

Variable Trends in Economic Time Series

James H. Stock and Mark W. Watson

The two most striking historical features of aggregate output are its sustained long run growth and its recurrent fluctuations around this growth path. Real per capita GNP, consumption and investment in the United States during the postwar era are plotted in Figure 1. Both growth and deviations from the growth trend—often referred to as “business cycles”—are apparent in each series. Over horizons of a few years, these shorter cyclical swings can be pronounced; for example, the 1953, 1957 and 1974 recessions are evident as substantial temporary declines in aggregate activity. These cyclical fluctuations are, however, dwarfed in magnitude by the secular expansion of output. But just as there are cyclical swings in output, so too are there variations in the growth trend: growth in GNP in the 1960s was much stronger than it was in the 1950s. Thus, changes in long run patterns of growth are an important feature of postwar aggregate economic activity.

In this article we discuss the implications of changing trends in macroeconomic data from two perspectives. The first perspective is that of a macroeconomist reassessing the conventional dichotomy between growth and stabilization policies. As an empirical matter, does this dichotomy make sense for the postwar United States? What is the relative “importance” of changes in the trend and cyclical swings in explaining the quarterly movements in economic aggregates? We next adopt the perspective of an econometrician interpreting empirical evidence based on data that contain variable trends. The presence of variable trends in time series data can lead one to draw mistaken inferences using conventional econometric techniques. How can these techniques—or our interpretation of them—be modified to avoid these mistakes?

■ *James H. Stock is Assistant Professor of Public Policy, Kennedy School of Government, Harvard University, Cambridge, MA. Mark W. Watson is Associate Professor of Economics, Department of Economics, Northwestern University, Evanston, IL. This article was written while Stock was a National Fellow at the Hoover Institution, Stanford University, Stanford, CA.*

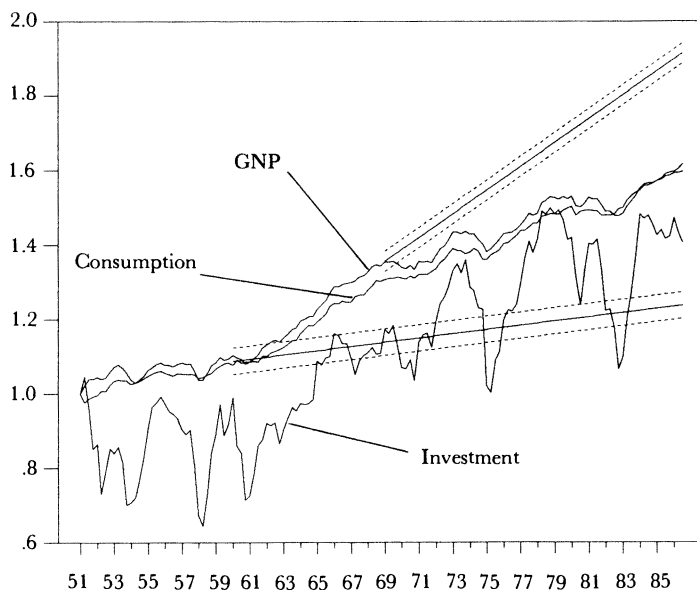


Fig. 1. Postwar real per capita U.S. GNP, total consumption, and gross private domestic investment (in logarithms)

All three series have been arbitrarily set to 1 in 1951:I. The straight solid lines represent two long-run forecasts of GNP, the lower using data from 1951:I–1959:IV and the upper using data from 1960:I–1969:IV. The forecasts were made by extrapolating GNP growth over these periods using a linear deterministic time trend. The dotted lines represent bands of \pm two standard deviations of quarterly GNP growth around the long-run forecasts. Were GNP a stationary series about a linear time trend, these bands would provide an approximate long-run 95% confidence interval for the respective GNP forecasts.

The macroeconomist's perspective is adopted in the first major section of the article. Discussions of macroeconomics often treat the concepts of “trends” and “cycles” in output separately. On the one hand, theories of growth examine the forces capable of changing long run trends, while on the other hand theories of the business cycle attempt to explain shorter run fluctuations and to determine when macroeconomic policy might stabilize or exacerbate the swings between expansion and recession. At one level this dichotomy seems natural, with different theories providing insights into macroeconomic movements over different horizons. But on another level this distinction is artificial: theories explaining only growth or only cycles cannot provide adequate macroeconomic insights if there are important interactions between the two.

A closer look at Figure 1 suggests that changes in growth trends are associated with some of the shorter, “cyclical” swings in the series. For example, a key turning point between the high-growth 1960s and the low-growth 1970s and 1980s seems to have been in the early 1970s, which was also followed by a major recession that saw a particularly sharp transitory drop in investment. When one formally defines trends using the “stochastic trends” concept discussed below, various statistical measures (introduced in the next section) indicate that a substantial fraction of the quarterly

variation in real GNP is associated with movements in long run trends rather than being purely transitory fluctuations. This agrees with the casual inference drawn from Figure 1 that shifts in trends are an important part of changes in GNP. Moreover, examinations of the co-movements of several macroeconomic variables indicate that unforecasted shifts in long run economic prospects are associated with short run fluctuations that are similar to common conceptions of the business cycle. These findings suggest two conclusions: first, that a key step in understanding the co-movements in aggregate economic variables is learning about the link between changes in long run economic trends and cyclical swings; and second, that a sharp dichotomy between growth and stabilization policies misses an important connection between the two policy goals.

The econometrician's perspective on variable trends is taken in the second major section of this article. Variable trends provide numerous econometric pitfalls and raise difficult methodological issues. Time series analysts have long recognized that regression analysis can be highly misleading when applied to series with variable trends. In some cases (as in Figure 1), the result can be dramatic errors in forecasting. In other cases, an improper treatment of variable trends can result in false conclusions about how the economy works. To illustrate this point, we construct a simple artificial economy and assign two hypothetical inhabitants the task of discovering its true structure. As in many model economies studied in the modern literature on macroeconomic theory, these inhabitants happen to be econometricians; but here their training often fails them simply because they mishandle the variable trends in their data. Finally, we draw on recent developments in econometric theory to provide some simple rules-of-thumb that a consumer or producer of econometric models can use in an attempt to avoid these pitfalls.

What is a "Variable Trend?"

The basic premise of econometric analysis is that, when viewed together, individual cases and experiences can provide insights into a deeper unifying structure. When analyzing cross sectional data, the individual experiences might be those of different workers or firms; but the only individual experiences we have with the operation of the U.S. economy are historical. For macroeconometric analysis to be of any value, then, it must be that the historical experiences comprising an economic time series are on the one hand sufficiently different from each other that more experiences provide additional information, but on the other hand sufficiently similar that combining individual experiences can elucidate the underlying economic structure. These two requirements are the essence of the technical assumption that a time series is ergodic and stationary.¹

The first requirement—that historical experiences be sufficiently unrelated—is unlikely to hold for GNP, consumption and investment in Figure 1. The key reason is

¹Formally, a time series random variable is said to be stationary if its distribution does not depend on time. As is the convention, we further assume that *stationary* variables have finite variances and autocovariances. See Andrew Harvey (1981, p. 22) or Clive Granger and Paul Newbold (1977, p. 4) for more details.

that the current trend levels of these variables arguably depend on the entire history of these series. For example, were the experiences of the 1950s essentially unrelated to those of the 1980s, except for a perfectly predictable trend, an econometric forecaster in 1960 might predict 1987 GNP by extrapolating trend growth during the 1950s out to 1987; an econometrician in 1970 might do the same, only using data from the 1960s. With unchanging trends, these forecasts should be similar, and indeed should be moderately accurate. But, as indicated in Figure 1, these forecasts are close neither to each other nor to the actual value of 1987 GNP. Just as important, the confidence bands around the trend extrapolations in Figure 1—designed to capture the uncertainty associated with these forecasts—clearly fail to provide a reliable estimate of the range of future GNP. Thus historical experiences—namely, unforecasted increases in the level of GNP in some past year—appear to have had a dramatic and persistent influence on production today.

On a casual level, the notion that GNP is composed of a variable trend plus some additional cyclical movements seems quite satisfactory. Time series variables can, without further restrictions, be thought of as composed of a part with a variable trend, plus a part that is not the trend. But a moment's reflection indicates that this decomposition lacks content; without a more precise definition, one economist's "trend" can be another's "cycle." The ambiguity surrounding the definition of a variable trend applies equally to the definition of a "business cycle." Indeed, one can imagine extreme views that there is no business cycle, in the sense that all economic fluctuations are merely movements in variable trends—or the reverse view that what appear to be variable macroeconomic trends are but very long cyclical swings. But these extreme views miss valuable notions traditionally associated with the business cycle: that the economy experiences protracted but nonetheless temporary periods of unusually high (or low) output, employment, and inflation across many or all sectors of the economy; that these protracted experiences have important implications for the well-being of people and institutions; and that these changes in well-being often have important political consequences. In referring to business cycles, then, we mean these protracted yet temporary swings in aggregate output.

Then what is a trend? Perhaps the theory of economic growth can provide an operational definition. For example, in a one-sector neoclassical growth model, trends can arise because of technical progress and an increasing labor force and capital stock. If the production technology has constant returns to scale, this suggests that—at least in theory—the trend in output per capita could be represented as a function of the capital-labor ratio, the labor force participation rate, and the stock of "technology." Unfortunately, in practice things are not so simple; for example, the capital-labor ratio and the labor force participation rate have cyclical as well as trend components, and the stock of technology cannot be measured directly. Thus, at least at this level of analysis, this approach based on growth theory has little empirical relevance.

