation. This class includes many of the statistics employed in the recent finance and macroeconomics literature. Furthermore, with respect to a popular mean reversion alternative model, we show that the asymptotically most powerful test is quite robust to distributional specifications.

*We would like to thank seminar participants at Duke University and Princeton University, and especially Rob Stambaugh for many helpful comments and suggestions. Research support received from the Geewax-Terker Research Fund and the University of Pennsylvania Research Foundation is gratefully acknowledged

1 Introduction

Researchers in economics and finance have become very proficient at calculating the asymptotic distributions of statistics under fairly weak assumptions. Nowhere is this more apparent than in the recent literature on tests for serial correlation. As a case in point, the asymptotic distribution of multiperiod autocorrelations, variance ratios, regressions of one period on multiperiod past values, and other statistics, have been derived. These distributions are easy to interpret since under the random walk null hypothesis they depend only on the multiperiod length. Moreover, these distributions tend to have asymptotic normal distributions, making formal inference straightforward to calculate. Small sample considerations aside, we, therefore, have a very good idea of the size of these tests.

Under these same fairly weak assumptions, what has been said about the other side of the coin — the power properties of these statistics? The common practice is to perform a Monte Carlo simulation under a specific distributional assumption and evaluate power that way. For example, Lo and MacKinlay (1989) and Poterba and Summers (1988) investigate the power of the variance-ratio statistic via Monte Carlo simulation. There are a number of disadvantages to analyzing the serial dependence tests within a Monte Carlo setting.

The first problem with the Monte Carlo procedure is that, without some theoretical justification, we have no way of knowing whether it is robust to alternative parametric assumptions. Alternative theories to the random walk model have usually been discussed in general terms; for example, the mixture of a permanent and transitory component of stock prices in Poterba and Summers (1988) and Fama and French (1988). These alternatives, however, place very little structure on the innovations in the stock price components other than mean zero and no serial correlation. Monte Carlo simulations, however, impose considerable structure by completely specifying and parameterizing the distribution of these innovations.

The second disadvantage is that the Monte Carlo results are difficult to interpret. The Monte Carlo results presented in the serial correlation literature clearly have practical value. For example, the small sample power of commonly used statistics under various distributional assumptions, sample sizes and alternatives have been documented. If these criteria are met for our particular problem, we have some idea of the usefulness of the hypothesis testing. These analyses, however, have trouble explaining *why* different decision rules are reached by different tests. (See, for example, the test statistic discussions and resulting conclusions given in Fama and French (1988), Lo and MacKinlay (1989) and Jegadeesh (1989)). The question, therefore, of why particular serial correlation tests get different power under various scenarios remains unanswered.

A notable exception to the Monte Carlo approach is given by Faust (1989). Under the assumption that returns are normally distributed, Faust (1989) derives the finite sample power properties of variance-ratio statistics against mean-reverting alternatives. In particular, for a given variance-ratio statistic, he finds the serial dependence alternative it is optimal against. Of special interest, Faust (1989) provides an intuitive feel for how close the variance ratio test is to the most powerful test against economically meaningful alternatives. There are some drawbacks to this approach. The first is that considerable structure (i.e. normally distributed stock returns) is placed on the null and alternative models. The second is that a number of the statistics currently used (e.g. the multiperiod autocorrelations, among others) do not fall into the variance ratio class. Comparisons, across these statistics are, therefore, more problematic.

The goal of this paper is to provide a new approach to analyzing the properties of the serial correlation statistics under interesting alternatives. In particular, using a procedure developed by Bahadur (1960) and Geweke (1981), we make asymptotic power comparisons and quantify the relative benefit of using a particular test for serial correlation.[1] This approach has a number of advantages:

- the optimal test is derived within a general class of statistics which captures recent test statistics for serial correlation in stock returns.
- the asymptotic power results are robust to parameterized distributions — in particular, the results depend only on the length of the multiperiod and the autocorrelation structure governing the alterantive.
- the results are intuitive, pointing to reasons why particular tests have power and quantifying how much power relative to other tests.

Whereas our approach provides more intuition and relies on weaker assumptions than the standard Monte Carlo approach, its major disadvantage is that the results are large sample

[1] Richardson and Smith (1989) introduce Gewke's (1981) approximate slope procedure to some recent tests for serial correlation. They explicitly quantify the benefits to overlapping data and compare the relative asymptotic power of j-period autocorrelations versus j-period variance ratios under various alternatives. Jegadeesh (1989) extends their discussion to regressions of j-periods on k-period returns.

based. This problem, however, is true of all asymptotic theory. We provide simulations and show that the intuition provided by the large sample analysis carries through to small samples.

The paper is organized as follows. Section 2 reviews the approximate slope procedure and discusses its application to tests for serial dependence. Section 3 derives the asymptoic distribution of a broad class of test statistics and shows how to apply the approximate slope procedure to this class. Of special interest, the asymptotically most powerful test is derived. Section 4 applies these results to an example of recent interest to financial economists, namely the mixture of a permanent and transitory component model of stock prices. Section 5 discusses some extensions. Section 6 concludes the paper.

2 Review of the Approximate Slope

Bahadur (1960) defines the approximate slope of the test to be the rate at which the logarithm of the asymptotic marginal significance level of the test decreases as sample size increases, under a given alternative. Consider some test i. If power against a given alternative is held constant,

$$\frac{-2ln[\alpha^i(\theta)]}{T} \xrightarrow{a.s.} c^i(\theta) \tag{1}$$

where

θ = vector of parameters

α^i = marginal significance level of ith test statistic

$c^i(\theta)$ = approximate slope of ith test statistic .

