WP/00/37

IMF Working Paper

Concordance in Business Cycles

C. John McDermott and Alasdair Scott

INTERNATIONAL MONETARY FUND

IMF Working Paper

Research Department

Concordance in Business Cycles

Prepared by C. John McDermott and Alasdair Scott[1]

Authorized for distribution by Peter Wickham

March 2000

Abstract

We study the properties of a test that determines whether two time series comove. The test computes a simple nonparametric statistic for "concordance," which describes the proportion of time that the cycles of two series spend in the same phase. We establish the size and power properties of this test. As an illustration, the procedures are applied to output series from selected major industrial countries. We find limited evidence of widespread concordance for these countries.

JEL Classification Numbers: C14, E32

Keywords: Concordance, phase, business cycles

Author's E-Mail Address: Mcdermottj@rbnz.govt.nz; scotta@govt.nz

---

[1] John McDermott and Alasdair Scott are advisors at the Reserve Bank of New Zealand. This paper was completed while John McDermott was a Visiting Scholar in the Commodities and Special Issues Division of the Research Department. We received valuable comments from Paul Cashin, Paul Conway, Weshah Razzak, Christie Smith, and Peter Wickham. The views expressed in the paper are those of the authors, and do not necessarily reflect those of the IMF or the Reserve Bank of New Zealand.

Contents Page

I. Introduction

In this paper we study the properties of a test that determines whether two time series comove. The test computes a simple nonparametric statistic for "concordance," which describes the proportion of time that the cycles of two series spend in the same phase. The concordance statistic was initially proposed by Harding and Pagan (1999), but they left the issue of the distributional properties of the statistic untouched. With no distribution theory, the statistic could only be used as an indicative measure of comovement rather than for making inferences. The contribution of this paper is that it establishes, using simulation methods, the distributional properties of the statistic and the size and power properties of this test.

Section II contains a description of an algorithm used for dating business cycles and a definition of the concordance statistic. In Section III, the properties of the concordance statistic are examined, including critical values for using this statistic as a test of comovement. A comparison with two alternative measures of comovement is contained in Section IV. In Section V, the algorithm and statistics are applied to output data for selected major industrial countries. Finally, conclusions are presented in Section VI.

II. Measuring cycles

In order to utilize the concordance statistic, we need to date the peaks and troughs of cycles in the data. Business cycles are commonly described in two different ways: the "classical" cycle, which refers to peaks and troughs in the *level* of a series, and the "growth" cycle, which refers to peaks and troughs in the level of the *detrended* series. In this paper, we employ the classical measure. The distinction is important: when people refer to the NBER dates for booms and recessions in the United States, they are referring to the classical cycle characteristics of the data. However, when researchers (such as those of the real business cycle school) refer to the business cycle in filtered data, they are referring to a growth cycle. The distinction between levels and detrended data is crucial: a recession can only occur in a classical cycle if growth is negative, while in a growth cycle, a recession is a phase where actual output is below the trend.[1]

Dating classical cycles is a nontrivial exercise. This is one possible reason why modern "business cycle" analysis has segued away from the original classical sense of Burns and Mitchell (1946) towards the use of simple detrending methods such as the HP filter. One conventional rule defines a classical recession as two consecutive periods of negative growth. However, the business cycle dates published by the NBER are the result of considerable judgement. Attempts have to be made to filter out false turning points from noisy data.

[1] For further discussion, see Pagan (1997a).

Nonetheless, an algorithm for monthly data set out by Bry and Boschan (1971) of the NBER is a good approximation to the judgmental procedure.[2] The algorithm proceeds in three basic steps. First, a potential set of peaks and troughs is determined by the application of a turning point rule. This defines a local peak in the time series x_t as occurring at time t whenever $\{x_t > x_{t\pm k}\}$, $k = 1,\ldots,K$, while a local trough occurs at time t whenever $\{x_t < x_{t\pm k}\}$, $k = 1,\ldots,K$. The second step enforces the condition that peaks and troughs must alternate. Thirdly, the peaks and troughs are revised, or "censored", according a range of criteria. Under these conditions, a complete cycle must be at least fifteen months long, while all phases must be at least six months. There are further rules designed to avoid spurious cycle dating at the ends of series.

The NBER rules were derived on the assumption that data would be at monthly frequency. One issue to be resolved, therefore, is how to apply these rules to quarterly data. Where K is generally set to five for monthly data, it is set here to two for quarterly data. As per the NBER convention, cycles are at least fifteen months or five quarters long.

When peaks and troughs have been dated—by using the Bry-Boschan algorithm, for example—we can compute some basic but informative statistics. First, we document the average duration and amplitude of the phases in individual series. A comparison of the time plots of the series and their peaks and troughs allows us to quickly assess whether recent cycles are in some way unusual, or whether there is a pattern in the evolution of the cycles. We also employ a test for regular periodic behavior in the cycles. Second, we are (more) interested in how the cyclical patterns of the series compare to each other. To facilitate this, we make use of the concordance statistic originally proposed by Harding and Pagan (1999). This is a simple nonparametric statistic that measures the proportion of time two series, x_i and x_j, are in the same state. Let $\{S_{i,t}\}$ be a series taking the value unity when the series x_i is in expansion and zero when it is in a contraction. Define the series $\{S_{j,t}\}$ in the same way. The degree of concordance is then

$$C_{ij} = T^{-1}\left\{\sum_{t=1}^{T}\left(S_{i,t} \cdot S_{j,t}\right) + \left(1 - S_{i,t}\right)\cdot\left(1 - S_{j,t}\right)\right\}, \tag{1}$$

where T is the sample size.

This approach accords us several technical advantages. First, the binary indicator series S_t are, by construction, stationary, allowing us to easily apply any number of statistical measures. Second, the Bry and Boschan algorithm and the concordance statistic are

[2] This algorithm has been used by King and Plosser (1994), Watson (1994) and Cashin, McDermott and Scott (1999a).

entirely nonparametric. This means that the dating will be almost entirely independent of the sample used, which is not the case in "global" parametric growth cycle models.[3]

III. Properties of the Concordance Statistic

In this section, we present the results of Monte Carlo simulations that examine the properties of this statistic. As a proportion, the values that the expression (1) may take are clearly bounded between zero and one. Faced with an empirical result of, say, 0.7, it is natural to assume that this is a large number relative to zero. However, there are factors that mean that the appropriate distribution of the statistic may be centered at 0.5 or higher.

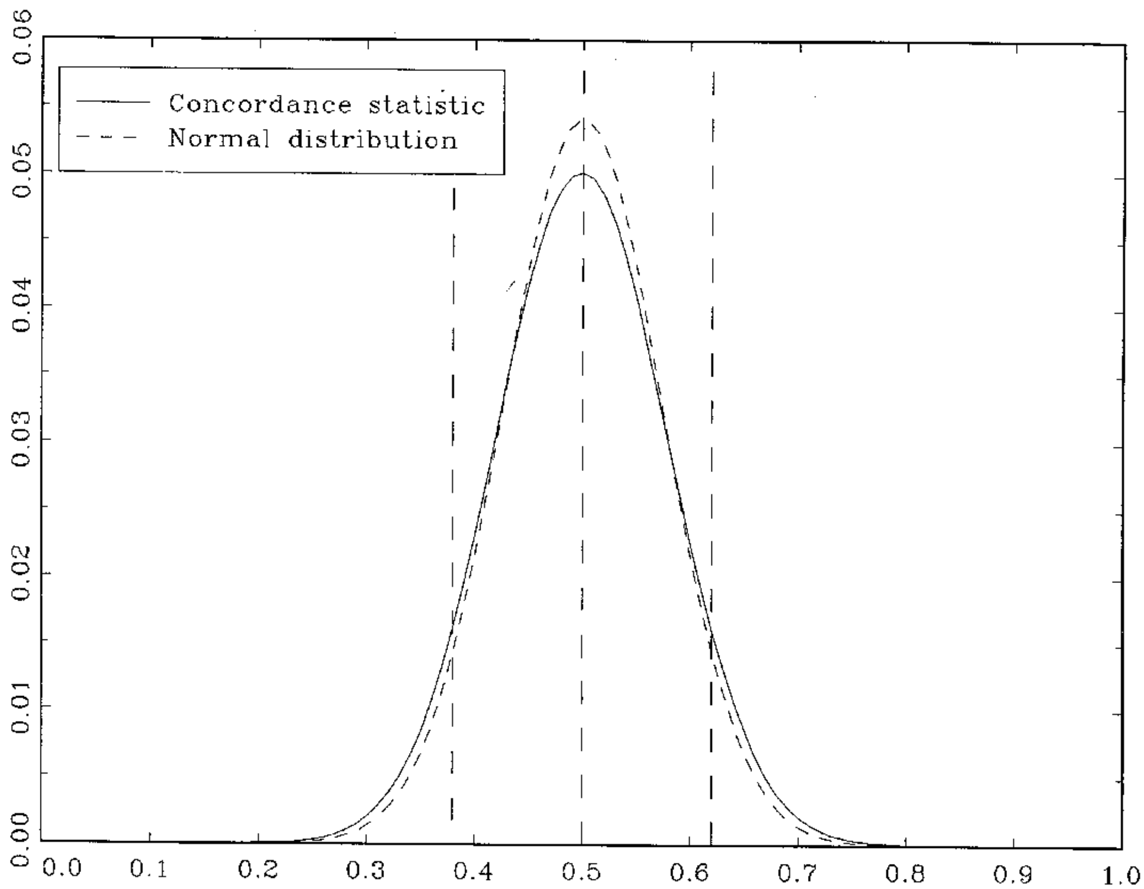
Consider a concordance statistic evaluated over two series, x_1 and x_2, when both series follow a Brownian motion. In this case there is an equal chance that the series are in or out of phase with each other, and the distribution of the statistic will be symmetric around 0.5. If the two series x_1 and x_2 were independent, then the variance of the concordance statistic would be $1/[4(T-1)]$, where T is the sample size of x_1 and x_2. However, since the Bry-Boschan algorithm involves censoring, then independence cannot be assumed. The censoring operation makes it difficult to produce an exact analytical solution to the distribution, so we use Monte Carlo simulation to generate the distribution of the concordance statistic. The distribution of the concordance statistic when the two series follow a Brownian motion is compared to the normal distribution in Figure 1. Even for this simple case with no drift, the distribution of the statistic deviates noticeably from the normal distribution. An increase in the minimum phase and the minimum cycle in the censoring rule increases the variance and thickness of the tails.

