Journal of Monetary Economics 4 (1978) 637–660. © North-Holland Publishing Company

MONEY, INCOME, AND SUNSPOTS: MEASURING ECONOMIC RELATIONSHIPS AND THE EFFECTS OF DIFFERENCING

Charles I. PLOSSER and G. William SCHWERT*

University of Rochester, Rochester, NY 14627, U.S.A.

This paper discusses the question of whether economic time series regression models should be estimated between the levels or the changes of the variables of interest. We argue that many economic models should be estimated between the changes of the variables, rather than the levels of the variables. In addition, comparisons of the levels and changes regressions can be used as a crude test of model specification. These issues are illustrated with examples from Friedman and Meiselman's (1963) study of annual income and consumption and with data on sunspot activity from 1897–1958.

1. Introduction

It is common to find economic relationships which are formulated and empirically investigated in terms of the levels of time series variables. Some simple examples might be money and prices, or income and consumption. Although at the theoretical level most model builders recognize that their models can be equivalently formulated in terms of the changes in the variables, for example the change in money and the change in prices, many do not admit this equivalence when it comes time to estimate the model. The reason appears to stem from the fact that many time series regressions that are computed using the levels of economic variables produce strong relationships as measured by R^2 or the adjusted R^2, $\bar{R}^2$, but when the same model is estimated in the changes, the relationship becomes weak and may even disappear entirely. This phenomenon has led some people to question the use of differencing because, they argue, it somehow removes part of the relationship between the variables.

This paper focuses on the effects of differencing in linear regression models. Although some of the technical aspects of this paper are well known in the econometrics literature, we feel that they are worthy of emphasis because they are often overlooked or ignored in much applied work. Our point of

*Plosser's participation in this research occurred while he was at the Graduate School of Business, Stanford University. We have benefited from discussions with Paul Evans, Kenneth Gaver, Martin Geisel, Michael Jensen, Ben McCallum, Charles Nelson, David Pierce, Christopher Sims, and Arnold Zellner, but retain responsibility for all errors.

view can be summarized by stating that the real issue is not differencing, but an appropriate appreciation of the role of the error term in regression models. This means that if the linear specification of the model is correct, the model can be estimated in the levels, in the first differences, or in the second differences with similar results as long as the autocorrelation properties of the regression disturbances are taken into account. In other words, differencing makes little difference.

On the other hand, we argue that the dangers associated with ignoring the effects of underdifferencing (that is, insufficient differencing) can be substantially greater than those associated with ignoring the effects of overdifferencing. For example, if the error term in the first differences regression is well behaved (e.g., serially uncorrelated), the error term in the levels (the underdifferenced) regression follows a nonstationary random walk. This can cause least squares estimators of the regression coefficients in the levels regression to be inconsistent. However, the error term in the second differences (the overdifferenced) regression follows a stationary first-order moving average process, so the estimators of the regression coefficients are unbiased and consistent, although they are not efficient.

Section 2 shows the effects of differencing on the linear regression model. In section 3 we present some examples to illustrate the arguments by reexamining some of the results of Friedman and Meiselman (1963). Section 4 presents a brief discussion of the relationship of differencing to certain types of specification errors; in particular, the effects of differencing on models where the variables are measured with error. Finally, section 5 provides some concluding remarks.

The question of whether to use the differencing transformation has important ramifications for empirical work in many areas of economics. Our goal in this paper is to clarify the issues related to differencing so that the empiricist can assess the costs and benefits of using differencing as an integral part of his data analysis. We show that differencing can often lead to important insights concerning the relationships between time series variables.

2. Some statistical issues associated with differencing

Many economic variables have a strong tendency to trend through time. Therefore, the levels of these variables can be characterized as nonstationary since they do not have a constant mean over time. The statistical problems associated with the estimation of regression relationships among nonstationary time series variables are documented by Yule (1926), and more recently by Granger and Newbold (1974). In particular, Granger and Newbold emphasize that, in many instances, regressions estimated among nonstationary variables are found to have residuals which are highly autocorrelated, as indicated by very low Durbin–Watson statistics.