Geweke (1981) shows that if the test statistic's limiting distribution under the null hypothesis is χ^2, then $c^i(\theta)$ equals the probability limit (i.e. *plim*) of the statistic divided by T.

For example, consider the sample autocorrelation of a continuously compounded j holding period return for a sample size T. Let P_t be the logarithm of price at time T. Then the sample autocorrelation can be written,

$$\hat{\beta}(j) = \frac{\sum_{t=j}^{T-j}(P_t - P_{t-j} - j\hat{\mu})(P_{t-j} - P_{t-2j} - j\hat{\mu})}{\sum_{t=j}^{T-j}(P_t - P_{t-j} - j\hat{\mu})^2},$$

where $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T}(P_t - P_{t-1})$.

Suppose we wish to test whether continuously compounded returns are uncorrelated, i.e. $\beta_j = 0$. Under the assumption that $P_t - P_{t-1}$ is stationary and ergodic, and under assumptions restricting conditional heteroskedasticity, Richardson and Smith (1989) show that the asymptotic variance of the autocorrelation is

$$Var[\hat{\beta}(j)] = \frac{2j^2+1}{3j}.$$

Using this result and Geweke (1981), the approximate slope of this statistic has a simple form:

$$c(\theta) = plim\left((\frac{3j}{2j^2+1})\hat{\beta}_j^2 \right).$$

Given a level of power and number of observations T, the interpretation of $c(\theta)$ is that the higher $plim(\hat{\beta}_j)$ is, the lower is α, the marginal significance level of the test. Intuitively, the greater the approximate slope the more incredible the null hypothesis becomes.

Of special interest, Geweke (1981) also provides formal justification for comparing the approximate slopes of different tests, subscripted by i and j. Specifically, let Λ_* equal the largest nonrejection region possible in a sample size of T given that the probability of not rejecting the null hypothesis is not to exceed a value $\gamma \in [0,1]$. Similarly, define T_* equal to the minimum number of observations required to insure that the probability that the statistic exceeds the nonrejection region Λ is $1-\gamma$. Then for any ith and jth test statistic with the same limiting χ^2 distribution,

Result 1. $$\lim_{\Lambda\to\infty} \frac{T_*^i}{T_*^j} = \frac{c^j}{c^i}$$

and

Result 2. $$\lim_{T\to\infty} \frac{\Lambda_*^i}{\Lambda_*^j} = \frac{c^i}{c^j}.$$

Either of the above results may be intrepreted in the following way. A test with greater approximate slope may be expected to reject the null hypothesis more frequently under that alternative than one with smaller approximate slope. That is, for given power fewer observations are needed to reject (see result 1) or a larger nonrejection region is allowable (see result 2).

For example, consider two particular autocorrelation statistics, $\hat{\beta}_j$ and $\hat{\beta}_{j'}$ where $j \neq j'$. Result 1 states that the ratio of their approximate slopes,

$$\frac{plim\left((\frac{3j}{2j^2+1})\hat{\beta}_j^2\right)}{plim\left((\frac{3j'}{2j'^2+1})\hat{\beta}_{j'}^2\right)} \approx \frac{j'\,plim[\hat{\beta}_j]^2}{j\;plim[\hat{\beta}_{j'}]^2},$$

has a special interpretation. If $j'plim(\hat{\beta}_j)^2 = 2 \times j\; plim(\hat{\beta}_{j'})^2$, then, if we employ the estimator $\hat{\beta}_{j'}$ instead of $\hat{\beta}_j$, we will need twice as many observations to reject the null hypothesis of no serial dependence. In practice, we might expect for small asymptotic marginal levels of significance that employing $\hat{\beta}_j$ with say 720 observations will achieve roughly the same power as $\hat{\beta}_{j'}$ with 1440 observations. Alternatively, for a large given number of observations, the nonrejection region would be about twice as large with $\hat{\beta}_j$. This type of comparison provides the researcher with a simple way to evaluate the relative power of different statistics.

Geweke (1981), however, points out two caveats with respect to this procedure. First, the approximate slope comparisons are only strictly valid asymptotically. It may, therefore, be inappropriate in small samples — this is true, however, of all asymptotic theory. Second, comparisons are made using the same critical points for both tests, which are known only asymptotically. For example, in this paper, all the test statistics have asymptotic χ^2 representations. In small samples, however, their distributions may look quite different.

3 The Optimal Test Statistic: Theory

More generally, if changes in a series P_t are uncorrelated, then the following restrictions hold:

$$\begin{aligned} & var(\sum_{i=1}^{j} R_{t+i}) = j \times var(R_t) \\ & cov(R_t, R_{t-k}) = 0 \quad \forall k \neq 0 \\ \text{where} \quad & R_t = P_t - P_{t-1}. \end{aligned}$$

These restrictions in turn imply a corresponding set of sample moment conditions:

$$g_T(\mu, \sigma_j^2, \rho_k(j), \rho_l(j)) \equiv \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} R_t - \mu \\ (\sum_{i=1}^{j} R_{t+i} - j\mu)^2 - j\sigma^2 \\ (R_t - \mu)(R_{t-k} - \mu) - \sigma^2 \rho_k(j) \\ (R_t - \mu)(R_{t-l} - \mu) - \sigma^2 \rho_l(j) \\ \vdots \end{pmatrix} = 0 \quad \forall j, k, l.$$