Now consider the case when x_1 and x_2 are Brownian motion with drift. The existence of a positive growth rate will shift the center of the distribution to the right, with the result that a large area of the distribution may be confined to the region just to the left of one. The skewness in the distribution caused by the drift can be seen in Figure 2.

---

[3] The exception comes about because of the censoring rules. To the extent that these affect the dating of peaks or troughs at the end of the series, an extended sample may result in slightly different datings. This reflects the "pattern-recognition" nature of the algorithm: more data mean that there will be more information with which to decide whether a local peak or trough is a cycle peak. A peak or trough which was not dated in the smaller sample may therefore be dated in the extended sample (note, however, that the reverse does not apply).

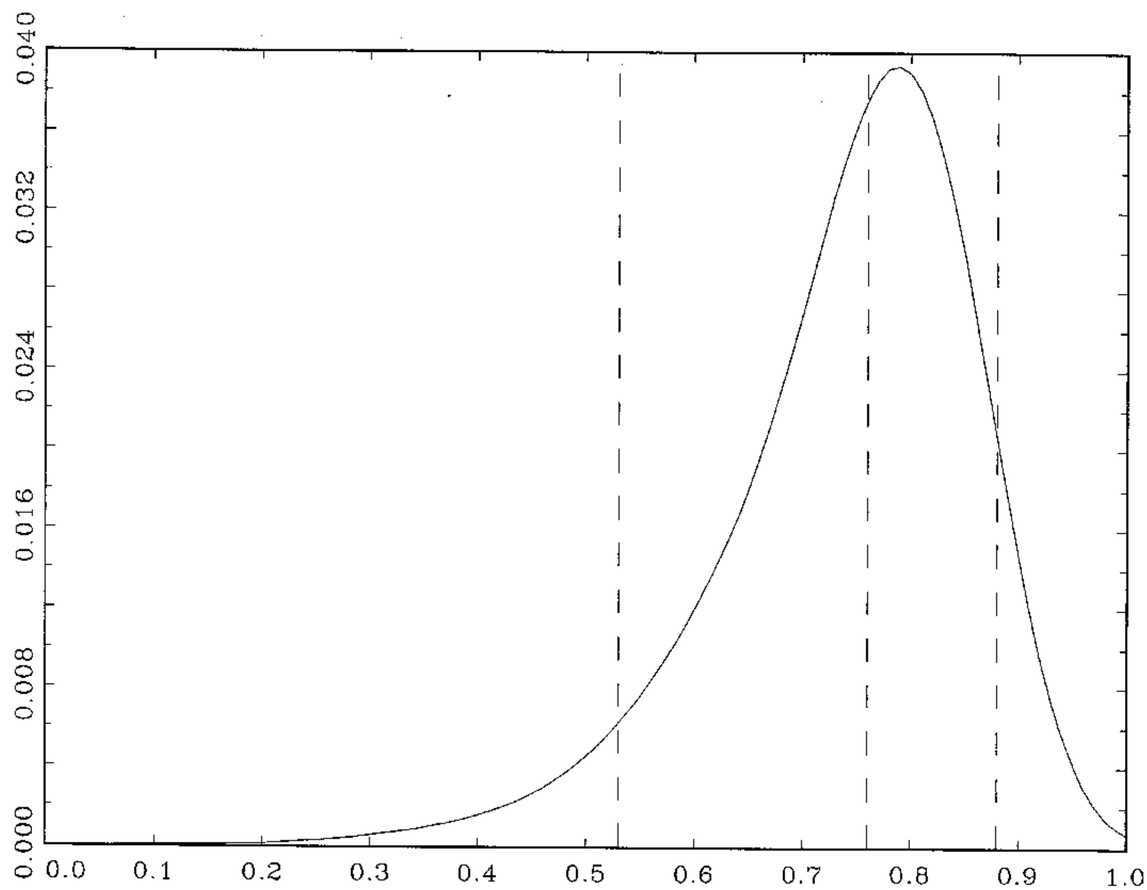
Figure 1. Distribution of Concordance Statistic, Zero Drift Case



Source: Authors' calculations.

Notes: Distribution computed from draws of two sets of 100 observations, each generated independently from a Brownian motion. Censoring rules: minimum cycle 5 periods, minimum phase 2 periods.

Figure 2. Distribution of Concordance Statistic, Positive Drift Case



Source: Authors' calculations.

Notes: Distribution computed from draws of two sets of 100 observations, each generated independently from $x_{i,t} = \mu + x_{i,t-1} + e_{i,t}$, $e_{i,t} \sim N(0, \sigma^2)$. Censoring rules: minimum cycle 5 quarters, minimum phase 2 quarters.

In the case where x_1 and x_2 are random walks with no drift, the variance of the innovations has no effect on the distribution. The intuition is simple: concordance summarizes the results of the turning points in x_1 and x_2, but not the amplitude of the swings. However, this is *not* the case when one or both of the series has drift: in that case the concordance will depend on how strong the trend is relative to the variance in the series.

Fortunately, when the two series x_1 and x_2 have different growth rates, the amount by which the distribution shifts depends solely on the smallest growth rate. An intuition for this can be gleaned from thinking of a hypothetical extreme case, where one series is drawn from a random walk with no drift and another from a random walk with infinite drift. We would expect the first series to be in each phase close to half the time, whereas the second would be in a permanent state of expansion. Nonetheless, the concordance between the two series would be roughly half; it takes drift common to both series for the expectation of the statistic to rise above 0.5.

This allows us some simplification. To calculate critical values, we assume that economic data can be well described by a random walk with drift. We regard this as a reasonable and sensible characterization. First, it is consistent with the well-known literature going back to Nelson and Plosser (1982) that identifies random walks in many economic time series. To this we add the stylized fact that many series do not exhibit duration dependence, so that there is no evidence from regular periodic behavior to support the notion of mean reversion (see Section V). Second, even if a time series does not have a unit root, the distinction is likely to be unanswerable on the basis of finite observations. For our purposes, the simple random walk with drift model is likely to be a good approximation to the properties of the data.[4]

Under this assumption, critical values for the concordance statistic can be simulated as a function of two parameters, the sample size, T, and the ratio of the drift, μ, to the standard error, σ. We confirmed this by Monte Carlo simulation.[5] The fact that it is only the

---

[4] For further discussion, see Pagan (1997b). Pagan (1999) and Harding and Pagan (1999) show that the random walk with drift model performs well compared to more complicated models (such as models with GARCH errors and regime-switching) for, respectively, output and stock price data.

[5] Two experiments were run. In the first, data from a random walk with zero drift were simulated so as to generate a distribution for the concordance statistic. The standard error of the data generating process was scaled up and down and the distribution re-generated. In the second experiment, data were generated from a random walk with positive drift and both μ and σ scaled up and down so as to maintain the same μ/σ ratio. Aside from minor numerical simulation error, the resulting distributions within the two experiments were identical. Tables of percentiles are available on request from the authors.

ratio μ/σ that matters reflects the notion that classical cycles can be summarized by their "triangular" properties—the amplitude of the cycle from peak to trough or trough to peak relative to the distance covered (that is, time elapsed). Equivalently, the parameters μ and σ summarize the probability of switching from one state to another—from a contraction to an expansion, for example. That is, if $\Delta \log x_t$ was $N(\mu, \sigma^2)$ then $\Pr(S_t = 1)$ would be a function of μ/σ exclusively (see Pagan (1999), p.14).

Instead of simply providing tables of critical values for a few specific sample sizes and μ/σ ratios, we estimate response surface regressions. These regressions have the advantages of both smoothing the simulation results and providing interpolated values for those critical values not provided by the original simulations. The regressions relate the 1 percent, 5 percent, and 10 percent upper-tail critical values for the concordance test statistic to the sample size and drift: standard error ratio.[6] Simulation experiments of 10,000 replications were conducted for a single value of T and a single value of μ/σ. The 1 percent, 5 percent, and 10 percent empirical quantiles for these data were then calculated, and each of these became a single observation in the response surface regression.