It is therefore common to take a different approach to the definition of a variable trend. The specific notion we adopt is a direct extension of a *deterministic linear time trend* (used to compute the forecasts in Figure 1), which increases by some fixed amount (say, 1 percent) every quarter. In contrast, we model a variable trend as increasing in

each quarter by some fixed amount (say, 1 percent) on average; however, in any given quarter the change in the trend will deviate from its average by some unforecastable random amount. Because it has this unpredictable random component, henceforth we refer to this formulation of a variable trend as a *stochastic trend*. The reader familiar with the random walk theory of stock prices will recognize that this notion of a trend corresponds to a random walk with drift. From a forecasting perspective a key feature of a random walk—in contrast to a stationary time series variable—is that, because of its random growth, forecasts of its level will entail uncertainty that increases as the forecast horizon recedes. With this formulation of a stochastic trend, a (random) change in the trend in one quarter provides a new base from which growth will occur in the next.²

While this definition of a stochastic trend might at first seem restrictive, a recent theorem by Stephen Beveridge and Charles Nelson (1981) suggests that it might be broadly applicable to U.S. data. Since George Box and Gwilym Jenkins (1970) proposed their influential autoregressive-integrated-moving average (ARIMA) models for forecasting time series variables, econometricians have generally recognized that many macroeconomic series (when considered one at a time) appear to be *integrated*, that is, their first differences are stationary. ARIMA models are sophisticated yet simple tools for forecasting a single series using only its history. Box and Jenkins referred to integrated variables with autoregressive (AR) or moving average (MA) components as ARIMA($p, 1, q$) processes, where p and q , respectively, denote the orders of the autoregressive and moving average terms and where “1” indicates that the variable is integrated of order one, which means that its first difference (quarterly change) is stationary. ARIMA models are extensions of conventional regression models to time series variables; the AR terms forecast the current variable using p of its lags, and the MA terms forecast using q lags of the error in the process.³

²A time series variable x_t is a *random walk with drift* if x_t evolves according to

$$x_t = \mu + x_{t-1} + e_t \text{ or } x_t - x_{t-1} = \mu + e_t$$

where e_t has mean zero and variance σ_e^2 , and where e_t is serially uncorrelated, i.e. e_t is uncorrelated with e_s for $s \neq t$, so e_t cannot be forecasted using past values of x_t . The “drift” in the random walk is μ , the average predictable increase in x_t in each period. (For example, from 1951:I to 1986:III, real GNP increased by 3.0 percent annually; on a per capita basis, the average annual increase was 1.7 percent). With this definition, a variable y_t that contains a stochastic trend can be written as $y_t = y_t^p + y_t^s$, where y_t^p is a random walk (possibly with drift) and y_t^s is a stationary time series variable; the superscripts “ p ” and “ s ” refer to the “permanent” (or trend) and “stationary” (or transitory) components of y_t . In contrast, a variable that contains a deterministic time trend can be written as $y_t = gt + y_t^s$, where y_t^s is stationary and g is the constant quarterly growth of the deterministic trend.

³*Some Useful Definitions*

(i) The variable x_t is *integrated of order one* (or simply *integrated*) if it is nonstationary and can be written,

$$x_t = \mu + x_{t-1} + u_t \text{ or } x_t - x_{t-1} = \mu + u_t$$

where u_t has mean zero and variance σ_u^2 , and where u_t is stationary. It is convenient to let Δx_t represent $x_t - x_{t-1}$. A variable is said to be *integrated of order d* if it must be differenced d times to be stationary; for example, if Δx_t is not stationary but $\Delta x_t - \Delta x_{t-1}$ is, then x_t is integrated of order 2. The distinction between an integrated process and a random walk is that, if x_t is integrated, u_t is stationary but might be correlated with lagged u_t ; if x_t is a random walk, u_t is serially uncorrelated.

Among time series analysts, a major attraction of ARIMA($p, 1, q$) models is their ability to forecast many macroeconomic variables with an accuracy that is impressive among univariate forecasting techniques. Here, however, the importance of ARIMA models is that Beveridge and Nelson prove that every variable having an ARIMA($p, 1, q$) representation contains a random walk stochastic trend. Since ARIMA($p, 1, q$) models seem to characterize many macroeconomic variables, it follows that the growth in these variables can be described by stochastic trends. Beveridge and Nelson's (1981) trend/cycle decomposition is presented in the Appendix for readers familiar with the mathematical particulars of ARIMA models.

A possible objection to the discussion so far is its emphasis on variables that are integrated of order one, so that the quarterly percentage growth in GNP is modeled as being stationary. Why not treat U.S. data as integrated of order two, so that the second difference of GNP is stationary but the quarterly growth rate itself contains a stochastic trend? Or why not model GNP as stationary but having coefficients that imply almost the same "stochastic trend" behavior as an integrated process? Our answer is that the preponderance of evidence currently suggests that the integrated model provides the best approximation of U.S. GNP. This is not to say that U.S. GNP is an integrated process, for this can never be learned with certainty by examining a finite time series; nor is it to say that future research using new techniques or more data could not change this assessment. But, given currently available statistical techniques, modeling GNP as an integrated process seems to provide a good approximation to its long run properties.⁴

(ii) Following Box and Jenkins, an ARIMA($p, 1, q$) model specifies x_t as being integrated of order one and as having a representation of the form,

$$\Delta x_t = c + a_1 \Delta x_{t-1} + \dots + a_p \Delta x_{t-p} + e_t + b_1 e_{t-1} + \dots + b_{q-1} e_{t-q}$$

where $a_1, \dots, a_p, b_1, \dots, b_q$, and c are constant parameters and where e_t is serially uncorrelated.

(iii) A variable that is integrated is said to have a *unit root in its autoregressive representation*. The term "unit root" refers to the unit coefficient on x_{t-1} in the formula defining an integrated process. The statements " x_t has a unit root" and " x_t is integrated of order one" are equivalent. Box and Jenkins use the term "nonstationary" to refer to an integrated process. This terminology, while conventional, is unfortunate: while all integrated processes are nonstationary, not all nonstationary processes are integrated.

(iv) *An integrated process has a variance that tends to infinity*. This is most easily demonstrated for a random walk. Let $x_t = x_{t-1} + e_t$ where e_t is serially uncorrelated and $\text{var}(e_t) = \sigma_e^2$, and let $x_0 = 0$. Then $x_t = \sum_{s=1}^t e_s$, so that $\text{var}(x_t) = t\sigma_e^2$, which tends to infinity with t .

⁴Guy Orcutt (1948) was among the first to suggest that GNP is an integrated process. Using annual U.S. data from 1919 to 1932, he found that many aggregate time series were well described by an ARIMA(1, 1, 0) model with an autoregressive coefficient of .3. This argument for the existence of stochastic trends in macroeconomic data is only partly convincing, since Orcutt's and Box and Jenkins' techniques for determining whether the process is integrated require the practitioner to make qualitative judgements. To overcome this drawback, Wayne Fuller (1976) and David Dickey and Fuller (1979) proposed several statistical tests for whether a variable is integrated, against the alternative that it is not (i.e. it is stationary). Nelson and Charles Plosser (1982) provided firm support for what Orcutt and the Box-Jenkins practitioners had suspected: upon applying the formal Dickey-Fuller tests to fourteen annual macroeconomic variables using 60 and 100 years of data, they could not reject the hypothesis that there is a stochastic trend in real and nominal output measures, wages, prices, monetary variables, and asset prices. In their analysis, only the unemployment rate failed to contain a stochastic trend over the twentieth century. Nelson and Plosser also emphasized that the presence of stochastic trends calls into question the validity of the traditional trend/cycle dichotomy used to describe aggregate time series.

The general success of ARIMA modeling therefore provides a technical motivation for the stochastic trend formulation. But the real motivation is provided by Figure 1, where the long-run forecasts based on a deterministic linear trend are simple, intuitively appealing, and wrong. With a deterministic linear trend, the uncertainty associated with a long run forecast is limited to the variation in the stationary deviations from that trend. In contrast, with a stochastic trend, the greater the forecasting horizon, the greater is the uncertainty associated with that forecast.

Relations Between Trends and Cycles in Macroeconomic Variables

Implicit in many models of the business cycle is the notion that macroeconomic fluctuations are, for the most part, caused by temporary rigidities or misperceptions. These models abstract from economic growth, operating on the implicit assumption that the growth process has little impact on the business cycle. In contrast, an alternative view (exposed clearly by Edward Prescott, 1987) explains economic fluctuations entirely as a reaction to changes in the long run growth prospects of the economy, with “business cycles” arising simply as adjustments to new long-run growth paths.

But theoretical debates linking changes in trends to cyclical fluctuations are of little practical interest unless empirical evidence suggests that there is such a link. In this section, we consider two related questions concerning the quantitative importance of changes in long run prospects for short run economic fluctuations. First, to what extent are quarterly movements in postwar U.S. real per capita GNP associated with variations in its trend? Second, are variations in macroeconomic trends linked to cyclical movements, or are the trend and cyclical variations largely unrelated?

Measuring the Trend in GNP

The first question seems simple enough, especially if we hold ourselves to the “random walk” concept of stochastic trends. It has, however, generated a heated debate over the importance of the trend component in GNP. Besides being of interest in its own right, this debate highlights central conceptual difficulties that arise in defining trends.