Let $\bar{R}_t = R_t - \mu$. Then under the assumptions that the R_t are stationary and ergodic and satisfy

$$\begin{aligned} 1. \quad & E[\bar{R}_t \bar{R}_{t-j}] = && 0 \\ 2. \quad & E[\bar{R}_t^2 \bar{R}_{t-j} \bar{R}_{t-k}] = && 0 \\ 3. \quad & E[\bar{R}_t^2 \bar{R}_{t-j}^2] = && \sigma^4 \quad \forall t, \forall j \neq k, \end{aligned}$$

and given results in Hansen (1982), it is possible to show that the asymptotic distribution of $\hat{\rho}(j) \equiv (\,\hat{\rho}_k(j) \quad \ldots \quad \hat{\rho}_l(j)\,)'$ is given by

$$\sqrt{T}\hat{\rho}(j) \overset{a}{\sim} N(\underline{0}, I)$$

where $\hat{\rho}_i(j) = \frac{\sum_{t=1}^{T}(P_t - P_{t-1} - \hat{\mu})(P_{t-i} - P_{t-i-1} - \hat{\mu})}{\sum_{t=1}^{T}(P_t - P_{t-j} - j\hat{\mu})^2/j}$ and $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T}(P_t - P_{t-1})$.

Note that this derivation required some restrictions on conditional heteroskedasticity of the series R_t. While this may be a reasonable assumption in many applications, recent evidence with respect to stock returns suggests violation of assumption 3 (see, for example, Schwert (1989)). It is possible, however, to derive the asymptotic distribution of $\hat{\rho}(j)$ with this assumption relaxed. Under this scenario, the asymptotic distribution for typical elements $\hat{\rho}_k(j)$ and $\hat{\rho}_l(j)$ is

$$\sqrt{T}\begin{pmatrix} \hat{\rho}_k(j) \\ \hat{\rho}_k(j) \end{pmatrix} \overset{a}{\sim} N(\underline{0}, I + C_{kl})$$

$$\text{where} \qquad C_{kl} = \begin{pmatrix} \frac{cov(\bar{\Delta}_t^2, \bar{\Delta}_{t-k}^2)}{\sigma^4} & 0 \\ 0 & \frac{cov(\bar{\Delta}_t^2, \bar{\Delta}_{t-l}^2)}{\sigma^4} \end{pmatrix}.$$

The variance of the autocorrelation (under heteroskedasticity) is equal to the variance under no heteroskedasticity plus an adjustment factor which represents the persistence in the conditional variance. The prevailing evidence suggests that for stock returns $cov(\bar{R}_t^2, \bar{R}_{t-k}^2) > 0$, which means that in practice the autocorrelation estimators will be less precise.

3.1 Class of Test Statistics

In order to avoid specifying this autocorrelation structure of the variance, however, we maintain the conditional heteroskedasticity restriction of assumption 3 throughout the rest of the paper. Note that since any linear combination of normals is normal, a more general result for $\hat{\rho}(j)$ is available:

$$
\begin{aligned}
& \sqrt{T}\underbrace{\hat{\rho}(j)}_{N\times 1} \sim N(\underbrace{0}_{N\times 1}, \underbrace{I}_{N\times N}) \\
\Longrightarrow \quad & \sqrt{T}\underbrace{D}_{M\times N}\underbrace{\hat{\rho}(j)}_{N\times 1} \sim N(\underbrace{0}_{M\times 1}, \underbrace{DD'}_{M\times M}) \\
\Longrightarrow \quad & J_T \equiv T\underbrace{(D\hat{\rho}(j))'}_{1\times M}\underbrace{(DD')^{-1}}_{M\times M}\underbrace{(D\hat{\rho}(j))}_{M\times 1} \sim \chi^2_M.
\end{aligned} \tag{2}
$$

Equation (2) represents a general class of statistics for testing the null hypothesis of no serial correlation — namely, particular linear combinations of sample autocovariances, $cov(\bar{R}_t, \bar{R}_{t-k})$, weighted by some measure of the variance of R_t. Under the null hypothesis, this class contains all linear combinations of consistent estimators of autocorrelations of the series R_t. As such, equation (2) contains either exact or approximate representations of many of the recent statistics used in the literature to test for serial correlation. For example, the multiperiod sample autocorrelation, the commonly used variance ratio statistic, and the Box-Pierce statistic, among others, fall into this class. Table 1 provides some examples of the weights D and the corresponding statistics.[2]

3.2 An Asymptotically Most Powerful Test

Equation (2) is a formula for a general class of test statistics which place different weights on consistent autocorrelation estimators over various multiperiod sums. As a special case, it captures many of the recent statistics employed in the finance and macroeconomics literature. Which statistic — that is, which weights D — should the econometrician choose?

Given that they all admit asymptotic χ^2 distributions, one criteria might be to choose the statistic based on power against a particular alternative. In the context of this discussion, choose D and j in order to maximize the approximate slope of the statistic. Under

[2]See also Cochrane (1988), Lo and MacKinlay (1989) and Poterba and Summers (1988) for representations of variance ratios in terms of autocorrelation estimators.

a given alternative and for a fixed number of restrictions M in equation (1), it is straightforward to apply Geweke's (1981) approximate slope procedure. In this case, choose D and j to maximize

$$c^j(D) \equiv \{D[plim(\hat{\rho}(j))]\}'\{DD'\}^{-1}\{D[plim(\hat{\rho}(j))]\}. \tag{3}$$

In order to solve the maximization of (3), all we need are the probability limits of the autocorrelation estimators $\hat{\rho}(j)$. Given a particular alternative, these can be readily calculated. One particular model which has received considerable attention in the recent literature is the mixture of permanent and stationary components of stock prices proposed by Fama and French (1988) and Poterba and Summers (1988). Below, we apply the approximate slope analysis to this alternative.