After some experimentation, the following functional form for the response surface regression was found to work well:

$$C_k(p) = \left[1 + \exp\left(-\beta_1 \cdot T_k^{-1/2} - \beta_2 \cdot \left(\frac{\mu}{\sigma}\right)_k - \beta_3 \cdot \left(\frac{\mu}{\sigma}\right)_k^2\right)\right]^{-1} + \varepsilon_k \tag{2}$$

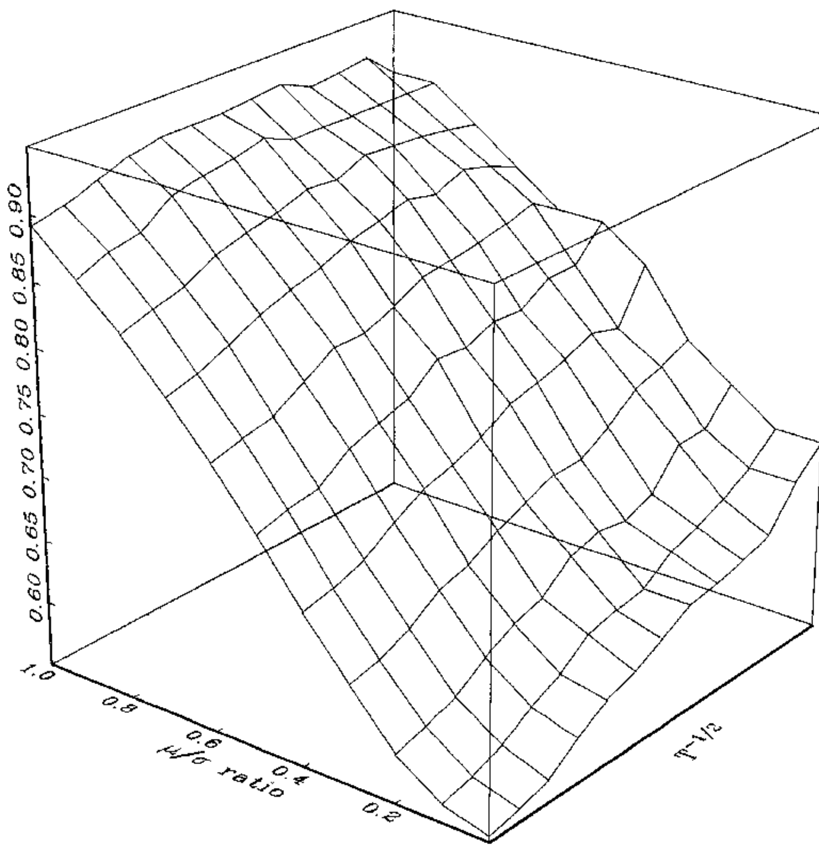
where $C_k(p)$ denotes the p percent quantile estimate for the k^{th} simulation experiment, T_k denotes the sample size for that experiment, $(\mu/\sigma)_k$ denotes the drift:standard error ratio for that experiment, the βs are parameters to be estimated, and ε_k is the residual. The form of the regression constrains the asymptotic critical value to 0.5 when there is zero drift. The parameters determine the shape of the response surface for finite T and nonzero drift.

As noted, the concordance statistic is bounded by construction in the range [0,1], and in simulation the values tend to "curl" asymptotically toward 1 as μ/σ increases and T decreases. This flattening can be seen in Figure 3a, which shows a surface plot of the naïve estimates of the critical values from the Monte Carlo simulation. This explains the use of the nonlinear logit function, which both provides a better fit and ensures that the smoothed

---

[6] We assume here that we are only interested in the existence of in-phase behavior and therefore in rejecting the null hypothesis that the two series are in phase by chance. The left tail is therefore not of any interest. We restricted ourselves to calculating the 1, 5 percent and 10 percent quantiles, because most hypothesis tests are conducted at one of 1 percent, 5 percent, or 10 percent significance levels.

Figure 3a. Response Surfaces of Naive 5 Percent Critical
Values of Concordance Test



Source: Authors' calculations.

response surface remains within [0,1]. The regression also smoothes out the surface: even after a simulation of 10,000 repetitions per point on the surface, the naïve surface is still very uneven. In empirical situations where we can expect high drift rates, this simulation noise could be severely detrimental to hypothesis testing. Hence the smoothed critical values (seen in Figure 3b) can be expected to be more accurate, as well as more convenient.[7]

The results from estimating the response surface regressions using maximum likelihood are shown in Table 1.[8] The influence of the sample size variable in the regression is strong, indicating the value of using finite sample distributional results. The estimated coefficients on the polynomial in μ/σ determine the final shape of the response surface. To illustrate the use of the response surface, if the user were dealing with a sample size of 155 and a minimum drift:standard error ratio of 0.32, the 5 percent significance level would be $\left[1+\exp\left(-4.78/\left(\sqrt{155}\right)-0.8\cdot 0.32-1.23\cdot(0.32)^2\right)\right]^{-1}=0.68$.

Table 1. Response Surface Parameters for the 1, 5, and 10 Percent Significance Levels of the Concordance Statistic

$$C_k(p)=\left[1+\exp\left(-\beta_1\cdot T_k^{-1/2}-\beta_2\cdot\left(\frac{\mu}{\sigma}\right)_k-\beta_3\cdot\left(\frac{\mu}{\sigma}\right)_k^2\right)\right]^{-1}+\varepsilon_k$$

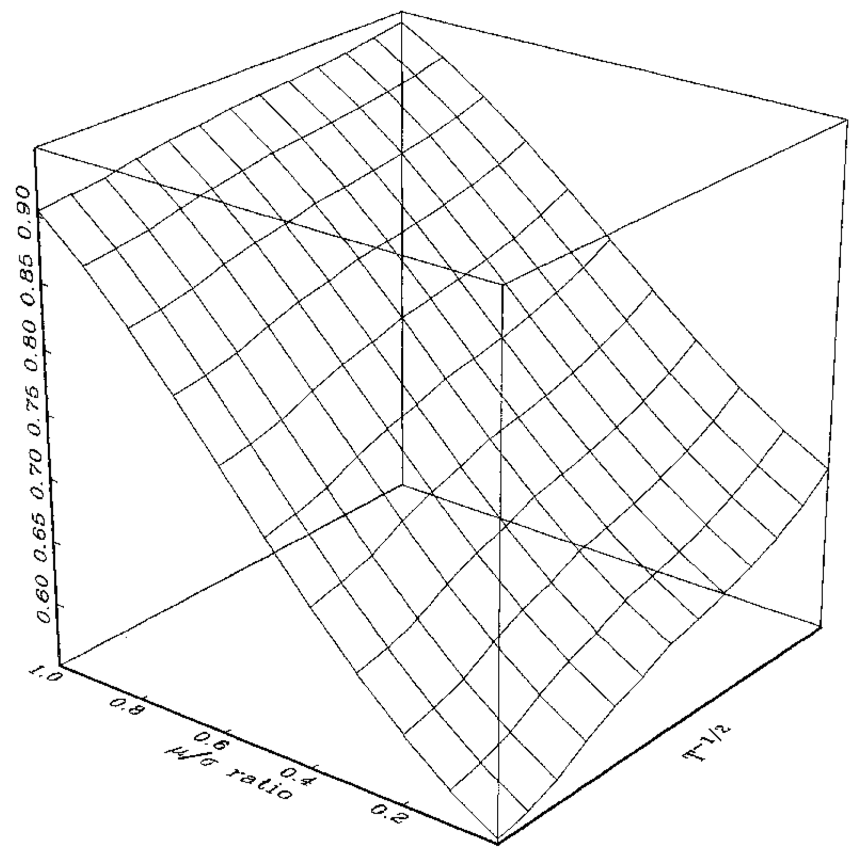
Significance level	β_1	β_2	β_3
10 percent	3.42	0.92	1.02
	(1.16)	(0.41)	(0.60)
5 percent	4.78	0.80	1.23
	(0.96)	(0.35)	(0.54)
1 percent	7.18	0.67	1.57
	(1.35)	(0.51)	(0.89)

Notes: Standard errors in parentheses.

[7] For a discussion of response surfaces, see Chapter 21 of Davidson and MacKinnon (1993).

[8] The starting values are taken from the results of GLS regressions of the three sets of critical values on a constant, $T^{-1/2}$, μ/σ, and $(\mu/\sigma)^2$. The weights for feasible GLS estimation were computed using the inverse of the fitted values of the auxiliary regression of the squared OLS residuals on the response surface regressors. The mean squared errors from the maximum likelihood estimations are 3 to 4 times smaller than the mean squared errors from the OLS or GLS estimations.

Figure 3b. Response Surfaces of Smoothed 5 Percent Critical Values of Concordance Test



Source: Authors' calculations.

To evaluate the usefulness of the concordance test we examine how often the test rejects the null hypothesis as the correlation between the innovation in the random walks increase. This is akin to examining the power of the test as the shocks to the two processes become more and more similar. The calculation of the critical values was made under the assumption that the innovations for the two series, e_1 and e_2, are completely independent. When e_1 and e_2 are nonindependent, we would hope that the statistic would correctly distinguish this from the independent case, and the calculated concordance value would be reliably greater than the relevant critical value. At the extreme, when the two series are perfectly correlated, we expect that the test would diagnose perfect concordance.