The starting point for this analysis is to suppose that real GNP consists of two parts: a stochastic trend, plus a part that is transitory, or more precisely, stationary. How important, then, are these two components? Answering this question requires first determining whether the unforecastable changes (technically, the innovations) in the two components are correlated, and second, developing an econometric framework (that is, a model) for interpreting the information contained in historical GNP data.

One would hope that an examination of historical data might shed some light on the correlation between the permanent and stationary innovations. Unfortunately this hope is vain, since this correlation cannot be estimated directly from a single time series; that is, this correlation is not identified in the usual econometric sense. This lack of identification should not be surprising. Broadly speaking, determining the correla-

tion between the trend and stationary innovations in GNP is much like deciding whether the 1975 downturn was a result of a permanent shift in the trend, a transitory fluctuation, or some combination of the two—using only the plot of GNP in Figure 1. Without embarking on a review of the literature, we follow previous authors and consider two extremes: that these innovations are either uncorrelated or perfectly correlated. In the context of the trend-stationary decomposition, assuming the innovations to be uncorrelated implies that the changes in the trend and the transitory fluctuations are unrelated, except of course that they both affect GNP. In contrast, assuming the innovations to be perfectly correlated implies that they arise from the same source.

Although the decision concerning the correlation between the trend and stationary innovations is conceptually distinct from the decision about which model to estimate, in practice the first decision suggests a class of models to choose from. In the introduction, we emphasized that Beveridge and Nelson's theorem shows that all ARIMA($p, 1, q$) models imply the presence of a stochastic trend. In fact, their theorem does more than this: it provides an explicit formula for computing a stochastic trend implied by an ARIMA model. An ARIMA model reduces all unforeseen economic events into a single innovation, and the Beveridge-Nelson trend and stationary components are both based on this innovation.

Although they sound complicated, the ARIMA framework and Beveridge-Nelson decomposition are in fact simple. As a concrete example, consider an ARIMA(0, 1, 1) model fit to GNP. Let y_t denote the logarithm of real U.S. GNP, let $\Delta y_t = y_t - y_{t-1}$ denote quarterly real GNP growth, let e_t denote an unobserved error term, and let SE denote the standard error of the estimate of Δy_t , that is, the standard deviation of the in-sample one-step-ahead forecast error. Because y_t is assumed to be an integrated process, the model is specified using the stationary growth rates, Δy_t . Using data from 1947:II to 1985:IV, we estimated the ARIMA(0, 1, 1) model,⁵

$$(1) \quad \Delta y_t = .008 + e_t + .3e_{t-1}, SE = .0106.$$

In this ARIMA(0, 1, 1) specification, GNP growth is expressed as a weighted moving average of the errors, e_t . According to (1), given e_{t-1} but not e_t , GNP growth would thus be forecasted by $.008 + .3e_{t-1}$. In addition, the Beveridge-Nelson decomposition is particularly easy to apply in this case: using (1) and (A.2) in the Appendix, the permanent, or trend, component in log output (call this y_t^p) is $y_t^p = 1.3\sum_{s=1}^t e_s$. The transitory, or stationary, component in log output (call this y_t^s), also given by (1) and (A.2), is $y_t^s = -.3e_t$. Evidently the innovations in the Beveridge-Nelson permanent and stationary components are both e_t , so that the permanent and stationary innova-

⁵Estimation was performed using the econometrics software package RATS. Let RGNP be the name used for the series, real GNP, in a RATS session. The ARIMA(0, 1, 1) model was estimated using the RATS commands:

```
set dlrgnp 47:2 85:4 = log(rgnp(t)) - log(rgnp(t - 1))
boxjenk(ar = 0, ma = 1, constant) dlrgnp 47:2 85:4
```


tions are perfectly correlated. Although this example is an ARIMA(0, 1, 1) model, their decomposition applies more generally to all ARIMA($p, 1, q$) models. In the “perfect-correlation” case then, the choice of model reduces to a choice of p and q .

The choice of a specific model to estimate in the “zero-correlation” case is conceptually similar. In this class of models, GNP is explicitly represented as the sum of its permanent component, modeled as a random walk with drift, and its transitory component, modeled as a stationary ARMA process (i.e. as an ARIMA($p, 0, q$) process). Since neither component is observed directly (only their sum, GNP, is), this model is called an unobserved components ARIMA (UC-ARIMA) model. As a specific example, Watson fit a simple UC-ARIMA model to log real GNP, where the stationary component was assumed to be a second order autoregression (i.e. to be ARMA(2, 0)). Let e_t^s and e_t^p respectively be the innovations in the stationary and permanent components. His model, estimated using data from 1949:I–1984:IV, is:⁶

$$(2) \quad y_t = y_t^p + y_t^s,$$

$$y_t^p = .008 + y_{t-1}^p + e_t^p, \text{ std. dev.}(e_t^p) = .0057$$

$$y_t^s = 1.5y_{t-1}^s - .6y_{t-2}^s + e_t^s, \text{ std. dev.}(e_t^s) = .0076$$

$$\text{cov}(e_t^s, e_t^p) = 0, SE = .0099.$$

The innovations e_t^s and e_t^p are uncorrelated by assumption. Summarizing, in (2) the permanent component is written as a random walk with a drift of .008, and the stationary term is predicted using two of its lags.

Given a choice of model, the question becomes how to measure the extent to which quarterly movements in GNP are associated with variations in its trend. Since there is no one answer, various measures addressing different aspects of this question are presented in Table 1. The results pertain to different models based on the two assumptions about the correlation between the trend and stationary innovations. The results in the first row are for the ARIMA(0, 1, 1) model reported in (1). The next two results are for other low-order ARIMA models of GNP of the type estimated by John Campbell and Gregory Mankiw (1987a) and Watson (1986), while the following two are for higher order autoregressive models of the type studied by John Cochrane (1986). The final model is a UC-ARIMA model estimated by Peter Clark (1987a) with a stochastic trend plus a stationary AR(2) component, in which the innovations in the trend and stationary components are uncorrelated by assumption.⁷

⁶ Estimation of the parameters of UC-ARIMA models is more involved than for ARIMA models. Harvey (1985) provides a complete discussion of specification and estimation of UC-ARIMA models.

⁷ Model 3 is taken from Campbell and Mankiw (1987a) and Model 6 is taken from Clark (1987a). Cochrane used annual data for his empirical work; the high-order autoregressive models reported here capture his notion of including many lags in an annual specification. Models 1 and 2 were also reestimated using the full data set. All models produce similar short run forecasts and fit the data well using standard time series diagnostic measures.

Table 1

**Measures of the importance of the trend in real log GNP,
estimated using data from 1947:I to 1985:IV**

Univariate Statistical Model	Long-run increase in GNP predicted from a 1% unforeseen increase in GNP in one quarter	Variance ratios	
		$R^2_{\text{corr}=1}$	$R^2_{\text{corr}=\text{min}}$
1. ARIMA(0, 1, 1)	1.3	.93	.73
2. ARIMA(1, 1, 0)	1.6	.86	.60
3. ARIMA(2, 1, 2)	1.5	.80	.70
4. ARIMA(12, 1, 0)	1.3	.75	.57
5. ARIMA(24, 1, 0)	.9	.79	.64
6. UC-ARIMA	.6	.84	.34

The univariate estimates are based on analysis of GNP. As discussed in the text, the UC-ARIMA model was estimated under the assumption that the trend and stationary innovations are uncorrelated. The R^2 statistics measure the fraction of the variance in the quarterly change in real GNP attributable to changes in its stochastic trend. The minimal correlations used to compute the $R^2_{\text{corr}=\text{min}}$ statistic for the univariate models are (in order): .85, .9, .95, .9, .8, 0.

Source: authors' calculations, drawing on Campbell and Mankiw (1987a); Clark (1987a); Cochrane; and Gagnon.

The first measure of the importance of the trend component in GNP addresses the following question (enunciated most clearly by Campbell and Mankiw, 1987a): supposing that in some quarter GNP were to increase by 1 percent above its forecasted amount, how would that change one's forecast of the long run level of GNP? If GNP were a stochastic trend with no stationary component, then the answer would be 1 percent: the best forecast of a random walk arbitrarily far in the future is its current value, and if that value were to change by 1 percent, so would the long run forecast. In contrast, if GNP were stationary around a purely deterministic time trend, then the answer would be zero percent: any unpredictable change would have only a transitory effect on forecasts of future GNP.

It turns out that the choice of model by the various researchers makes a big difference in estimating the long run effect of an innovation in the trend; the estimates vary by a factor of two. Although this is unsatisfying, it is useful to understand why these estimates vary. Consider first the estimates based on the assumption of a perfect correlation between the trend and stationary innovations. Using the results in the first row of Table 1, suppose that GNP were to grow by an unforecasted 1 percent in some quarter. This growth will, on average, arise partly from an innovation in the trend and partly from innovations in the stationary component. Since the stationary innovation is perfectly correlated with the trend innovation, the change in the stationary component will either augment or partially offset the permanent innovation, respectively depending on whether the trend innovation induces an increase or a decrease in the stationary component. Using the ARIMA(0, 1, 1) model (1), since $y_t^p = 1.3\sum_{s=1}^t e_s$ and $y_t^s = -.3e_t$, an increase in the trend is associated with a decrease in the stationary component using the Beveridge-Nelson decomposition: for a 1.3

percent increase in the trend, the stationary component would initially drop by .3 percent, leaving a net unforecasted increase in GNP of 1 percent. Eventually the effect of this innovation on the stationary component will vanish, leaving only the 1.3 percent permanent increase.