4 Mean-reversion in Stock Prices

To see the approximate slope procedure in practice, consider fixing the number of restrictions, M, equal to one and consider the alternative hypothesis proposed by Poterba-Summers (1988) and Fama-French (1988):

$$\begin{aligned} P_t &= \mu + q_t + z_t \\ \text{where } q_t &= q_{t-1} + \epsilon_t, \quad E[\epsilon_t] = E[\epsilon_t \epsilon_{t-j}] = 0 \\ z_t &= \lambda z_{t-1} + \eta_t, \quad E[\eta_t] = E[\eta_t \eta_{t-j}] = 0, \ |\lambda| < 1. \end{aligned}$$

This model contains a stationary component, z_t, and a permanent component, q_t. In this system, prices mean-revert and the speed of this reversion depends on the size of the autocorrelation parameter, λ, and the share of the variance captured by the stationary component (denote σ_η^2).

Under the assumptions of the model, the probability limit of $\hat{\rho}_k(j)$ can be written as

$$\begin{aligned} plim[\hat{\rho}_k(j)] &= \frac{-\gamma\lambda^{k-1}(1-\lambda)^2}{2(1-\gamma)(1-\lambda) + 2\gamma\frac{(1-\lambda^j)}{j}} \\ \text{where } \gamma &= \text{share of variance captured by mean-reverting component } z_t \\ &= \frac{2\sigma_\eta^2}{(1+\lambda)\sigma_\epsilon^2 + 2\sigma_\eta^2}. \end{aligned}$$

Substituting this formula for $plim[\hat{\rho}_k(j)]$ into equation (3), we can then maximize over D and j, given values for λ and γ. Specifically, for the single restriction case (i.e. $M = 1$), we want to

$$\max_{(D_k \ldots D_l), j} \underbrace{\left[\frac{\gamma(1-\lambda)^2}{2(1-\gamma)(1-\lambda) + 2\gamma\frac{(1-\lambda^j)}{j}} \right]^2}_{A} \underbrace{\frac{[\sum_{i \in (D_k \ldots D_l)} D_i \lambda^{i-1}]^2}{\sum_{i \in (D_k \ldots D_l)} D_i^2}}_{B}. \tag{4}$$

This maximization problem involves two separate parts. First, which value of j maximizes A? It is clear from the above equation that A reaches its maximum as j gets large. The intuition for this result is straightforward. Since $\hat{\rho}_k(j)$ converges in probability to $\frac{cov(R_t, R_{t-k})}{var(\sum_{i=1}^{j} R_{t+i})/j}$, and all the covariances are negative under this alternative, it is apparent that as $j \longrightarrow \infty$ the variance measure $\frac{var(\sum_{i=1}^{j} R_{t+i})}{j}$ approaches σ_ϵ^2 from above. The null theory above holds for j fixed and letting $T \to \infty$; hence, the choice of j cannot approach infinity.

However, small sample considerations aside, the fixed value of j should be "large". The question of how large is "large" presents a potential problem. The difficulty arises because once we have fixed a j, say j^*, we could always achieve greater asymptotic power by choosing $j^* + 1$. With respect to this particular mean-reversion example, it turns out that, although we increase the approximate slope as we increase j, the marginal gain to increasing j declines. Therefore, at some value of j, the efficiency gain is essentially zero. Moreover, how quickly this occurs depends only on two parameters governing the stationary component, that is, γ and λ.

With respect to part B, the maximization problem can be reduced to choosing the elements of D which maximize the sum of the covariances or in this particular case:

$$\max_{(D_k \ldots D_l)} \frac{\sum_{i \in (D_k \ldots D_l)} D_i \lambda^{i-1}}{\sqrt{\sum_{i \in (D_k \ldots D_l)} D_i^2}}. \tag{5}$$

Our choice of D depends only on λ, the mean-reversion parameter. It does not depend on any parameters governing the distribution of ϵ_t or η_t (other than recognizing that the null hypothesis includes the assumptions given in Section 3).

The effect on the approximate slope as we add additional weights $D_i \in D$ is ambiguous. The reason for this is straightforward: as we add elements to D, we apparently pick-up more mean-reversion in the series (i.e. the numerator in (5)) but also more variation due to estimation of additional parameters (i.e. the denominator in (5)). The choice of D, therefore, is a tradeoff between gaining more information at the cost of that information being noisy. The exact choice of D will depend on the magnitude of λ — that is, on how quickly prices mean revert. In general, this method can be used to compare the relative asymptotic power of any linear combination of $\hat{\rho}_i$s. This procedure, therefore, has widespread applicability.

For a given value of λ, it is possible to solve the maximization problem (5) to find the optimal number and values of the elements D_i in D. After working through the maximization problem, the solution is

$$D_i^* = \lambda^{i-1} \quad \forall i.$$

The intuition behind this result is that although there is useful information in each autocorrelation the econometrician needs to put less weight on the longer autocorrelations as they contain less information. The values of these weights are determined by how fast the autocorrelations approach zero, that is, by λ^{i-1}.