We examine the power properties of the test by simulating two dependent random walks.[9] Figure 4 shows the proportion of rejections as a function of ρ (the correlation between the innovations in the random walk series). We observe that the probability of rejecting the hypothesis of no concordance increases in the correlation, ρ, between e_1 and e_2 (this simulation is based on the 5 percent significance level calculated earlier). For a given sample size T and a given μ/σ, the curves show the increase in the power of the test as the correlation between e_1 and e_2 increases. We also observe that for time series of length typical for economic data we require a high degree of correlation in the underlying shocks in order to have a high probability of rejecting no concordance. The power of the test increases sharply as the sample size grows.

One further way to examine the accuracy of the concordance test is to examine the size of the test under different assumptions. Specifically, we examine the robustness of the preceding calculation to nonnormality in the innovations of the random walk series. This exercise is then repeated 10,000 times. The number of times the test rejects the hypothesis of no concordance between two independent random walks should remain at 500 when using the 5 percent significance level. The entries in Table 2 are under two headings: "gap" refers to the distance between the 5 percent critical value calculated under the assumption of normally-distributed innovations and the 5 percent value calculated with the alternative distribution, and "proportion" describes the fraction of times that the test rejects the null hypothesis.

[9] The errors in the random walks were generated using a bivariate normal distribution with a correlation of ρ.

Table 2. Size of Concordance Test Under Different Distributional Assumptions

	Distribution							
μ/σ	t_{10}		t_3		$-\chi_4^2$ (with mean zero)		χ_4^2 (with mean zero)	
	Gap	Proportion	Gap	Proportion	Gap	Proportion	Gap	Proportion
0.0	0.00	0.050	0.00	0.050	0.02	0.080	0.03	0.084
0.1	0.00	0.046	0.01	0.059	0.05	0.144	-0.01	0.037
0.2	0.00	0.045	0.03	0.088	0.06	0.175	-0.03	0.015
0.3	0.01	0.056	0.06	0.181	0.07	0.222	-0.05	0.008
0.4	0.01	0.064	0.08	0.256	0.07	0.222	-0.05	0.006
0.5	0.01	0.056	0.07	0.256	0.05	0.166	-0.05	0.005
0.6	0.02	0.075	0.08	0.284	0.05	0.168	-0.02	0.017
0.7	0.01	0.068	0.07	0.243	0.03	0.116	-0.01	0.030
0.8	0.01	0.065	0.05	0.203	0.02	0.083	0.01	0.071
0.9	0.00	0.047	0.03	0.138	0.00	0.049	0.02	0.106
1.0	0.01	0.061	0.03	0.127	0.01	0.050	0.04	0.207

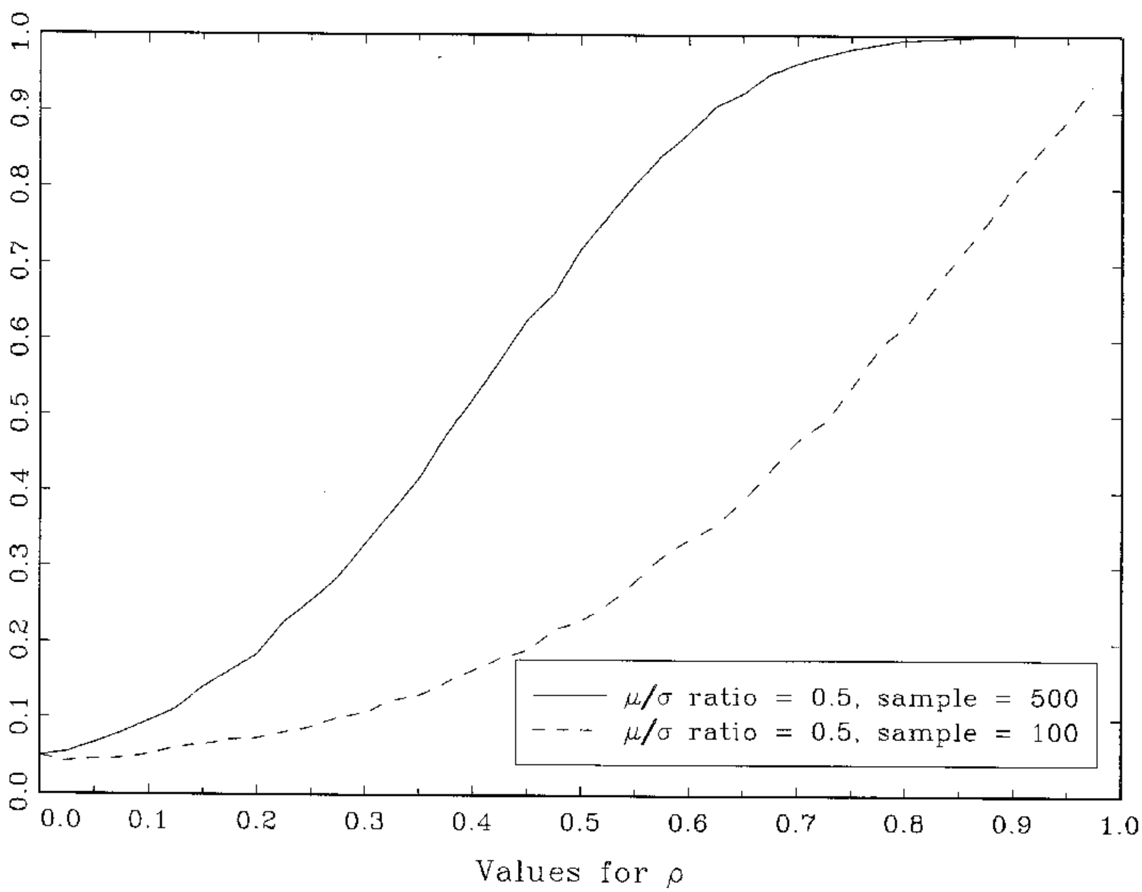
Notes: Results were computed using 10,000 Monte Carlo draws with a sample size of T=100 and a nominal size of 0.05. For each alternative distribution, the innovations are scaled such that their variance is the same as the variance of the normally-distributed innovations used to generate the critical values. "Gap" refers to the distance between the 5 percent critical value calculated under the assumption of normally-distributed innovations and the 5 percent value calculated with the alternative distribution. "Proportion" refers to the proportion of time the null hypothesis is rejected; that is, the actual size.

The distributions considered were the t_{10}, t_3, and χ_4^2 (shifted to have mean zero). The t distributions exhibit moderately and extremely thick tails, respectively. For the χ^2 distribution, we also reversed the signs on the demeaned innovations so that we have both right- and left-skewed distributed innovations. For each alternative distribution, the innovations are scaled such that their variance is the same as the variance of the normally-distributed innovations used to generate the critical values (see Figure 5).[10]

The results indicate that the test is sensitive to nonnormality of the innovations. In all cases, there are interesting interactions with the drift parameter that explain the observed size distortions.

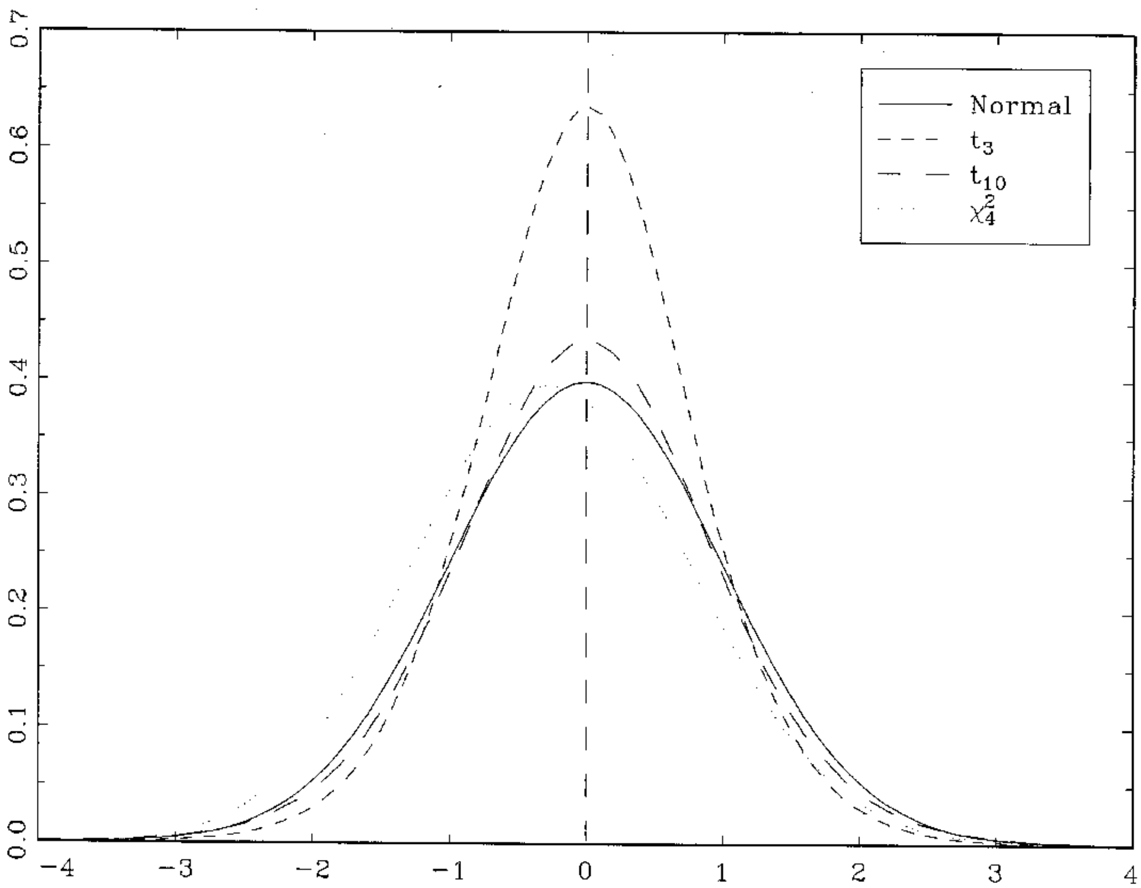
[10] For the t_k distribution, the appropriate rescaling is $\left(\sigma\sqrt{\frac{k}{k-2}}\right)^{-1}$. For the χ_k^2 distribution, the rescaling is $\left(\sigma\sqrt{2k}\right)^{-1}$.