In contrast, when an unforecasted increase in GNP consists of the two uncorrelated innovations, less of an unforecasted quarterly change will be attributed to shifts in the trend. A 1 percent increase in GNP might have come from a trend or a stationary innovation; since these two innovations on average neither reinforce nor cancel each other, the best guess of either innovation will be less than one. Equivalently, with uncorrelated innovations, the fraction of an unforecasted change that is likely to be permanent will be less than one. Under the uncorrelated assumption, a 1 percent increase in GNP today will on average therefore lead to a long run increase of less than 1 percent. In general, then, the initial assumption about this correlation is crucial to ascertaining the relative importance of the two components.

The second pair of measures in Table 1 answers a related question about the relative importance of the trend and stationary components: what fraction of the quarterly variation in GNP is attributable to permanent shifts? Unfortunately, the answer to this is ambiguous, even given an estimated model for GNP, since the fraction of the variance accounted for by the trend depends on the correlation between the trend and stationary innovations. The two measures in the table are “R-squared” statistics that would arise from regressing the quarterly change in GNP against the change in its true trend which in general is unobserved where the trend is computed under different assumptions about this correlation. The first statistic, $R^2_{\text{corr}=1}$, is from the regression using the trend computed assuming the trend and stationary innovations to be perfectly correlated, so that this trend is computed using Beveridge and Nelson’s formula. The second measure is based on the observation that a given ARIMA model cannot be distinguished from (formally, is observationally equivalent to) a UC-ARIMA model with an assumed correlation falling in a certain range, where this range depends on the parameters of the ARIMA model. Accordingly, $R^2_{\text{corr}=\text{min}}$ is computed using the smallest correlation that is capable of generating the estimated ARIMA model from a UC-ARIMA model. Thus these R^2 measures provide a rough range of the fraction of movements in quarterly GNP attributable to shifts in the trend.⁸

⁸The observational equivalence of ARIMA and UC-ARIMA models for a range of correlations is not obvious, although as an illustration it is readily derived for the UC-ARIMA model, $y_t = y_t^p + e_t^s$, where $y_t^p = y_{t-1}^p + e_t^p$, $\text{cov}(e_t^s, e_t^p) = 0$, and where e_t^p and e_t^s are serially uncorrelated. This model is observationally equivalent to the ARIMA(0, 1, 1) model: $\Delta y_t = e_t + b e_{t-1}$, where $b/(1 + b^2) = -\text{var}(e_t^s)/[\text{var}(e_t^p) + 2\text{var}(e_t^s)]$. The observational equivalence follows because both models imply the same autocovariances for Δy_t .

Our variance measures are inspired by Masanao Aoki’s (1987) measure of the relative variance of the two components. The correlation for the $R^2_{\text{corr}=\text{min}}$ measure was computed by matching the autocovariances implied by the estimated ARIMA model to those implied by a UC-ARIMA model with various correlations between the two innovations. The R^2 refers to the trend from the UC-ARIMA model with the smallest correlation for which this match was possible, assuming an ARIMA(0, 100) structure for the stationary component in the UC-ARIMA model. The numerical search over trial correlations was carried out using steps of .05. The search was initialized by setting the autocovariances of the stationary process equal to the autocovariances of the innovation in GNP growth.

According to the first, “perfect correlation” measure, all the univariate models attribute much of the quarterly movements in GNP to shifts in its trend. Indeed the theoretical maximum of this R^2 is attained in the perfect correlation case. Examining the range of R^2 measures indicates that although the assumed correlation makes a quantitative difference to the estimated fraction of the variance associated with movements in the trend, in all cases it exceeds one-third, and in all but one it exceeds one-half.⁹

Do these calculations resolve the basic question of the extent to which quarterly movements in GNP are associated with variations in its trend? The answer must be both yes and no. Certainly these results provide further substance to the original findings of time series analysts that aggregate economic variables, and in particular GNP, seem to be integrated processes and therefore contain a stochastic trend. Indeed, even the smallest of the R^2 measures (.34) and of the estimates of the change in the long run forecast (.6) are qualitatively quite different from the values of zero that would obtain were the trend in GNP deterministic rather than stochastic. These statistical measures therefore provide further confirmation of the conclusions drawn from our initial examination of GNP in Figure 1: changes in trends seem to be an important feature of the postwar U.S. experience.

Unfortunately, the ranges of the statistics in Table 1 indicate that the answer must also partially be no. There is nothing intrinsic to this question that rules out the possibility of reaching agreement among the various models. But, at least for the U.S., the different modeling strategies lead to very different quantitative conclusions.¹⁰

We have emphasized that a key reason for this ambiguity is the choice of the correlation of the innovations in the permanent and transitory components, likening this to the difficulty of deciding whether the 1975 downturn was a shift in the permanent or transitory components based solely on visual examination of GNP in Figure 1. Ultimately, however, as economists we should object to having one hand tied behind our back when sorting out trend and stationary movements, in the sense that this analysis focuses solely on the evidence contained in historical GNP. Rather, we might ideally wish to use qualitative evidence about the behavior of the monetary

⁹One approach to resolving the ambiguous results in Table 1 is to develop sensitive statistical measures to assist in choosing one model from another. For example, Cochrane suggested examining the variance of the growth in GNP over several years as a means of estimating the importance of stationary fluctuations in determining the long run behavior of GNP; Joseph Gagnon (1986) has taken a different statistical approach to the same question. Both provide evidence that the low order ARMA models overstate the importance of the trend component in GNP, essentially for the reasons discussed above. However, Campbell and Mankiw (1987b) provide evidence that variance ratios involving long differences might exhibit substantial small sample bias when based on quarterly data covering 30–40 years.

¹⁰These different models might yield similar conclusions when applied to data from different countries. In an initial comparison of the stochastic trend behavior across countries, Clark (1987b) has provided evidence of striking international differences in the relative importance of the trend and stationary components. Indeed, for some countries he suggests that the notion of a stochastic trend as it has been discussed here is inappropriate; following Harvey, he argues for a model in which the trend itself follows a random walk. In this case, GNP would be modeled as being integrated of order two.

authorities in response to the quadrupling of oil prices in 1973–1974, or about the lasting effect on investment, technology, and human capital of the adjustments made in response to the price hike. Since such information is difficult to use in the context of statistical models, we still might hope to perform a less ambitious analysis using additional aggregate variables such as employment, consumption and investment.

Using Additional Information To Study Macroeconomic Trends

Casual inspection of Figure 1 suggests not only that GNP, consumption and investment appear to contain a stochastic trend, but that they contain a common stochastic trend. Indeed, one indication that the “trend” in GNP changed between the 1960s and the 1970s is that this shift appears in consumption and investment as well. This suggests that consumption and investment data contain important information about the trend in GNP, so that there is likely to be a statistical payoff to analyzing these series jointly.¹¹ This section briefly describes some recent work that uses multiple aggregate variables to assess the importance of the trend component in output.

One device used in these multivariate studies to measure the importance of the innovation to the permanent component is the concept of “variance decompositions,” which Christopher Sims (1980a, 1980b) has profitably applied to Vector Autoregressive (VAR) multivariate time series models. When a forecast of GNP (econometric or otherwise) misses its target, and when this forecast error can be traced to a particular factor, it is reasonable to conclude that this factor is important in determining the evolution of GNP—at least within the context of the model that generated the forecast. Forecast error variance decompositions build on this intuition by quantitatively attributing the errors in forecasting the different variables to the various innovations in the system. For example, suppose that, in an empirical multivariate model with permanent and transitory components, forecasts of GNP two years hence are typically off by ± 1.5 percent. These forecast errors will, by construction, arise from errors in forecasting the trend, the stationary component, or both. If the errors from forecasting the trend generally exceed the errors from forecasting the transitory component, then one might attribute more importance to the permanent than the transitory component in determining the evolution of GNP two years hence. In multivariate models with stochastic trends, Sims’ forecast error variance decomposition performs this calculation.

Several recent studies (Olivier Blanchard and Danny Quah, 1987; Clark, 1987b; Campbell and Mankiw, 1987b; Andrew Harvey and Stock, 1987; Matthew Shapiro and Watson, 1988) have used multiple aggregate variables to shed greater light on the importance of the permanent component in GNP. Blanchard and Quah (1987) and Robert King, Plosser, Stock, and Watson (1987) both use these forecast error variance decompositions as a guide to assessing the importance of the permanent component.

¹¹Since government expenditures and net exports are not considered, the national income accounting identity for GNP imposes no restrictions on the number of common trends in GNP, consumption and investment.

Blanchard and Quah (1987) identify this component by assuming that it has no permanent effect on unemployment. In contrast, King, Plosser, Stock, and Watson identify the permanent component in output by assuming that it is also the permanent component in consumption and investment. Despite these and other differences, both papers conclude that between 60 percent and 80 percent of the movements in output at the two- to four-year horizon are explained by movements in the permanent component. Although this literature is still developing, an emerging conclusion from these multivariate studies is that permanent innovations in output play an important role in determining the movements of GNP at horizons typically associated with the business cycle.

Implications

There is a large body of evidence that macroeconomic variables behave as if they contain stochastic trends. Moreover, the empirical research outlined here suggests that the innovations in these stochastic trends play an important role in short run cyclical movements. Multivariate empirical analysis suggests that trend variations and business cycle movements appear to be related. One interpretation of this link is that business cycle fluctuations might be caused by innovations in growth. An alternative explanation—equally consistent with the empirical results—is that cyclical fluctuations cause changes in long run growth. This latter view is consistent with James Tobin's (1980) argument, "With respect to human capital, as well as to physical capital, demand management has important long run supply-side effects. A decade of slack labor markets, depriving a generation of young workers of job experience, will damage the human capital stock far beyond the remedial capacity of supply-oriented measures." Given the challenges associated with differentiating the trend and cyclical components in the time series models discussed above, however, the logical next step is to bring additional information to bear on distinguishing the complex trend-cycle interactions with which Tobin was concerned.