4.1 Example

In this subsection, we find the asymptotically most powerful test against the alternative given in Section 4. Specifically, to coincide with the finance literature, we choose values of λ and γ from $(.90, .95, .98)$ and $(\frac{1}{4}, \frac{1}{2}, \frac{3}{4})$ respectively.[3] These values suggest a large, but relatively slow, mean-reverting component to stock prices. The practical implications of such a model is that these low frequencies can only be picked up by multiperiod data — for example, by looking at sums of autocorrelations.

Substituting the parameter values for λ and γ into the solution for the optimal test given above, we calculate the test statistic with the maximal approximate slope. Table 2 compares the approximate slope of this optimal test to some of the existing tests in the literature. For low values of γ (i.e. for alternatives close to the random walk), the optimal test's approximate slope is almost twice the slope of these other tests. For high values of γ (in which mean-reversion plays a dominant role), the approximate slope is over six times that of the other tests. For example, consider the results for $\lambda = .95$ and $\gamma = \frac{1}{2}$. One of the most popular test statistics in the literature, the variance ratio, is only 24.2%

[3]These values coincide with those given in Poterba and Summers (1988) and implied in the Fama and French (1988) analysis.

of the optimal test's approximate slope value of .00641.[4] The interpretation, therefore, is that approximately 4 times as many observations are needed by the variance ratio to achieve the same power.

For $\lambda = .95$, Figure 1 plots the number of $D_i \in D$ and their respective values for various test statistics; including the optimal test, the optimal variance ratio statistic, the optimal multiperiod autocorrelation and the optimal J-statistic.[5] The optimal weights most resemble those of the variance-ratio statistic. In contrast, the J-statistic places too much weight on higher order autocorrelations relative to low order ones. The multiperiod autocorrelation fares even worse by placing relatively little weight on the informative low order autocorrelations. For example, it places over 10 times more weight on ρ_{14} than on ρ_1 while the optimal statistic places only $\frac{1}{2}$th as much weight on ρ_{14} than on ρ_1.

Note though that the approximate slope results in Table 2 do not correspond exactly with the optimal weights given in Figure 1 — for example, in Table 2 the variance-ratio statistic has less asymptotic power than the J-statistic. The reason for this is that the approximate slope in (4) has two components, A and B. While the variance-ratio statistic's weights are declining and therefore close to optimal with respect to B, the choice of $j = 1$ in A is far from optimal. In contrast, the J-statistic chooses large j in A, but fails to capture the declining weights required by optimization of B. As an illustration, for $\lambda = .95$ and $\gamma = \frac{1}{2}$, about $\frac{3}{4}$th of the loss in efficiency for the variance ratio comes from component A, while only about $\frac{1}{2}$th comes from component A for the J-statistic.

This analysis also helps explain the different conclusions of various authors regarding the power of statistics against mean-reversion. For example, consider the J-statistic and the variance ratio statistic (see, for example, Lo and MacKinlay (1989) and Jegadeesh (1989)). The somewhat similar power of these tests is due to different reasons, namely either the optimal declining weights in A (i.e. as with the variance ratio statistic) or the large j multiperiod variance dividor in B (i.e. as with the J-statistic). Of course, the optimal statistic captures both the declining weights and the large choice of j, leading to a much higher approximate slope and as we shall see below higher power.

[4] Note that we chose the optimal variance ratio (i.e. the j-period variance ratio with maximum approximate slope over all j) for comparison with the optimal test. For example, for $\lambda = .95$ and $\gamma = \frac{1}{2}$, $j = 42$.

[5] The J-statistic is the t-statistic from a regression of one period on multiperiod past values given in Jegadeesh (1989). See Table 2 for its exact form.

4.1.1 Simulation results

Since the approximate slope is a large sample theory, the results in Table 2 are valid only asymptotically. It is of some interest, therefore, to study the small sample behavior of the statistics via Monte Carlo simulation. The simulation consists of 5000 replications with $\lambda = .95$, $\gamma = \frac{1}{2}$ and ϵ_t and η_t drawn from i.i.d. normal distributions. Table 3 reports the small sample power of the statistics given in Figure 1 and the size of the nonrejection region Λ^* (under fixed 90% power) for values of $T = 360$, 720, 1440, 2880 and 5760.[6]

As the approximate slope theory suggests, the optimal test has higher power than the other statistics. The most comparable statistic in terms of power is the J-statistic which interestingly enough is also the most comparable in terms of approximate slope. Of particular interest, the ratio of the nonrejection regions Λ^* are of similar magnitude to the ratio of approximate slopes. For example, at $T = 1440$, the ratio of the J-statistic's Λ^* to the optimal statistic's is .417 while the ratio of their approximate slopes is .359. Hence, even in light of the well-documented small sample problems of multiperiod autocorrelation statistics, the approximate slope results provide a fairly accurate assessment of the relative power of these statistics.

5 Extensions

5.1 Alternative Models

It is straightforward to calculate the approximate slopes of the statistics in (3) under most interesting alternatives. All that the econometrician requires is the $plim(\hat{\rho}(j))$.