Figure 4. Power of Concordance Statistic at 5 Percent Significance Level



Source: Authors' calculations.

Figure 5. Distributions Used for Size Testing



Source: Authors' calculations.

The t_{10} distribution exhibits mildly leptokurtic behavior. The numbers in the third column are erratic and appear within the range of simulation error, which would lead us to believe that there is no appreciable size distortion. Nonetheless, while the effect is not large, there is some sign of size distortion corresponding with the middle range of μ/σ values, where there is some tendency to overreject the null for values around $\mu/\sigma = 0.5$. This effect can also be seen in the "gap" between the critical values and the actual point where the true 5 percent of rejections occurs.

This effect becomes quite clear with the t_3 distribution. The t_3 has fatter tails than the t_{10}, and this leads to a large rightwards "bulge" of the gap associated with the midrange of μ/σ values. This is the case even though the gaps look small, which is an indication of the way in which the distribution of the statistic becomes more left skewed as the drift increases.

For the t distributions, there is no impact on the critical value at zero drift. However, drift does matter as μ/σ rises: when the innovations are scaled to have the same variance as in the normal case, this concentrates more draws in a range close to zero, therefore lowering the typical magnitude of a given shock. This means that the drift is relatively more dominant; all else being equal, we will see longer phases and cycles with this distribution. The point that gives a true 5 percent rejection rate will be greater than the critical values calculated using the normal distribution. The test therefore overrejects the null hypothesis of no concordance. However, as μ rises, concordances are pushed towards upper bound of unity. The gap between the 5 percent critical value calculated under the assumption of normality and the 5 percent values calculated with the alternative assumption closes, and we see less size distortion. For example, as we move down the fifth column of Table 2, we see that the proportion of rejections rises steadily, reaching a maximum of 28 percent when the μ/σ ratio is 0.6. The distortion then declines as the μ/σ ratio rises to 1.0.

The χ_4^2 distribution creates a distortion even without any drift. This is because it is skewed: in the case of the "negative" χ_4^2, the large rescaling required to make the variances equal brings most of the innovations into a range that is positive but close to zero. The effect is exactly as before. In the case of the "positive" χ_4^2 distribution, the effect is reversed: the effective μ/σ ratio is lowered and we observe underrejection in the midrange of μ/σ values.

IV. ALTERNATIVE MEASURES OF COMOVEMENT

Our measure of comovement is not unique, and deserves comparison with two other measures that have been used: conformity and correlation.

King and Plosser (1994) discuss a quantitative device, developed by Burns and Mitchell (1946), for assessing the degree to which a series comoves with the reference business cycle. This device, which they called conformity, is similar to our measure of concordance in that it aims to measure how much one series move with another; in their case one of the series is always the reference cycle.

Conformity between two series is determined by their relative behavior. One series is taken to be a *reference* series; the dates of peaks and troughs for the analysis are determined from this series. A series is said to conform to the reference series if it expands and contracts at broadly the same time as the reference series. That is, if during an expansion phase of the reference series the series grows at a higher rate than during the period of the contraction phase that follows, the *whole cycle* is assigned a value of 1. If, on the other hand, the series grows at a higher rate during the reference series contraction period than during the reference series expansion period, a value of -1 is assigned to the whole cycle. The degree of conformity for the series is then given by the average of ones and minus ones (multiplied by 100).

The conformity measure is therefore somewhat difficult to interpret.[11] The series and the reference series may conform perfectly, even if, say, the series is in an expansion phase while the reference series is in a contraction phase, as long as the series' expansion during the reference series contraction phase is less than its expansion during the reference series expansion phase. Thus, conformity and concordance clearly differ with respect to what they measure; which metric one should use depends on the question at hand.

One obvious reaction to our discussion of the concordance technique might be: "Why not simply calculate correlations?" There are many responses that can be given to this question. First, the two series would need to be rendered stationary. However, determining whether a time series is trend or difference stationary can be difficult given the low power of unit root tests in small samples.[12] Nor is it safe to assume differencing is required; over-differencing can be as troublesome as underdifferencing.

---

[11] Furthermore, the distributional properties of the conformity statistic are unknown.

[12] See De Jong et al. (1992).

We prefer to focus on an alternative issue. Imagine two random series. Following the discussion in Section III, we would expect the concordance statistic for these two series to be near 0.5. We would also expect their correlation of the stationary forms of the series to be zero. However, the correlation statistic is easily affected by particular, single events in the time series, which are arguably irrelevant to inferences of comovement.

To be more precise, consider an example with two independent random walks with zero drift. Figure 6 depicts such an example, where the variances of the innovations have been chosen arbitrarily in order to generate series which look like typical economic series. Allowing for some high frequency noise, the series exhibit periods of general expansion and contraction, which appear to be roughly periodic. The Bry-Boschan algorithm has been applied to the series with a minimum phase rule of two periods and a minimum cycle rule of five periods, and the peaks and troughs marked as solid and dashed lines respectively. The "bar code" at the bottom of the Figure shows when the two series are in the same phase according to the dating—that is, when the two series are in contraction or expansion at the same time, the bar code is solid, and blank when out of phase. It can be seen that the amplitudes of swings do not matter under our definition: if the series are in an expansion at the same time, then they share the same phase, even if the growth rates of the expansion are very different.[13] The concordance statistic is simply the proportion of black (or white) over the whole area of the bar graph. In the case of these two series, the concordance is exactly 0.5.

We will assume that the order of integration is known, so that in order to avoid spurious correlation, the series have been rendered stationary by first differencing. As expected, the correlation statistic is close to zero (0.12) and not significant.[14]

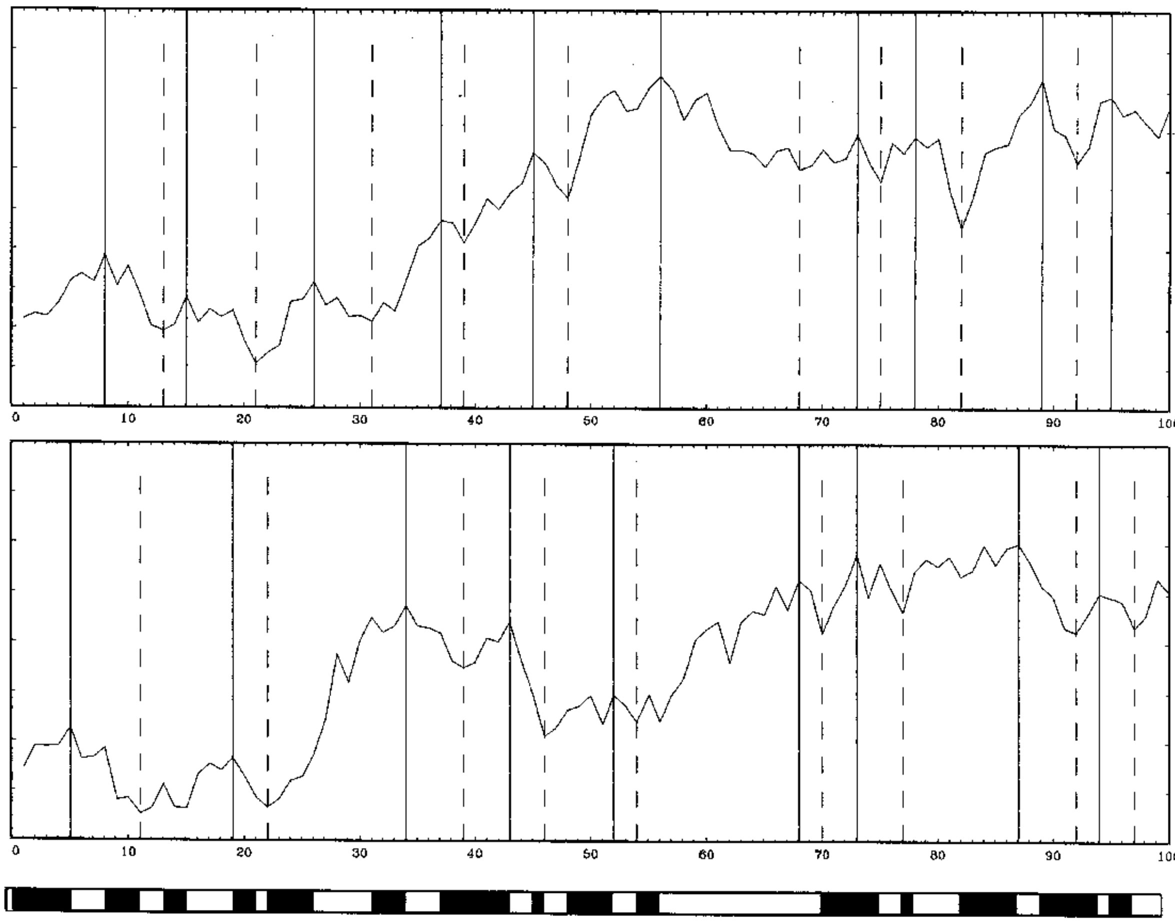
Figure 7 shows the same two series with a step function added; the jump point is exactly halfway through both series. The datings picked by the Bry-Boschan algorithm remain exactly as before. Accordingly, the concordance statistic is the same, and has not been affected by the increased amplitude in one of the phases. However, the correlation is now large (0.60) and clearly significant, even though the two series are otherwise random. This simply reflects the fact that correlation, as scaled covariance, mixes the concepts of duration and amplitude into one measure. The statistic is therefore not easily interpreted: a high number may be the result of significant comovement through time, or, as here, the result of a single large event that is common to the two series.