Interpreting Econometric Evidence When Variables Have Stochastic Trends

A Tale of Two Econometricians

Consider the plight of two hypothetical econometricians studying aggregate consumption—but ignorant of the pitfalls that can arise when performing econometric analysis with integrated variables. The econometricians do not, of course, know the true structure of the economy which produced their data; but we do. Specifically, we have constructed for them a simple artificial economy, focusing on aggregate real per capita consumption (C_t), aggregate disposable income (Y_t), and a price index (P_t).

This artificial economy has several key features. First, disposable income consists of two parts: in Milton Friedman's (1957) terminology, permanent and transitory

income. The permanent component of disposable income is assumed to follow a random walk, while the transitory component is an independently and identically distributed random variable that is independent of the permanent component. Using the terminology of the previous section, we therefore assume that disposable income has a stochastic trend plus a stationary component, that this stationary component has no serial correlation, and that the correlation between the trend and stationary innovations is zero.

Second, we suppose that consumers know their permanent income and that they behave according to a narrow interpretation of Friedman’s permanent income hypothesis, thereby consuming precisely the permanent component of their disposable income (which changes from period to period, since it is a random walk). Finally, price changes—assumed to be random and unforecastable with mean zero—do not

Table 2
Regression results using data from the artificial economy (3)–(7)

Typical Regression Results Estimated Regression Equation	Results based on 1000 Replications		Other statistics
	R ²	D.W.	
<i>Econometrician #1</i>			
1. $C_t = 9.16 + .40 P_t$ (28.7) (5.12)	.15	.08	Testing $\beta_P = 0$: 81% Median $R^2 = .15$
2. $C_t = 2.48 + .069 t$ (6.35) (16.9)	.66	.16	Testing $\beta_t = 0$: 91% Median $R^2 = .42$
3. $\Delta C_t = .048 + .28 \Delta Y_t$ (.81) (8.06)	.31	2.27	— Median $\hat{\beta}_{\Delta y_t} = .33$
4. $\Delta Y_t = .41 - .041 C_{t-1}$ (2.33) (-2.15)	.03	1.98	Testing $\beta_{C_{t-1}} = 0$: 48% Median $\hat{\beta}_{C_{t-1}} = -.035$
<i>Econometrician #2</i>			
5. $C_t = .51 + .94 Y_t$ (2.60) (-2.74)	.93	1.88	Testing $\beta_{Y_t} = 1$: 84% Median $\hat{\beta}_{y_t} = .94$
6. $C_t = .45 + .97 C_{t-1} - .01 C_{t-2}$ (2.52) (11.7) (-.13)	.94	2.01	Testing $\beta_{C_{t-2}} = 0$: 9% Median $\hat{\beta}_{C_{t-2}} = -.001$
7. $C_t = .41 + 1.03 C_{t-1} - .07 Y_{t-1}$ (2.36) (14.3) (-.97)	.94	1.97	Testing $\beta_{Y_{t-1}} = 0$: 10% Median $\hat{\beta}_{Y_{t-1}} = .005$
8. $C_t = .47 + .95 C_{t-1}$ (2.21) (45.0)	.94	2.00	Testing $\beta_{P_{t-1}} = 0$: 27% Median $\hat{\beta}_{P_{t-1}} = -.003$
+ .004 $P_{t-1} + .06 \Delta P_{t-1}$ (.17) (.87)			Testing $\beta_{\Delta P_{t-1}} = 0$: 9% Median $\hat{\beta}_{\Delta P_{t-1}} = -.008$

Notes: The regression results in the left half of the table are based on a typical draw of 150 observations constructed according to (1)–(4) using a random number generator in the statistical package RATS (see footnote 12). R² and D.W. respectively denote the regression R-squared and the Durbin-Watson statistic. t-statistics are given in parentheses. The t-statistic on the coefficient on Y_t in regression 5 tests the hypothesis that the coefficient equals one, while all other t-statistics refer to the hypothesis that the coefficient equals zero. The entries in the right half of the table summarize the results of repeating these regressions 1000 times with independently drawn series.

confuse these consumers, in the sense that real consumption and disposable income are determined independently of the price level or its changes.

To summarize these assumptions mathematically, let Y_t^p and Y_t^s respectively denote the permanent and stationary (or transitory) components of disposable income (Y_t). The equations describing output, consumption and the price level in this artificial economy are:

$$(3) \quad Y_t = Y_t^p + Y_t^s$$

$$(4) \quad Y_t^p = Y_{t-1}^p + u_t$$

$$(5) \quad C_t = Y_t^p$$

$$(6) \quad P_t = P_{t-1} + v_t.$$

The innovations Y_t^s , u_t , and v_t are assumed to be mutually independent and, for convenience, to be normally distributed with mean zero and unit variance. According to (3), disposable income (Y_t) is the sum of its permanent (Y_t^p) and transitory (Y_t^s) components where, according to (4), the permanent component evolves according to a random walk. The consumption function (5) states that consumers set their consumption (C_t) equal to their permanent income, so that the marginal propensity to consume out of permanent income is one. Finally, (6) states that, like permanent income, the price level (P_t) is a random walk.

We provide our two econometricians with a typical time series comprised of 150 observations (coincidentally, the number of quarters between 1950:I and 1987:II) on the three variables Y_t , C_t and P_t , generated according to (3)–(6) using a pseudo-random number generator on a portable computer.¹² They do not know but seek to uncover the true relations among these variables. What might they learn?

The first econometrician begins by investigating whether consumers change their consumption patterns based on the price level, which he does by regressing C_t against a constant and P_t . Since the t -statistic on P_t (given on the left hand side of Table 2) of

¹²The following computer program, written for the econometrics package RATS, will produce random time series data according to equations (3)–(6):

```

cal 50 1 4
all 0 87:2
dec rect err(150, 3)
matrix err = ran(1.0)
zer yp ; zer ys ; zer p
eval yp(1) = err(1, 1) ; eval ys(1) = err(1, 2) ; eval p(1) = err(1, 3)
do i = (50:2), (87:2)
  eval yp(i) = yp(i - 1) + err(i, 1)
  eval ys(i) = err(i, 2)
  eval p(i) = p(i - 1) + err(i, 3)
end do i
set y / = yp(t) + ys(t) ; set c / = yp(t)
diff y / 1 dy ; diff c / 1 dc ; diff p / 1 dp

```


5.12 far exceeds the conventional 5 percent two-sided critical value of 1.96, a standard interpretation of his regression is that nominal prices affect consumption. He then checks whether consumption has a linear deterministic time trend by regressing C_t on a constant and time; upon checking the t -statistic, he concludes that it does. However, he recognizes that the low Durbin-Watson statistic from these regressions indicates substantial serial correlation in the residuals (e.g. Robert Pindyck and Daniel Rubinfeld, 1981, pp. 158–164). Troubled by these low Durbin-Watson statistics and thinking consumption to have a deterministic time trend, he differences the data and attempts to estimate the marginal propensity to consume by regressing the change of consumption on the change of income; he finds that the marginal propensity is small indeed. Finally, he recalls Robert Hall's (1978) famous argument that Friedman's Permanent Income Hypothesis implies that consumption follows a random walk, so that the first difference of consumption should be unpredictable. Accordingly, he checks whether lagged consumption is a useful predictor of future changes in consumption; based on the t -statistic of -2.15 , he rejects the random walk hypothesis at the 5 percent significance level. Summarizing, he concludes that consumers have money illusion, that consumption contains a linear time trend, that the marginal propensity to consume is .28, and that past values of consumption are useful in predicting future consumption. That he drew these conclusions is not an artifact of the particular series we gave him to analyze. Repeating this experiment 1000 times using independent draws (the results are summarized on the right hand side of Table 2) indicate that his findings were typical.

Each of his conclusions is wrong.

The second econometrician estimates different regressions. She estimates the marginal propensity to consume by regressing consumption against income and finds it large, but significantly less than one using the usual 5 percent critical value for the t -statistic. When she tests the random walk hypothesis by regressing consumption on two of its lags, the second lag has no statistically significant predictive content; the same conclusion obtains if a lag of income is used. Finally, she finds no additional forecasting value of lagged price changes (although, with the benefit of 1000 replications, we know she would incorrectly reject the hypothesis of no predictive content of the lagged price level 27 percent of the time using the 10 percent critical value). She concludes that Hall's interpretation of Friedman's theory is valid and that the marginal propensity to consume is less than one, although this latter finding seems to be more a matter of statistical than economic significance.

Her conclusions, then, are largely right. Participants at a conference at which these two econometricians present their results might find the exchange entertaining. But they might also long for a systematic way to decide which regression results could be trusted and which could not.