For example, there is recent evidence suggesting that stock returns are positively autocorrelated at short-horizons (see, for example, Lo and MacKinlay (1989)). One suggested model has been a first-order autoregressive process in returns:

$$P_t - P_{t-1} = \mu + \rho(P_{t-1} - P_{t-2}) + \epsilon_t, \quad E[\epsilon_t] = 0, E[\epsilon_t \epsilon_{t-j}] = 0, \quad \rho < 1.$$

Choosing j and D to maximize (3), it is possible to show that the optimal statistic picks $j = 1$ and the weights $D_i = \rho^i$. While this process is nonstationary in prices,

[6] With respect to the optimal test statistic, we let $j = 120$ and let the last weight equal D_{120}. While these values were chosen somewhat arbitrarily, they should capture most of the optimal statistic's implied power. However, since we limit $j = 120$ we should not expect the ratio of the nonrejection region Λ^* to exactly equal the ratio of approximate slopes.

the optimal weights are still declining in the autocorrelations. The primary difference is that the econometrician uses the single-period sample variance (rather than the sample multiperiod variance) as the estimator for the variance of returns. This analysis suggests that the standard variance ratio statistic should be close to the optimal test statistic. The variance ratio captures both the declining weights (although not the exact ones) and the choice of $j = 1$.

5.2 Alternative Classes of Statistics

The class of statistics studied in this paper are linear combinations of consistent estimators of autocorrelations. As special cases, this class captures many of the statistics studied in the recent finance and macroeconomics literature. These statistics are especially suited to testing no serial correlation against vague alternatives.

If we imposed the alternative model directly, more powerful tests may be available. For example, one could impose the Fama-French/Poterba-Summers mean-reversion alternative, estimate γ and λ, and test for no serial correlation in this framework. Conceivably, the approximate slope analysis could also be conducted for these alternative test statistics. With respect to finite sample results, Faust (1989) finds that under the assumption of normality the most powerful statistics against mean-reversion resemble filtered variance-ratios (i.e. weighted squared multiperiod variances). Note that filtered variance-ratios fall into the class of statistics analyzed in this paper. There is some reason to suspect, therefore, that the optimal statistic will have comparable power to statistics designed against specific alternatives. Nevertheless, this remains an open question for further research.

6 Conclusion

One way to evaluate the merits of a test statistic is to consider its power against "interesting" alternatives. When evaluating power via Monte Carlo simulation, the econometrician faces an identification problem. By having to specify the complete structure of the alternative, he cannot determine whether the statistic's power is due to "essential" parameters or simply nuisance parameters governing the alternative specification. For example, consider the stock price mean reversion literature and specifically the alternative model proposed by Fama and French (1988), among others. It is arguable that we have information concerning only a few parameters governing this model's stock price distribution (e.g. perhaps the mean, variance and mean-reversion parameter); other assumptions like normally distributed innovations go beyond the theory, yet are commonly made in Monte Carlo simulations.

This paper proposes an alternative method for evaluating power based on Geweke's (1981) approximate slope procedure. Consistent with the theory behind the proposed alternative models, it places weak assumptions on the alternative's specification. In addition, as well as being robust to various parameterizations, the procedure is simple to use in practice. With respect to tests for serial correlation, all that the econometrician needs to calculate are the true autocorrelations of the series under the alternative model. Of special interest to the resent finance and macroeconomics literature, the relative asymptotic power of statistics within a general class is explicitly quantified.

TABLE 1

Table 1 provides examples of a general class of statistics for testing whether a series $P_t - P_{t-1}$ is uncorrelated. Specifically, this class can be represented as linear combinations of consistent autocorrelation estimators:

$$
\begin{aligned}
D\hat{\rho}(j) &= \sum_{i \in (D_k \ldots D_l)} D_i \hat{\rho}_i(j) \\
\text{where } \hat{\rho}_i(j) &= \frac{\sum_{t=1}^{T}(P_t - P_{t-1} - \hat{\mu})(P_{t-i} - P_{t-i-1} - \hat{\mu})}{\sum_{t=1}^{T}(P_t - P_{t-j} - j\hat{\mu})^2/j} \\
\hat{\mu} &= \sum_{t=1}^{T}(P_t - P_{t-1})
\end{aligned}
$$

This class includes many of the statistics currently employed in the finance and macroeconomics literature; including the variance-ratio statistic and multiperiod autocorrelation, among others.

Statistic	$\hat{\rho}_i$ **Representation** $D\hat{\rho}(j)$	**Typical Weights** $D_i \in D$
Autocorrelation $\hat{\beta}_j = \frac{\sum_{t=1}^{T}(P_t-P_{t-j}-j\hat{\mu})(P_{t-j}-P_{t-2j}-j\hat{\mu})}{\sum_{t=1}^{T}(P_t-P_{t-j}-j\hat{\mu})^2}$	$\sum_{i=1}^{2j-1} min(i, 2j-i)\hat{\rho}_i(j)/j$	$D_i = min(i, 2j-i)/j$
Variance Ratio $\hat{V}_j = \frac{\sum_{t=1}^{T}(P_t-P_{t-j}-j\hat{\mu})^2}{j\sum_{t=1}^{T}(P_t-P_{t-1}-\hat{\mu})^2} - 1$	$2\sum_{i=1}^{j-1}\frac{(j-i)}{j}\hat{\rho}_i(1)$	$D_i = 2(j-i)/j$
J Statistic $\hat{\beta}(1,j) = \frac{\sum_{t=1}^{T}(P_t-P_{t-1}-\hat{\mu})(P_{t-1}-P_{t-j-1}-j\hat{\mu})}{\sum_{t=1}^{T}(P_t-P_{t-j}-j\hat{\mu})^2}$	$\sum_{i=1}^{j}\hat{\rho}_i(j)/j$	$D_i = 1/j$
Box-Pierce $Q_j = \sum_{i=1}^{j-1}\left\{\frac{\sum_{t=1}^{T}(P_t-P_{t-1}-\hat{\mu})(P_{t-j}-P_{t-j-1}-\hat{\mu})}{\sum_{t=1}^{T}(P_t-P_{t-1}-\hat{\mu})^2}\right\}^2$	$\sum_{i=1}^{j-1}\hat{\rho}_i(1)^2$	$D = I$ (in eqn.2)