---

[13] In this way, this measure of comovement is firmly associated with the classical definition of the business cycle rather than the growth cycle measure.

[14] Since the test is two-sided, with a sample of 100 observations, the 95 percent confidence level is $1.96/\sqrt{T} = 0.20$.

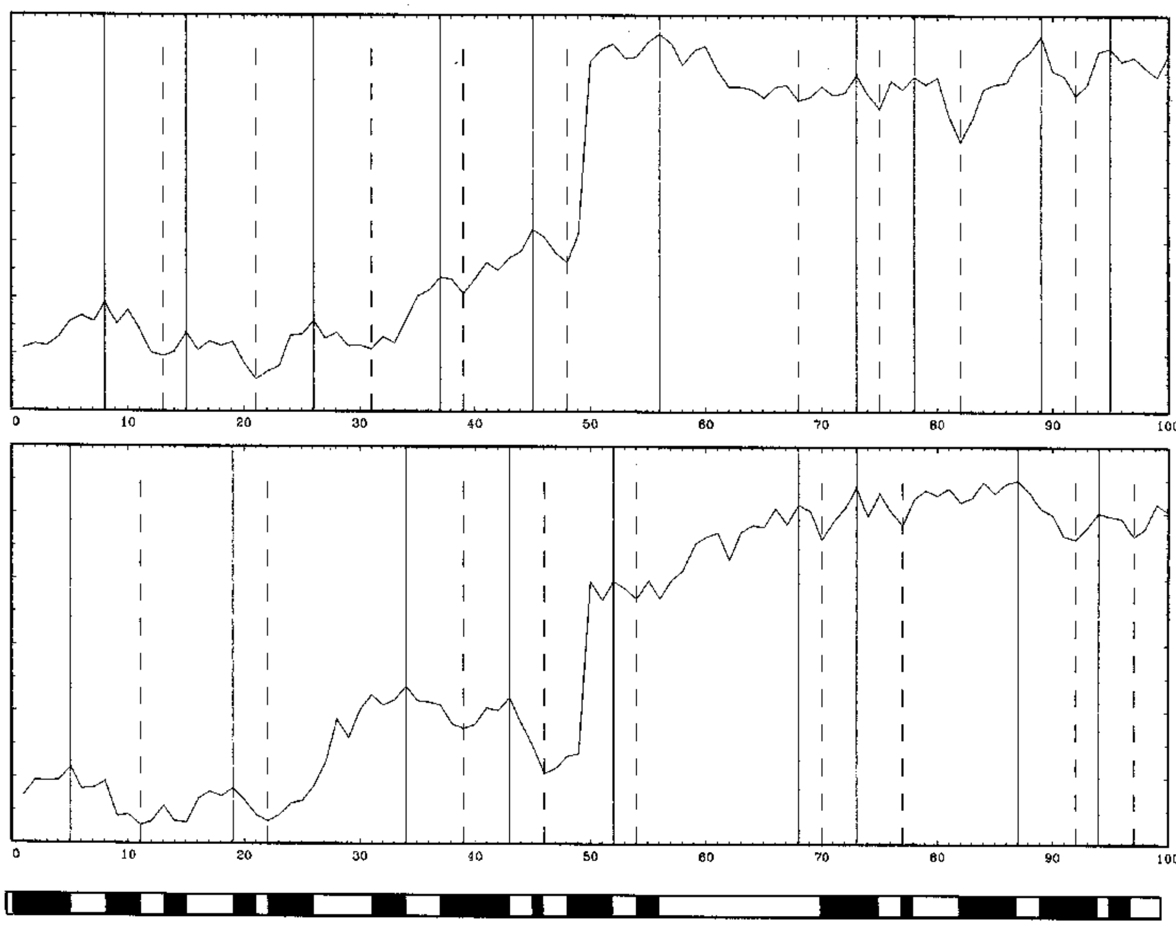
Figure 6. Peaks and Troughs of Two Random Walks (Top and Middle Panels) with Phase Indicator (Bottom Panel)



Source: Authors' calculations.

Notes: Peaks are denoted by solid lines; troughs are denoted by dashed lines. Periods from peaks to troughs are contractions, while periods from troughs to peaks are expansions. The "bar code" at the bottom of the Figure shows when the two series are in the same phase according to the dating—that is, when the two series are in contraction or expansion at the same time, the bar code is solid, and blank when out of phase.

Figure 7. Peaks and Troughs of Two Random Walks with Common Break (Top and Middle Panels) with Phase Indicator (Bottom Panel)



Source: Authors' calculations.

Notes: Peaks are denoted by solid lines; troughs are denoted by dashed lines. Periods from peaks to troughs are contractions, while periods from troughs to peaks are expansions. The "bar code" at the bottom of the Figure shows when the two series are in the same phase according to the dating—that is, when the two series are in contraction or expansion at the same time, the bar code is solid, and blank when out of phase.

It is theoretically possible to correct this problem with the use of dummy variables (see Perron (1989)), but the break point is typically unknown. Noting this problem, Zivot and Andrews (1992) suggest a modified procedure, but nonetheless both methods allow for only one possible jump point. In practice, therefore, correcting with dummy variables is highly problematic.

We avoid such problems by construction. If we define comovement in terms of the coincidence of phases, the concordance measure is both more appropriate and more easily interpretable. Given that regime shifts are a stylised fact of economic series, the concordance measure avoids the possibility of being mislead by correlations that have been highly affected by the amplitudes of single events, when the series are random and otherwise independent.

V. APPLICATION[15]

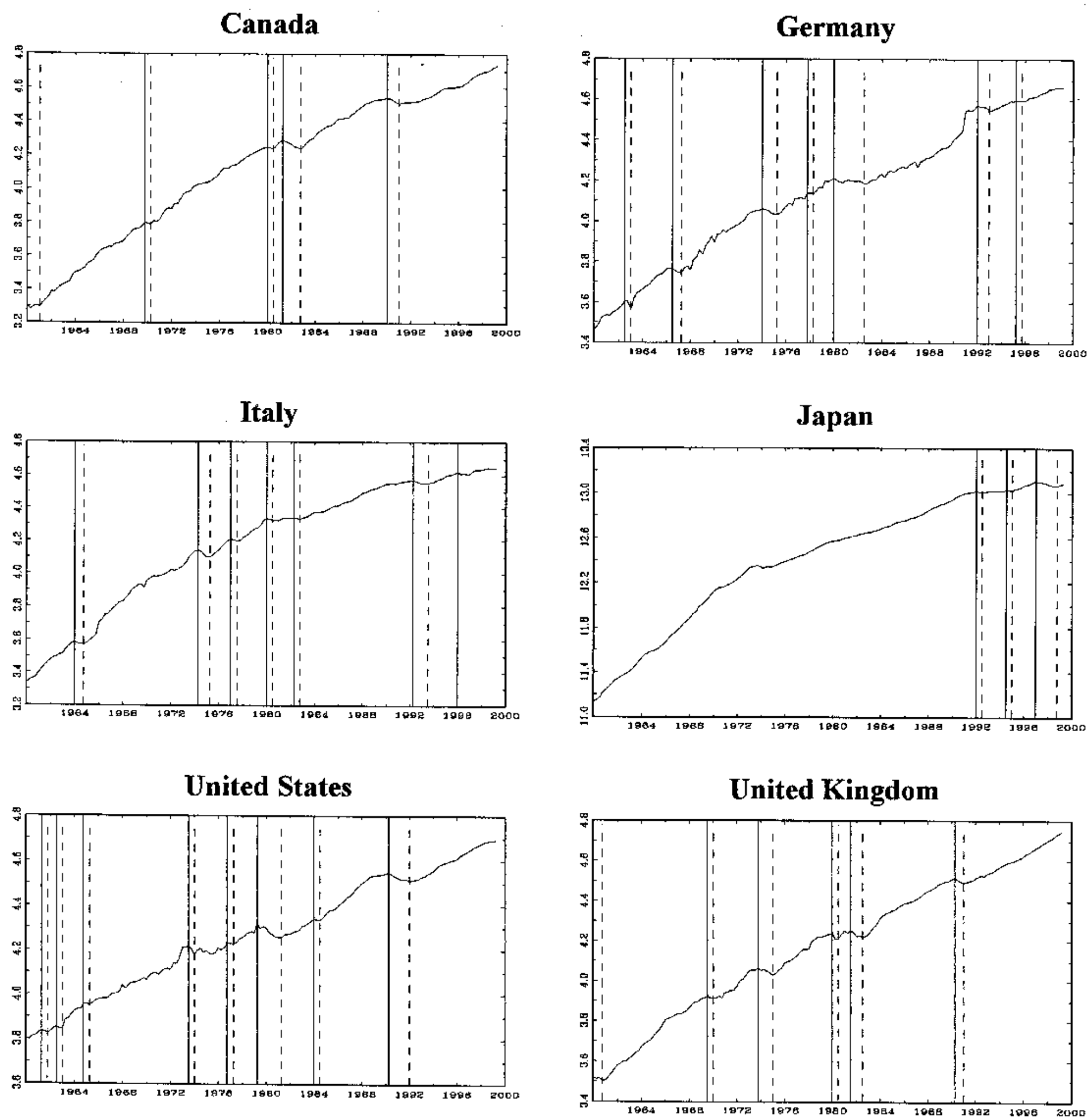
To illustrate the use of the concordance statistic, we apply the procedures discussed above to the logarithms of real output obtained from the IMF *International Financial Statistics* database for six selected major industrial countries: Canada, Germany, Italy, Japan, United Kingdom and the United States. The data for the six series are available from 1960:1 to 1999:2, affording us a sufficiently long data span to have a good impression of properties of the average classical cycle for each. The dating procedure as described in Section II is applied and the series plotted in Figure 8. As in Figures 6 and 7, solid vertical lines indicate peaks while dotted vertical lines indicate troughs.