Recent Developments in the Theory of Regression with Integrated Variables

It has long been recognized that the usual techniques of regression analysis can result in highly misleading conclusions when the variables contain stochastic trends. In the econometrics literature, since Clive Granger and Paul Newbold's (1974) influential

simulation study, this has been known as the problem of “spurious regressions.” Largely influenced by the techniques of Box and Jenkins, the accepted “solution” to the “problem of nonstationarity” has been to transform the variables so that they appear to be stationary; in practice this typically means using first differences of the series. Unfortunately, by sidestepping the issues raised by stochastic trends, this approach has little to say about the regressions in Table 2. Moreover, simply using first differences of the data in the regressions generally will not suffice to uncover the true relations in the economy, as the first econometrician’s regression of changes in consumption on changes in income makes clear.¹³

The past few years have seen important progress associated with the specification and analysis of multivariate models with integrated processes. Much of this stems from Fuller’s (1976) and Dickey and Fuller’s (1979) development of the first formal tests for the existence of stochastic trends in a single time series, and from three seminal papers by Granger and Andrew Weiss (1982), Granger (1983), and Robert Engle and Granger (1987). These latter papers provide a mathematical framework for analyzing variables that contain common stochastic trends. Specifically, they consider the case that two (or more) variables might each contain a stochastic trend, i.e. appear to be integrated: consumption and GNP in Figure 1 are good examples. However, casual inspection suggests that these series contain a common trend—and that, by subtracting out this trend, the difference between the two variables is stationary. Formally, they define two integrated process to be *cointegrated* if there is some linear combination (that is, weighted average) of them that is stationary. Thus consumption and GNP are arguably individually integrated; indeed, Dickey and Fuller’s tests fail to reject this null hypothesis. But, assuming that log consumption less log output is stationary, they are jointly cointegrated; that is, they share a common stochastic trend. These theoretical developments (tests for a series being integrated and the concept of cointegration) spurred an enormous amount of recent research into econometric issues that arise when cointegrated processes contain unit roots.¹⁴

To understand the regression results in Table 2, we focus on one of the key lessons of this research: in certain circumstances, even if the right-hand variables (the regressors) are integrated, the usual procedures of OLS analysis can still provide a

¹³In a lively discussion of the history of spurious regression, David Hendry (1986) traces the recognition of this problem in the context of integrated processes to G. Yule (1926). In many circumstances, using differences of time series variables has proven very successful. Aside from having the intuitive appeal of modeling the rates of change of variables (when first differences of logarithms of the series are used), modeling differences of variables is arguably well-suited to producing short run forecasts—as dramatized by the early success of Box-Jenkins time series methods when pitted against the large Keynesian econometric models of the 1960s and 1970s. However, restricting econometric attention to differences in series rules out direct examination of the relation among the levels of the series. Even if the objective is to produce short run forecasts, the econometrician might wish to follow Sims (1980a) and use the additional information available in the levels of variables.

¹⁴The empirical results of James Davidson, David Hendry, Frank Srba, and S. Yeo (1978) provided an important motivation for the development of the theory of cointegration by developing an empirical model of consumption in which consumption and income were implicitly modeled as cointegrated. Granger (1986) and Hendry review recent developments in this area and discuss the link between error-correction models and cointegration.

satisfactory framework for evaluating econometric evidence and for producing forecasts. Furthermore, this research has produced some simple rules-of-thumb that suggest when this is—and is not—likely to be the case.

We first state these rules-of-thumb in a general context, then apply them to the regressions in Table 2.¹⁵ Broadly speaking, the usual assumptions of time series analysis are:

- (i) The error term is serially uncorrelated and is uncorrelated with the regressors (i.e. the regressors are either exogenous or predetermined and the error term is i.i.d.).
- (ii) All the regressors are either deterministic or stationary random variables.

Under these circumstances, the estimated coefficients will become arbitrarily close to their true values in increasingly large samples (that is, they will be consistent). Furthermore, in large samples the null distribution of regression t - and F -statistics can be approximated by normal and F -distributions, respectively.

When some or all of the regressors are integrated processes, condition (ii) clearly is violated. Perhaps surprisingly, however, in many cases the usual techniques of regression analysis will still apply. Specifically, suppose that one is interested in interpreting a particular coefficient or set of coefficients in the regression equation. Although (ii) does not hold, suppose that (ii') does:

- (ii') If there are integrated regressors, either (a) the coefficients of interest are coefficients on mean zero stationary variables; or (b) even if some or all of the coefficients of interest are coefficients on integrated regressors, the regression equation can nevertheless be written in such a way that all the coefficients of interest become coefficients on mean zero stationary variables.

Even if condition (a) in (ii') does not hold, condition (b) still might; an example of this is given below. Under (i) and (ii'), the OLS estimator will be consistent. Moreover, the t - and F -statistics for the coefficient(s) of interest have their usual large sample distributions, so (for example) the standard critical values apply.

In some cases it might be impossible to express the coefficient of interest as a coefficient on a mean zero stationary variable, and instead (ii'') might hold:

- (ii'') The parameter of interest is a coefficient on an integrated process and cannot be written as a coefficient on a stationary variable.

¹⁵These rules-of-thumb are drawn from Sims, Stock and Watson (1986). For expositional simplicity, we assume throughout that all regressions include a constant term for reasons discussed in that paper. All statements assume that the variables are either stationary or integrated of order one, and that the integrated variables have zero drift. The situation with nonzero drift or multiple orders of integration is somewhat more complicated, and the reader is referred to Sims, Stock and Watson or Stock and Kenneth West (1988) for a discussion of this case.

Under (i) and (ii''), the estimator of the coefficient is consistent, but it does not have the usual normal asymptotic distribution, so the usual critical values do not apply.

When the level of one variable is regressed against the level of another, it might well be the case that the error term is not i.i.d. or independent of the regressors. However, suppose that (i) does not hold, but that (i') does:

- (i') The integrated dependent variable is cointegrated with at least one of the integrated regressors, so that the error in the regression equation is stationary but not necessarily serially uncorrelated or independent of the regressors.

Under (i') and (ii'), unless the regressor is strictly exogenous the stationary regressor will typically be correlated with the error term and the parameter estimate will be inconsistent. This is the usual source of "simultaneous equations bias," "omitted variables bias," and "errors-in-variables bias." However, under (i') and (ii''), rather remarkably the estimator of the coefficient of interest is consistent, although it does not have an asymptotic normal distribution.

Finally, suppose that the regressor and at least one dependent variable are integrated, but that there is no cointegrating relationship between the dependent variable and the regressors; it follows that the error term in the regression is integrated. In this case the estimated coefficients on regressors satisfying (ii'') will not be consistent; indeed, these coefficients and the R^2 of the regressions converge to random variables. An important example of this case is the regression of one random walk on another independent random walk, which is Granger and Newbold's (1974) spurious regression problem.

While these rules-of-thumb are rather involved, some intuition for why they work can be developed by first considering the familiar case of a single regressor that is stationary and uncorrelated with a serially uncorrelated error term. Under these assumptions, the OLS estimator is consistent because in large samples the average squared residual formed using any trial coefficient is minimized at or near the true coefficient value; for all coefficients except the true one, the average squared residual is larger but finite. In contrast, in time series regression with two cointegrated variables (so that the dependent variable and the regressor are both integrated but the true error term is stationary), because the two series are trending together, the residual constructed using other than the true coefficient will itself be integrated. Since an integrated process has a variance that tends to infinity, the average squared residual formed using a trial coefficient will grow arbitrarily large as the sample size increases, and will remain finite (and therefore be minimized) only for the true parameter value. This suggests that the coefficient on an integrated regressor can be estimated unusually precisely, as long as the error term is stationary. Furthermore, this unusual behavior of the average squared error suggests that the standard Gaussian asymptotic theory might not apply when there are integrated regressors. In summary, the coefficient in a regression of one integrated variable on another will be consistent if the

two variables are cointegrated although, as it turns out, the asymptotic theory is nonstandard.¹⁶

This intuition can be extended to a regression that includes one stationary regressor, one integrated regressor, and a serially uncorrelated error term that is uncorrelated with either regressor. The preceding reasoning suggests that the coefficient on the integrated regressor will be estimated more precisely than usual (and will have a nonstandard asymptotic distribution); indeed, by the logic of cointegrated regressions, this would be so even if the stationary regressor were omitted from the specification. Turning to the coefficient on the stationary regressor, suppose that the stationary and integrated regressors are uncorrelated. Then the usual reasoning for OLS with uncorrelated regressors suggests that this coefficient will have the conventional large-sample properties. While this argument assumed the two regressors to be uncorrelated, this turns out not to be restrictive: it can be shown that, since the stochastic trend dominates the behavior of an integrated process, the sample correlation between an integrated variable and any mean zero stationary variable tends to zero as the sample size increases. Thus the usual tools of time series regression can be used to examine the coefficient on the mean zero stationary regressor even if there are other regressors that are integrated. Finally, this reasoning can be extended to the case that the coefficient can be written as a coefficient on a mean zero stationary regressor by recognizing that the sum of a stationary variable and an integrated variable is itself integrated, so that the previous logic applies directly to the rearranged regression.

We illustrate these general principles by returning to the curious regressions in Table 2.

Understanding the Two Econometricians' Results

These rules-of-thumb provide a simple framework for explaining the results of our two econometricians, especially since we know the true economic structure that generated their data.

By construction, consumption and the price level are integrated processes that do not share a common stochastic trend, so they are not cointegrated. Thus the first regression clearly falls into the category of a spurious regression. As both Granger and Newbold (1974) and Peter Phillips (1986a) have emphasized, an indication of this situation is the extremely low Durbin-Watson statistic—although this statistic does not provide a formal test for a relation between integrated processes.