TABLE 2

Table 2 presents approximate slope calculations for various statistics under the null hypothesis that stock prices have a slowly mean reverting component. Specifically, the null and alternative hypotheses are given by.

$$\begin{aligned}
\textbf{Null} : P_t &= \mu + q_t \\
\textbf{Alternative} : P_t &= \mu + q_t + z_t \\
\text{where } q_t &= q_{t-1} + \epsilon_t, \quad E[\epsilon_t] = E[\epsilon_t \epsilon_{t-j}] = 0 \\
z_t &= \lambda z_{t-1} + \eta_t, \quad E[\eta_t] = E[\eta_t \eta_{t-j}] = 0, \ |\lambda| < 1.
\end{aligned}$$

The approximate slope measures the relative asymptotic power of the statistics when testing the null (under a particular alternative). The table provides calculations for different values of λ and γ, where γ measures the share of the variance captured by the mean-reverting component z_t. Note that the optimal test is the test statistic with the highest approximate slope, $\hat{V}_j$ is the optimal variance ratio, $\hat{\beta}_j$ is the optimal multiperiod autocorrelation, $\hat{\beta}(1,j)$ is the optimal coefficient from a regression of one-period on past j-period values, and $\hat{\beta}_1$ is the single period autocorrelation.

		Slope	**Relative % of Optimal Test**			
Parameters		**Optimal**	$\hat{V}_j$	$\hat{\beta}_j$	$\hat{\beta}(1,j)$	$\hat{\beta}_1$
$\lambda = .90$	$\gamma = \frac{1}{4}$	.00146	54.5%	39.5%	57.8%	10.7%
	$\gamma = \frac{1}{2}$	.01316	24.2%	20.0%	35.4%	4.7%
	$\gamma = \frac{3}{4}$	.11842	6.1%	5.9%	15.3%	1.2%
$\lambda = .95$	$\gamma = \frac{1}{4}$	.00071	54.6%	38.3%	58.3%	5.5%
	$\gamma = \frac{1}{2}$	.00641	24.2%	19.5%	35.9%	2.4%
	$\gamma = \frac{3}{4}$	.05769	6.1%	5.8%	15.7%	0.6%
$\lambda = .99$	$\gamma = \frac{1}{4}$	.00028	54.6%	37.5%	58.5%	2.2%
	$\gamma = \frac{1}{2}$	.00253	24.2%	19.2%	36.2%	1.0%
	$\gamma = \frac{3}{4}$	.02273	6.1%	5.8%	15.9%	0.2%

TABLE 3

Table 3 compares the small sample power of the statistics in Table 2 under Table 2's alternative with $\lambda = .95$, $\gamma = \frac{1}{2}$ and ϵ_t and η_t drawn from normal distributions. The statistics have the form:

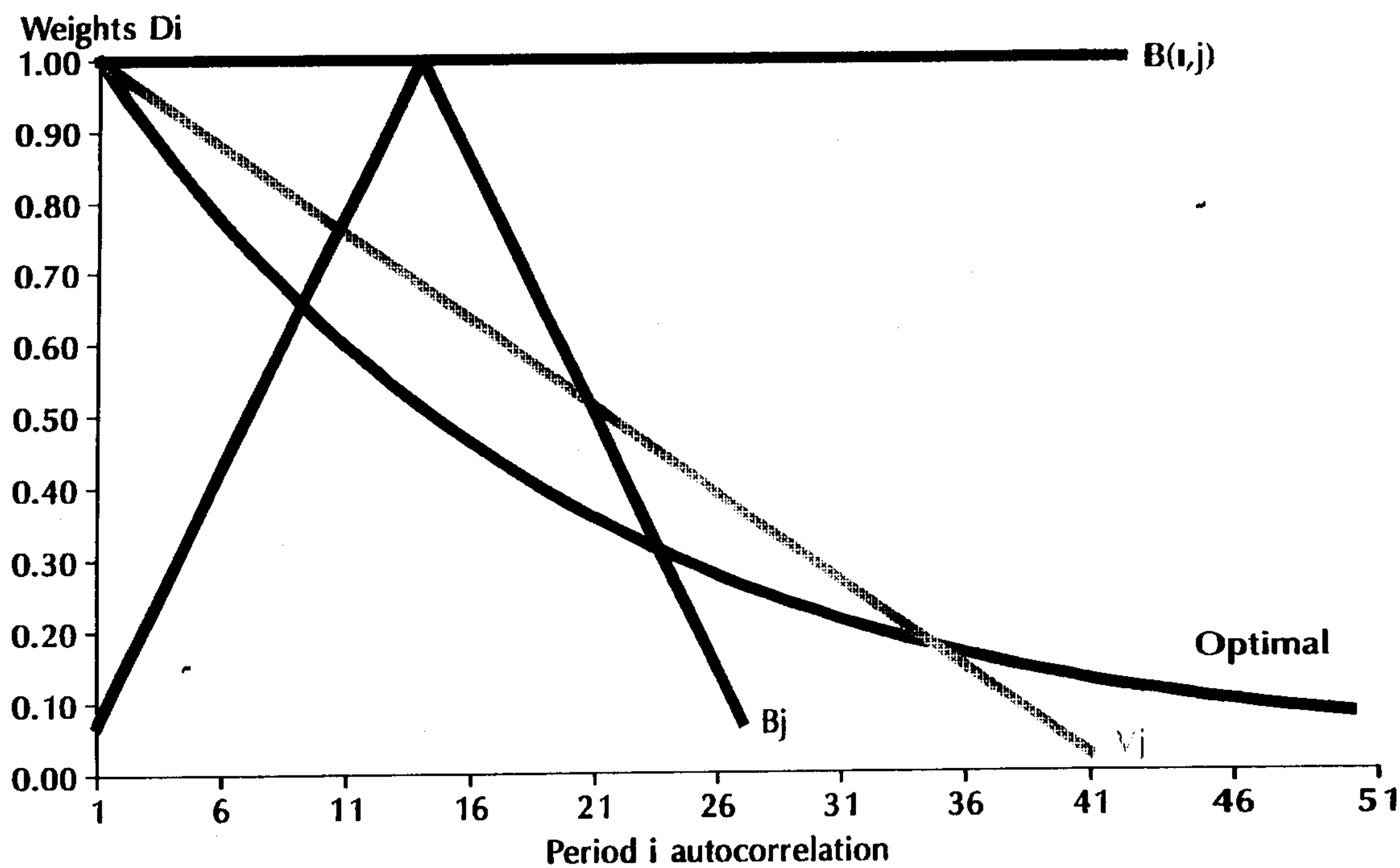
$$J_T \equiv T \frac{\left[\sum_{i \in (D_k \ldots D_l)} D_i \hat{\rho}_i(j)\right]^2}{\sum_{i \in (D_k \ldots D_l)} D_i^2}.$$