Several features are immediately apparent from these plots. First, they confirm the stylized fact that classical output cycles are asymmetric: expansions are generally longer than contractions. Second, the timing of the recessions in the early 1980s and early 1990s is similar (but not identical) for all countries except Japan.[16]

[15] See also Cashin, McDermott and Scott (1999b) for an application of the concordance statistic in examining comovement in world commodity prices.

[16] The plot of the data for Japan is notable for the extremely long expansion dating from 1960:1 to 1992:1. Within this period there are two points at which growth was negative, one at 1974:1 and the other at 1974:4. However, because these periods of negative growth each last only one quarter, under the NBER dating rule they are not counted as a recession.

Figure 8. Datings for Peaks and Troughs, Selected Major Industrial Country Output Series, 1960:1-1999:2
(Logarithm of real GDP)



Sources: IMF, *IFS* and Authors' calculations.

Notes: Peaks are denoted by solid lines; troughs are denoted by dashed lines. Periods from peaks to troughs are contractions, while periods from troughs to peaks are expansions.

Once the dates of peaks and troughs have been determined, summary statistics can be calculated (see Table 3). We first report the average duration of expansions and contractions. For the six countries in our sample, the average contraction lasts around 3 to 4 quarters while the average expansion ranges from 11.5 quarters in the United Kingdom to 48 quarters in Japan. The average rise (decline) in an expansion (contraction) is shown by the amplitude measure. This measure indicates that in the major industrial countries contractions typically reduce real output by about 2 percent, while expansions increase real output by about 20 percent. However, a striking feature in these results is that the per quarter amplitudes are quite similar: around 1 percent per quarter in expansions, and ½ percent in contractions.

Table 3. Descriptive Statistics for Selected Major Industrial Country Output Series

	Contractions				Expansions			
	Duration	Amplitude	Amplitude per quarter	Brain-Shapiro test	Duration	Amplitude	Amplitude per quarter	Brain-Shapiro test
Canada	3.50	-2.50	-0.59	0.66	26.50	26.88	1.17	-1.10
Germany	4.00	-2.05	-0.61	1.92	17.50	16.60	0.95	0.95
Italy	3.00	-1.74	-0.56	1.18	18.33	16.01	0.89	1.02
Japan	3.67	-1.76	-0.39	0.77	48.00	31.67	0.65	0.77
United Kingdom.	3.44	-2.07	-0.57	2.66 1/	11.50	9.99	1.03	1.17
United States	3.20	-2.37	-0.75	0.50	21.00	19.32	0.94	-0.73

Notes: For each of the two phases (expansions and contractions), and for each of the six series, four results are presented. First, the average duration (in quarters) of the phase. Second, the average amplitude of the aggregate phase movement (in percentage changes). Third, the average amplitude per quarter (amplitude divided by duration). The fourth statistic (Brain-Shapiro test) is an examination of duration dependence. The null hypothesis of the Brain-Shapiro statistic is that the probability of exiting a phase is independent of the length of time a series has been in that phase. Using a 5 percent critical value for a two-tailed test, any result greater than 1.96 (in absolute value) indicates duration dependence in the output series. In the case of Japan, a subjective judgement was made to count the period from 1960:1 to 1992:2 as a complete trough-to-peak expansion.

1/ Significant at the 1 percent level

Following Diebold and Rudebusch (1990), the table also presents the Brain-Shapiro statistic for duration dependence, which tests whether the probability that an expansion or contraction will end changes as the expansion or contraction continues. The statistic has an asymptotically standard normal distribution and can be interpreted like a t-statistic for significant dependence in duration. The overwhelming result is that there is no duration dependence in expansions and significant duration dependence in contractions only for the United Kingdom. Hence, while the recent expansion phase in United States output is clearly exceptional—some 33 quarters against an average of 21 quarters—there is no evidence from past history that this will have increased the probability of switching to a contractionary phase. We take this as further evidence that the random walk with drift model is a good characterization of the data.

Table 4 presents concordance statistics. The i,j^{th} cell represents the concordance between the i^{th} and j^{th} countries. The numbers along the diagonal are therefore unity. The concordance numbers are very high for all 15 permutations—all of the series spend at least two thirds of the time in the same phase as each other. The highest concordance of 0.93 is recorded for Canada and the United States, and this is significant at the 1 percent level. This indicates that the two countries' business cycles are in phase most of the time. Given very strong trade links and their close proximity to each other, this is to be expected.

Table 4. Concordance Statistics for Selected Major Industrial Country Output Series

	Canada	Germany	Italy	Japan	United Kingdom	United States
Canada	1.00	0.79 1/	0.73	0.81	0.75 3/	0.93 1/
Germany		1.00	0.75 2/	0.78 2/	0.75 2/	0.82 1/
Italy			1.00	0.84 2/	0.68	0.75
Japan				1.00	0.73	0.80
United Kingdom					1.00	0.78 2/
United States						1.00

1/ Significant at the 1 percent level
2/ Significant at the 5 percent level
3/ Significant at the 10 percent level

It is perhaps more surprising, however, that for all of the combinations only three statistics are significant at the 1 percent significance level, with 8 statistics significant at the 5 percent level. This is because of trend growth in all of the series, which is documented, in the first column of Table 5. The estimated drift:standard error ratio ranges from 0.50 for Germany to 1.18 for Japan. This highlights the need to use hypothesis testing procedures rather than relying on the point estimate of concordance; using the measure of concordance without an understanding of its distributional properties may lead to misleading conclusions. For example, the concordance between Canada and Germany (0.79) is significant at the 1 percent level, while the concordance between Canada and Japan (0.81) is not significant even at the 10 percent level.

Table 5 also reports statistics for excess skewness and kurtosis. The null of no excess skewness is rejected for 3 out of 6 countries, while the null of no excess kurtosis is rejected for 4 out of 6 countries. Our test for concordance will therefore suffer from size distortion for many of the combinations in Table 4. Further, the range of μ/σ ratios in the data means that the size distortion will be relatively large. The excess kurtosis will lead us to reject the null of no concordance too many times. However, the data for Germany, Italy and the United Kingdom the data are left skewed, which works in the opposite direction; this may well offset the size distortion.

Table 5. Descriptive Statistics for Selected Major Industrial Country Output Series

	μ/σ	Skewness	Kurtosis
Canada	0.92	0.25	0.01
Germany	0.50	7.03 1/	19.76 1/
Italy	0.72	8.11 1/	22.13 1/
Japan	1.18	0.50	0.79
United Kingdom	0.54	2.48 1/	6.83 1/
United States	0.86	-1.47	2.83 1/

Notes: Three results are presented. The first column reports the ratio of the drift to the standard error from an estimated random walk with drift model, where Δx_t is $N(\mu,\sigma)$. The second and third columns report the Kiefer-Salmon statistics (as reported in Davidson and MacKinnon, 1993, pp.568-69) for skewness and kurtosis, respectively. The statistics are distributed as $N(0,1)$.