¹⁶This argument applies both when the regressor and the dependent variable are cointegrated and when the regressor is a lagged value of the (integrated) dependent variable. In the latter case, the estimated coefficient will converge to one, even if additional lags in the regression have been omitted, so that the error is stationary but serially correlated. Assuming the process to have no drift, in this case the estimated coefficient will have an asymptotic “unit root” distribution as discussed by Fuller (1976) and Dickey and Fuller (1979) which differs dramatically from the usual normal distribution, implying substantial bias towards zero in moderate sample sizes. The estimator converges to this distribution at the rate T rather than $T^{1/2}$ as in the case of conventional time series regression. Also see Fuller's discussion of the case that the variable contains a drift.

The second regression also fails to meet any of the criteria for applying the usual asymptotic approximation to the t -statistic, since consumption is not cointegrated with a linear time trend; indeed, from the definition of cointegration, it cannot be, since a linear time trend is not itself an integrated stochastic process. This regression was analyzed theoretically in an influential article by Nelson and Heejaon Kang (1981); in addition to producing a random R^2 , with 100–150 quarterly observations the residuals from this regression seem to exhibit cycles with a period similar to one associated with the business cycle. Thus a linearly detrended random walk is likely to exhibit spurious periodicity.

Had our econometrician read Friedman closely, he would have recognized the problem with the third specification: the change in disposable income measures the change in permanent income, but with error, since it includes the change in transitory income as well. Thus the coefficient estimator in the third regression is biased downwards, even in arbitrarily large samples.

His final regression satisfies condition (i) (the true coefficient is zero and the error term is the i.i.d. error in permanent income); however, the regressor C_{t-1} is integrated. Thus the coefficient on C_{t-1} satisfies (ii''). It follows that this coefficient is consistent (indeed, its median estimate is close to its true value of zero), but that it will not have the usual asymptotic distribution. For the hypothetical economy (3)–(6), the t -statistic on C_{t-1} is in fact Dickey and Fuller's proposed test for a unit root in consumption; according to Fuller (1976, Table 8.5.2, p. 373) the correct 10 percent critical value for a test against the hypothesis that consumption is stationary is -2.57 . Had the first econometrician used this critical value, he would have failed to reject the hypothesis that consumption follows a random walk.

Why did the second econometrician fare better? By construction, consumption and income are cointegrated in the artificial economy, so the theoretical rules-of-thumb indicate that her first regression will result in a consistent estimator with a nonstandard distribution. Indeed, the median estimate of the 1000 replications is close to—but less than—one, although the t -statistic clearly has a nonstandard distribution. This is a specific example of Stock's (1987) result about the consistency of OLS estimators of cointegrating vectors, even if the error in the regression equation is serially correlated or not independent of the regressors. Additionally, a moment's reflection will indicate that Trygve Haavelmo's (1943) argument that the OLS estimator of the MPC is consistent because of simultaneous equations bias does not hold when consumption and income are cointegrated; rather, the estimator will be consistent, but will exhibit small sample bias.¹⁷

Her next test involved a coefficient on an integrated process in a regression equation with an i.i.d. error. However, the coefficient on the second lag of consump-

¹⁷It is interesting to note that this small sample bias provides an explanation of the difference between estimates of the marginal and average propensities to consume that provided one of the original motivations for Friedman's study of consumption. Specifically, applying nonstandard distribution theory to Friedman's original data, Stock (1986) shows that the marginal propensity to consume has a negative bias of approximately .15; upon adjusting for bias, the marginal and average consumption propensities are both approximately .9.

tion can be rewritten as a coefficient on a stationary variable, so that condition (ii') applies. To see this, consider that part of the regression involving the lags of consumption, and denote the two coefficients as α and β . Algebraic manipulations show that $\alpha C_{t-1} + \beta C_{t-2} = (\alpha + \beta)C_{t-1} - \beta(C_{t-1} - C_{t-2})$; but $C_{t-1} - C_{t-2}$ is stationary, so β clearly can be written as a coefficient on a stationary variable. Thus the theory predicts that the usual t - and F -distributions will apply, and the simulation results on the right side of Table 2 support this prediction. This argument applies equally to her next regression, except that the stationary combination of regressors is $Y_{t-1} - C_{t-1}$, which is stationary because consumption and income are cointegrated.

Her final regression satisfies (i) and, since ΔP_{t-1} is stationary, its coefficient satisfies (ii'), so that the usual asymptotic theory applies to this coefficient (and indeed appears to work well in moderately sized samples, at least according to these calculations). However, the coefficient on P_{t-1} cannot be written as a coefficient on a mean zero stationary regressor, since there are no other regressors with which P_{t-1} is cointegrated. It follows that this estimator is consistent (converging to zero in this case) but has a nonstandard distribution, so that the usual critical values do not apply.¹⁸

These rules of thumb thus provide a general framework for evaluating these regression results with integrated variables, assuming the true economic structure to be known.¹⁹

Evaluating Actual Regression Results

The previous discussion emphasizes the importance of uncovering the stochastic trend properties of actual data to be used in regression analysis. In certain circumstances, economic theory might suggest orders of integration and cointegration among the variables. For example, Campbell (1987) argues that consumption and income being cointegrated is plausible both theoretically and empirically. Often, however, economic theory provides no clear guidance in determining which variables have stochastic trends, which do not, and when the trends are common among those that do.

There are no simple "recipes" for performing time series analysis with integrated variables. One sensible starting point in analyzing time series data in which there might be stochastic trends, however, is to perform a series of initial tests on the data. In particular, Dickey and Fuller's test for a unit root in a single series can provide an important piece of evidence about whether the variable is integrated. Typically an analysis will involve multiple time series, in which case it is important to know their cointegration properties as well. Diagnostic measures for the existence of cointegration

¹⁸Interpreting P_t not as the price level but as the Standard and Poor's Stock Price Index, the reader might recognize her final three equations as simplified versions of Hall's tests of the Permanent Income Hypothesis. Stock and West provide further evidence on the finite sample performance of the asymptotic distribution theory that arises in interpreting Hall's regressions.

¹⁹This discussion has focused on the case of roots exactly equaling one, rather than (say) .999 or 1.001. Mathematically, however, similar warnings about the use of usual regression techniques arise for roots close to one; for example, see Cavanagh (1986) or Phillips (1986b). Thus the rules-of-thumb for identifying those coefficients with asymptotic normal distributions can be seen as tools to guard against misleading inferences if one suspects there to be roots close to one.

have been proposed and analyzed by Engle and Granger, Søren Johansen (1987), Phillips and Ouliaris (1987), and Stock and Watson (1986).²⁰ While these tests certainly are not foolproof, their judicious application can shed considerable light on the way that stochastic trends enter the time series. This in turn provides a framework for implementing the rules of thumb described in this section.

Conclusions

Macroeconomic time series appear to contain variable trends. Moreover, modeling these variable trends as random walks with drift seems to provide a good approximation to the long run behavior of many aggregate economic variables, at least in the U.S. While this general observation is over thirty years old, the application of recently developed statistical “magnifying glasses” has led to several important conclusions for developing and testing macroeconomic theories and for formulating macroeconomic policies.

First, variations in growth trends constitute a quantitatively large part of the movements in real per capita GNP in the United States. Thus the importance of shifts in long run prospects must be recognized even if one is primarily concerned with a relatively short specific historical episode.

Second, the presence of stochastic trends requires careful thought to avoid important econometric pitfalls. If an econometrician wishes to exploit the additional information contained in levels of variables rather than their differences, it is possible to apply a variety of tests and some simple rules-of-thumb to reduce the possibility of making dramatic errors in inference.

Finally, there is evidence not only that aggregate variables contain a substantial stochastic trend component, but that there is a link between changes in this stochastic trend and business cycle movements. This emphasizes the importance of assessing both the short run implications of growth policies and the long run implications of stabilization policies.

■ *The authors thank Francis Bator, Peter Clark, Francis Diebold, Eric Fisher, Joseph Gagnon, Deborah Haas-Wilson, Andrew Harvey, Charles Manski, Danny Quah, Carl Shapiro, Joseph Stiglitz, and Timothy Taylor for helpful comments.*

²⁰Nelson and Plosser provide a clear discussion of the particulars of applying the Dickey-Fuller test to a single series. One shortcoming of integration or cointegration tests is that, with sample sizes typically encountered in macroeconomic research, they can have a fairly low ability to discriminate between the various hypotheses, particularly with multiple variables. For example, with 150 observations it is difficult to distinguish an integrated process from one that is stationary but highly serially correlated. (One implication of this is that tests based on theoretical cointegrating vectors, if they are known, typically have greater power than tests based on estimated cointegrating vectors.) This is closely related to the statistical difficulties discussed in the previous section with distinguishing stochastic trends from stationary components, since both econometric enterprises involve extracting information about relations among variables over the very long run using only 30 to 40 years of data.

Appendix

Stochastic Trends and Integrated Processes

This appendix presents the mathematical link between ARIMA models and the concept of stochastic trends used in the text.