Using empirical 5% cut-off levels, power is evaluated for different values of T. In addition, for a fixed power of 90% and given T, the size of the nonrejection region Λ^* over 5000 replications is calculated. Note that the optimal test is the test statistic with the highest approximate slope, $\hat{V}_j$ is the optimal variance ratio, $\hat{\beta}_j$ is the optimal multiperiod autocorrelation, $\hat{\beta}(1,j)$ is the optimal coefficient from a regression of one-period on past j-period values, and $\hat{\beta}_1$ is the single period autocorrelation.

# Observations		Optimal	$\hat{V}_j$	$\hat{\beta}_j$	$\hat{\beta}(1,j)$	$\hat{\beta}_1$
$T = 360$	Emp. Power	28.26%	5.82%	11.50%	13.30%	6.28%
	$\frac{\Lambda_i^*}{\Lambda_j^*}$	$\Lambda_j^* = .16$	.438	.188	.430	.125
$T = 720$	Emp. Power	34.26%	14.72%	17.66%	23.34%	6.14%
	$\frac{\Lambda_i^*}{\Lambda_j^*}$	$\Lambda_j^* = .27$	.481	.185	.519	.074
$T = 1440$	Emp. Power	58.16%	31.94%	26.96%	45.56%	6.90%
	$\frac{\Lambda_i^*}{\Lambda_j^*}$	$\Lambda_j^* = 1.12$	.473	.134	.438	.018
$T = 2880$	Emp. Power	84.64%	65.02%	48.38%	71.94%	9.40%
	$\frac{\Lambda_i^*}{\Lambda_j^*}$	$\Lambda_j^* = 3.69$	.455	.176	.455	.005
$T = 5760$	Emp. Power	98.68%	94.48%	79.40%	94.76%	16.82%
	$\frac{\Lambda_i^*}{\Lambda_j^*}$	$\Lambda_j^* = 11.24$	.426	.228	.495	.003

FIGURE 1

Normalized Optimal Weights for Test Statistics



j = Multiperiod Autocorrelation
j = Variance Ratio
(i,j) = Regn. Coefficient

8 G T 1 5

References

Bahadur, R. R., 1960, Stochastic comparison of tests, *Annals of Mathematical Statistics*, 31,276–295.

Cochrane, John H., 1988, How big is the random walk in GNP?, *Journal of Political Economy*, 96,893–920.

Fama, Eugene F. and Kenneth R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy*, 96,246–273.

Faust, Jon, 1989, A variance ratio test for mean reversion: statistical properties and implementation, Federal Reserve Bank of Kansas City.

Geweke, J., 1981, The approximate slope of econometric tests, *Econometrica*, 49(6),1427–1442, November.

Hansen, L., 1982, Large sample properties of generalized method of moment estimators, *Econometrica*, 50(4),1029–1054, july.

Jegadeesh, Narashimhan, 1989, Seasonality in stock price mean reversion: Evidence from the U.S. and U.K., Working paper (Anderson Graduate School of Management, University of California at Los Angeles, Los Angeles, CA).

Lo, Andrew W. and A. Craig Mackinlay, 1989, The size and power of the variance ratio test in finite samples: a Monte Carlo investigation, *Journal of Econometrics*, 40,203–238.

Poterba, James M. and Lawrence H. Summers, 1988, Mean reversion in stock prices: Evidence and implications, *Journal of Financial Economics*, 22,27–59.

Richardson, Matthew and Tom Smith, 1989, Tests of financial models in the presence of overlapping observations, forthcoming in Review of Financial Studies.

Schwert, G. William, 1989, Why does stock market volatility change over time?, *Journal of Finance*, 44,1115–1154.