1/ Significant at the 1 percent level

Table 6. Correlation Statistics for Selected Major Industrial Country Output Series

	Canada	Germany	Italy	Japan	United Kingdom	United States
Canada	1.00	0.01	0.19 2/	0.22 1/	0.17 2/	0.51 1/
Germany		1.00	0.16 2/	0.21 1/	0.20 1/	0.01
Italy			1.00	0.36 1/	0.02	0.20 1/
Japan				1.00	0.21 1/	0.18 2/
United Kingdom					1.00	0.26 1/
United States						1.00

1/ Significant at the 1 percent level
2/ Significant at the 5 percent level

We contrast these findings with correlation statistics for the first differences of the output series, reported in Table 6. Significant correlation at the 1 percent level is recorded for 8 out of the 15 combinations and 12 out of the 15 at the 5 percent level. It is *not* the case, however, that the concordance test is simply "harder" than the correlation test. Quite different conclusions about comovement could be reached from the two measures. For example, an examination of the plots for German and United States output would probably lead most to conclude that the series generally move together. Indeed, as seen in Figure 9, the phase indicator for the two series shows that the concordance measure corroborates this impression, with 130 out of 158 periods spent in the same phase. As noted, the concordance statistic for these two countries is not merely large but significant at the 1 percent level, despite the obvious occurrence of trend growth in both series. In stark contrast, however, the correlation statistic is 0.01 and statistically indistinguishable from zero. This comes from a sample covariance statistic that is also close to zero. However, it would seem difficult to assert, on this basis, that the two series we see in Figure 9 are orthogonal to each other.

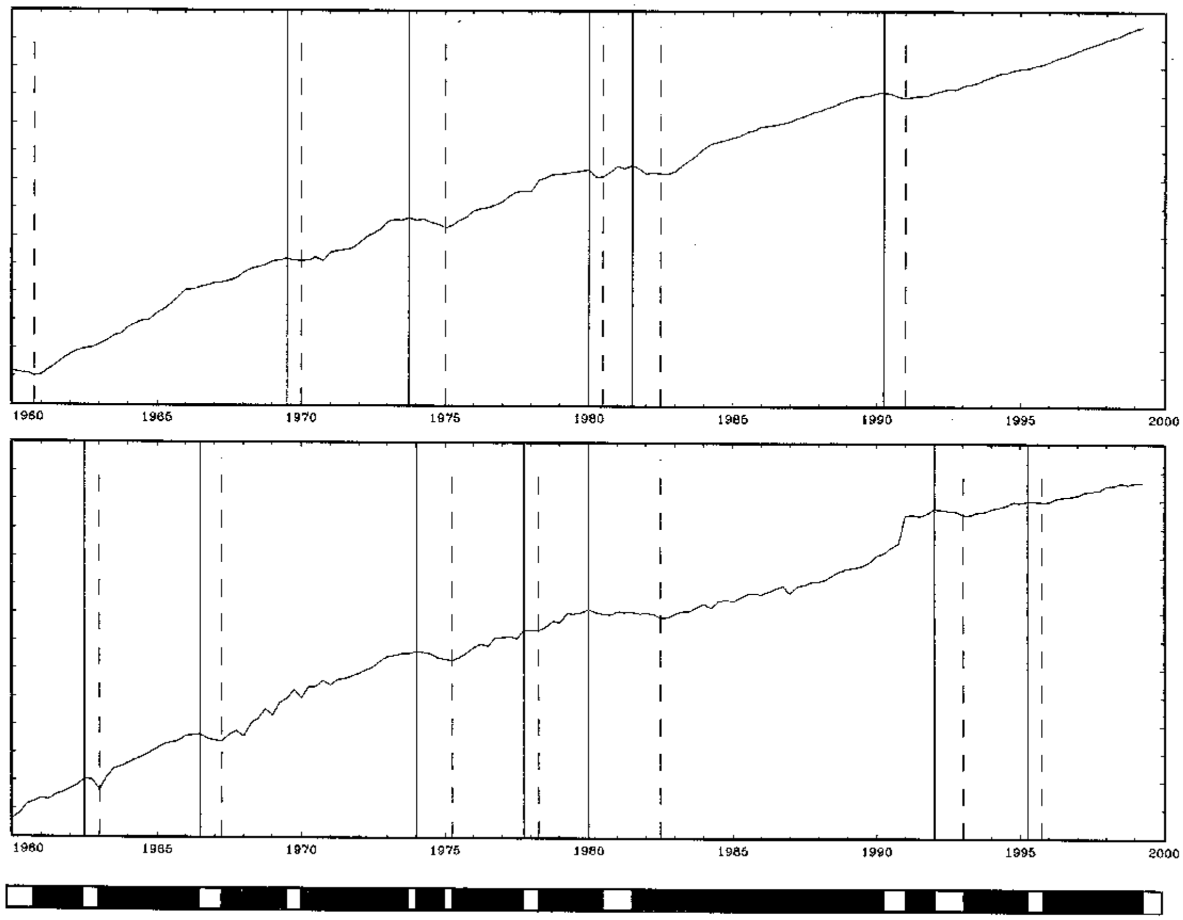
We hesitate to make comments on the evidence for or against the existence of an international business cycle from what is intended to be merely an illustrative exercise of the concordance statistic. While the concordance statistic is nonparametric, inference is unavoidably parametric. A more thorough analysis would therefore have to consider the impact of differing subsamples on inference. Nonnormality of the estimated residuals would clearly have to be addressed. Nonetheless, while significant concordance is found in the case where it would be most expected—the neighboring economies of Canada and the United States, for example—it is quite surprising how little evidence there is for widespread coincidence of expansions and contractions across these major industrial countries.

VI. Conclusions

In this paper we show that the concordance test is a useful means to gauge whether two series comove; that is, to what extent they are typically together in an expansionary or contractionary phase. We have also shown how to infer the significance of the statistic against the null hypothesis that the concordance is the result of pure chance. For this purpose, we utilized Monte Carlo simulations and response surfaces. These regressions are an effective means of summarizing results from such simulations, because they smooth and interpolate critical values for all sample sizes.

The properties of the concordance test were examined in two ways: first, by using simulations to calculate size distortions when the innovations are nonnormal, and second, by using simulations to calculate power as a function of the correlation between the innovations. The simulations show the test is somewhat sensitive to nonnormal innovations, so that it is important to conduct diagnostic checks on estimated innovations. The test has reasonable power properties when the correlation between the innovations is high and improves substantially as the sample size grows.

Figure 9. Datings of Peaks and Troughs for Germany (Top Panel), United States (Middle Panel), and Phase Indicator (Bottom Panel), 1960:1-1999:2
(Logarithm of real GDP)



Sources: IMF, *IFS* and Authors' calculations.

Notes: Peaks are denoted by solid lines; troughs are denoted by dashed lines. Periods from peaks to troughs are contractions, while periods from troughs to peaks are expansions. The "bar code" at the bottom of the Figure shows when the two series are in the same phase according to the dating—that is, when the two series are in contraction or expansion at the same time, the bar code is solid and blank when out of phase.

As an empirical illustration, the Bry-Boschan algorithm and the concordance statistic were applied to output data from selected major industrial countries. The algorithm and concordance tests are easy to apply systematically to large data sets. In the case of the small number of output data series considered here, it proved a good way to compare business cycles across countries. Overall, we find the concordance statistic to have the appealing virtue that it means what it says—the statistic relates well to our intuitive impressions about series "moving together", and can be literally read as a measure of the extent of comovement. We therefore hope that it will form a useful addition to the applied economist's toolkit.

References

Bry, Gerhard, and Charlotte Boschan, 1971, *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, (New York: National Bureau of Economic Research).

Burns, Arthur F. and Wesley C. Mitchell, 1946, *Measuring Business Cycles*, NBER Studies in Business Cycles No.2, (New York: National Bureau of Economic Research).

Cashin, Paul, C. John McDermott, and Alasdair Scott (1999a), "Booms and Slumps in World Commodity Prices," IMF Working Paper 99/155, (Washington: International Monetary Fund).

_____, (1999b), "The Myth of Comoving Commodity Prices," IMF Working Paper 99/169, (Washington: International Monetary Fund).

Davidson, Russell, and James G. MacKinnon, 1993, *Estimation and Inference in Economics*, New York: Oxford University Press.

De Jong, David N., John C. Nankervis, N.E. Savin, and Charles H. Whiteman, 1992, "The Power Problems of Unit Root Tests in Time Series with Autoregressive Errors," *Journal of Econometrics*, Vol. 53, pp. 323-43.

Diebold, Francis X. and Glenn D. Rudebusch, 1990, "A Nonparametric Investigation of Duration Dependence in the American Business Cycle," *Journal of Political Economy*, Vol. 98 No. 3, pp. 596-616.

Harding, Don and Adrian Pagan (1999), "Knowing the Cycle," mimeo, (Melbourne, Australia: University of Melbourne).

King, Robert G. and Charles I. Plosser, 1994, "Real Business Cycles and the Test of the Adelmans," *Journal of Monetary Economics*, Vol. 33 No. 2, pp. 405-38.

Nelson, Charles R. and Charles I. Plosser, 1982, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, Vol.10, pp.139-62.

Pagan, Adrian, 1997a, "Towards an Understanding of Some Business Cycle Characteristics," *The Australian Economic Review*, Vol. 30 No. 1, pp.1-15.

_____, 1997b, "Policy, Theory and the Cycle," *Oxford Review of Economic Policy*, Vol. 13 No. 3, pp.19-32.

_____, 1999, "A Framework for Understanding Bull and Bear Markets," mimeo, (Canberra, Australia: Australian National University).

Perron, Pierre, 1989, "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, Vol. 57, pp.1361-1401.

Watson, Mark, 1994, "Business Cycle Durations and Postwar Stabilization of the U.S. Economy," *The American Economic Review*, Vol. 84 No. 1, pp. 24-46.

Zivot, Eric and Donald Andrews, 1992, "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Journal of Business and Economic Statistics*, Vol. 10, pp. 25