Beveridge and Nelson show that any ARIMA model can be represented as a stochastic trend plus a stationary component, where a stochastic trend is defined to be a random walk, possibly with drift. This representation is most easily obtained for an ARIMA(0, 1, 1) model. Specifically, suppose that Δy_t (where $\Delta y_t = y_t - y_{t-1}$) is a MA(1) process, so that $\Delta y_t = e_t + be_{t-1}$, where e_t is i.i.d. and b is a constant. Let $y_0 = e_0 = 0$, so that y_t can be written as

$$\begin{aligned}
 \text{(A.1)} \quad y_t &= y_{t-1} + e_t + be_{t-1} \\
 &= y_{t-2} + (e_{t-1} + be_{t-2}) + (e_t + be_{t-1}) \\
 &= \sum_{r=1}^t e_r + b \sum_{r=1}^{t-1} e_r \\
 &= (1 + b) \left(\sum_{r=1}^t e_r \right) - be_t.
 \end{aligned}$$

Letting $y_t^p = (1 + b)\sum_{r=1}^t e_r$ and $y_t^s = -be_t$, one can rewrite the final expression in (A.1) as

$$\text{(A.2)} \quad y_t = y_t^p + y_t^s, \quad \text{where } y_t^p = y_{t-1}^p + (1 + b)e_t.$$

Evidently y_t^p is a random walk with no drift and y_t^s is stationary (indeed, here y_t^s is serially uncorrelated). Equation (A.2) gives the Beveridge-Nelson representation of an ARIMA(0, 1, 1) process in terms of a stochastic trend (y_t^p) and a stationary component (y_t^s). Note that the innovations in the two components are both proportional to e_t , i.e. they are perfectly correlated; if $b > 0$, the correlation is -1 , whereas if $b < 0$, the correlation is $+1$.

The representation can in fact be obtained for general ARIMA($p, 1, q$) processes. This is most easily shown using lag (or “backshift”) operator notation, where by definition $L^j x_t = x_{t-j}$ and $a(L) = \sum_{j=0}^p a_j L^j$. With this notation, an ARIMA($p, 1, q$) model can be written,

$$\text{(A.3)} \quad a(L)\Delta y_t = f + b(L)e_t$$

where f is a constant and $a(L)$ and $b(L)$ are lag polynomials of order p and q , respectively. Inverting $a(L)$, one can write (A.3) in its infinite moving average form,

$$(A.4) \quad \Delta y_t = g + c(L)e_t$$

where $g = f/\sum_{j=0}^p a_j$ and $c(L) = b(L)/a(L)$. Next, recursively substitute lagged Δy_t as was done in (A.1) and assume that $y_0 = 0$ and $e_r = 0$ for $r \leq 0$, so that an expression similar to the final one in (A.1) obtains:

$$(A.5) \quad y_t = gt + h \sum_{r=1}^t e_r + d(L)e_t$$

where $h = \sum_{j=0}^{\infty} c_j$ and $d_i = -\sum_{j=i+1}^{\infty} c_j$. Thus (A.5) can be rewritten as

$$(A.6) \quad y_t = y_t^p + y_t^s, \quad \text{where} \quad y_t^p = g + y_{t-1}^p + he_t \quad \text{and} \quad y_t^s = d(L)e_t$$

which provides the Beveridge-Nelson decomposition for general ARIMA($p, 1, q$) processes, where the stochastic trend has drift g . Note that h is easily calculated as $\sum_{j=0}^q b_j/\sum_{i=0}^p a_i$. As in the ARIMA(0, 1, 1) case, the innovations in the trend and the cyclical components are both proportional to e_t , and thus are perfectly correlated.

References

- Aoki, Masanao**, "How to Build State Space Models for Nonstationary Time Series and How to Measure Random Walk Components," University of California-Los Angeles, Department of Economics Working Paper No. 438, March 1987.
- Beveridge, Stephen, and Charles R. Nelson** "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle,'" *Journal of Monetary Economics*, 1981, 7, 151-174.
- Blanchard, Olivier J., and Danny Quah**, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," manuscript, MIT, 1987.
- Box, George E. P., and Gwilym M. Jenkins**, *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1970.
- Campbell, John Y.**, "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica*, 1987, 55, 1249-1274.
- Campbell, John Y., and N. Gregory Mankiw** (1987a), "Are Output Fluctuations Transitory," *The Quarterly Journal of Economics*, 1987, 102, 857-880.
- Campbell, John Y., and N. Gregory Mankiw** (1987b), "Permanent and Transitory Components in Macroeconomic Fluctuations," *The American Economic Review*, 1987, 77, 111-117.
- Cavanagh, Christopher L.**, "Roots Local to Unity," Harvard University Department of Economics Discussion Paper, 1986.
- Clark, Peter K.** (1987a), "The Cyclical Component of U.S. Economic Activity," *The Quarterly Journal of Economics*, 1987, 102, 797-814.
- Clark, Peter K.** (1987b), "Trend Reversion in Real Output and Unemployment," Discussion Paper, Graduate School of Business, Stanford University, 1987.
- Cochrane, John H.**, "How Big is the Random Walk Component in GNP," *Journal of Political Economy*, forthcoming.
- Davidson, James E. H., David F. Hendry, Frank Srba, and S. Yeo**, "Econometric Modelling

of the Aggregate Time-Series Relationship Between Consumer's Expenditure and Income in the United Kingdom," *Economic Journal* 1978, 86.

Dickey, David A. and Wayne A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 1979, 74, 1057-1072.

Engle, Robert F., and Clive W. J. Granger, "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 1987, 55, 251-276.

Friedman, Milton, *A Theory of the Consumption Function*. Princeton: Princeton University Press, 1957.

Fuller, Wayne A., *Introduction to Statistical Time Series*. New York: Wiley, 1976.

Gagnon, Joseph E., "Short Run Models and Long Run Forecasts," *Quarterly Journal of Economics*, forthcoming.

Granger, Clive W. J., "Co-integrated and Error Correcting Models," UCSD Discussion Paper no. 83-13a, 1983.

Granger, Clive W. J., "Developments in the Study of Cointegrated Economic Variables," *Oxford Bulletin of Economics and Statistics*, August 1986, 48, No. 3, 213-228.

Granger, Clive W. J., and Andrew A. Weiss, "Time Series Analysis of Error-Correction Models," UCSD Discussion Paper no. 82-28, 1982.

Granger, Clive W. J., and Paul Newbold, "Spurious Regressions in Econometrics," *Journal of Economics*, 1974 2, 111-120.

Granger, Clive W. J., and Paul Newbold, *Forecasting Economic Time Series*. New York: Academic Press, 1977.

Haavelmo, Trygve, "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 1943 11, no. 1, 1-12.

Hall, Robert E., "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 1978, 86, no. 6, 971-87.

Harvey, Andrew C., "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 1985, 3, 216-227.

Harvey, Andrew C., *Time Series Models*. New York: Wiley, 1981.

Harvey, Andrew C. and James H. Stock, "Continuous Time Autoregressive Models with Common Stochastic Trends," manuscript, Harvard University, 1987; forthcoming, *Journal of Economic Dynamics and Control*.

Hendry, David F., "Econometric Modelling with Cointegrated Variables: An Overview," *Oxford Bulletin of Economics and Statistics*, August 1986, 48, No. 3, 201-212.

Johansen, Søren, "Statistical Analysis of Cointegrating Vectors," manuscript, University of Copenhagen, 1987; forthcoming, *Journal of Economic Dynamics and Control*.

King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson, "Stochastic Trends and Economic Fluctuations," NBER Discussion Paper No. 2229, 1987.

Nelson, Charles R., and Heejoon Kang, "Spurious Periodicity in Inappropriately Detrended Time Series," *Econometrica*, 1981, 49, 741-751.

Nelson, Charles R., and Charles I. Plosser, "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 1982, 129-162.

Orcutt, Guy H., "A Study of the Autoregressive Nature of the Time Series Used for Tindbergen's Model of the Economic System of the United States, 1919-1932," *Journal of the Royal Statistical Society B*, 1948, 10, no. 1, 1-45.

Phillips, Peter C. B. (1986a), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 1986, 33, 311-340.

Phillips, Peter C. B. (1986b), "Towards a Unified Asymptotic Theory of Autoregression," Yale University, Cowles Foundation Discussion Paper no. 782, 1986.

Phillips, Peter C. B., and S. Ouliaris, "Asymptotic Properties of Residual Based Tests for Cointegration," Yale University, Cowles Foundations Discussion Paper no. 847 (1987).

Pindyck, Robert S., and Daniel L. Rubinfeld, *Econometric Models and Economic Forecasts*, second edition. New York: McGraw-Hill, 1981.

Prescott, Edward C., "Theory ahead of Business Cycle Measurement," Carnegie-Rochester Conference on Public Policy, 1987.

Shapiro, Matthew D., and Mark W. Watson, "Sources of Business Cycle Fluctuations," *NBER Macroeconomics Annual*, 1988, forthcoming.

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica*, January 1980, 48 (1980a), 1-48.

Sims, Christopher A., "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered," *American Economic Review*, May 1980, 70 (1980b), 250-257.

Sims, Christopher A., James H. Stock, and Mark W. Watson, "Inference in Linear Time Series Models with Some Unit Roots," manuscript, Stanford University, 1986.

Stock, James H., "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, 1987, 55, 1035-1056.

Stock, James H., "A Reexamination of Fried-

man's Consumption Puzzle," *Journal of Business and Economic Statistics*, forthcoming.

Stock, James H., and Mark W. Watson, "Testing for Common Trends," *Journal of the American Statistical Association*, forthcoming.

Stock, James H., and Kenneth D. West, "Integrated Regressors and Tests of the Permanent Income Hypothesis," *Journal of Monetary Economics*, 1988, in press.

Tobin, James, "Stabilization Policy Ten Years

After," *Brookings Papers on Economic Activity*, 1980, 1, 19–90.

Watson, Mark W., "Univariate Detrending Methods with Stochastic Trends," *Journal of Monetary Economics*, 1986, 18, 49–75.

Yule, G. U., "Why Do We Sometimes Get Nonsense-Correlations Between Time-Series?" *Journal of the Royal Statistical Society B*, 1926, 89, 1–64.