Furthermore, they go on to show that usual tests of hypotheses about the regression coefficients in the presence of autocorrelated residuals can produce misleading results by increasing the probability of falsely rejecting the null hypothesis that there is no relationship between the variables (Type I error).[1] Although the dangers associated with autocorrelated residuals have been recognized for some time, we agree with Granger and Newbold that it is distressing that much applied work consistently ignores these dangers.[2]

One solution to the problems associated with estimating and interpreting regression equations among nonstationary variables which is tentatively suggested by Granger and Newbold is to difference the regression relationship until each variable is stationary prior to estimating the regression equation. The use of the difference transformation eliminates a linear time trend and/or a stochastic trend (such as exemplified by a random walk) from each of the variables in the regression relationship.

Consider the effect of differencing on the properties of the regression model. For example, suppose a correctly specified regression model is given by

$$y_t = \alpha + \beta x_t + \varepsilon_t, \qquad t = 1, 2, \ldots, T, \tag{2.1}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables, independent of $\{x_t\}$, with mean zero and constant variance σ_ε^2. Furthermore, note that we have not made any assumptions about the properties of y_t or x_t. They might be the levels or the changes of economic variables. In fact, economic theory does not in general tell us which transformations of the variables are linearly related with a stationary disturbance term, yet the stationarity of the error term can be a very important factor in the estimation of the regression equation. The ordinary least squares estimator for β is:

$$\hat{\beta} = \sum_{t=1}^{T} \tilde{x}_t \tilde{y}_t \Big/ \sum_{t=1}^{T} \tilde{x}_t^2, \tag{2.2}$$

where a tilde '~' indicates the variables are measured as deviations from their sample means. This estimator for β is consistent, unbiased and efficient.

[1]Vinod (1976) presents bounds for t-ratios in the presence of autocorrelated errors which depart substantially from the Student t-distribution in some cases.

[2]Although autocorrelated residuals in the levels regression may indicate other types of misspecification, for example, omitted variables, in this paper we are primarily concerned with the specification of the disturbance term and the effects of differencing. Therefore, we assume that all of the relevant variables have been included in the equation and the functional form of the relationship is correctly specified.

2.1. The effects of overdifferencing

Now suppose we difference (2.1) to obtain

$$\Delta y_t = \alpha' + \beta \Delta x_t + u_t, \qquad t = 2, 3, \ldots, T, \tag{2.3}$$

where

$$u_t = \Delta \varepsilon_t = \varepsilon_t - \theta \varepsilon_{t-1}, \tag{2.4}$$

and $\Delta = (1-L)$, $\theta = 1$ and L is the lag operator such that $L^j z_t = z_{t-j}$. Eq. (2.3) is the regression equation that relates the changes in y_t to the changes in x_t. Furthermore, the changes in these variables are related in the same way as the levels of the variables, except for the constant term. Therefore, whether we estimate (2.1) or (2.3) we are estimating the same parameter, β. The only way that (2.3) differs from (2.1) is in the constant term and the error structure. The error structure of (2.3) is a first-order moving average process (MA(1)) as seen in (2.4), with a MA parameter equal to one.[3] Since (2.1) is the correct specification of the relationship, (2.3) can be referred to as the *overdifferenced* model and the moving average nature of the errors in (2.4), with $\theta = 1$, is characteristic of the type error structure found in overdifferenced regression equations.[4]

Suppose that we estimate (2.3) using ordinary least squares. In this case the estimator for β is

$$\hat{\beta}^* = \sum_{t=2}^{T} \Delta \tilde{x}_t \Delta \tilde{y}_t \Big/ \sum_{t=2}^{T} (\Delta \tilde{x}_t)^2 . \tag{2.5}$$

This is an unbiased and consistent estimator of β. Conditional on $\{x_t\}$, the usual least squares estimator of the sampling variance of $\hat{\beta}^*$ is

$$s^2(\hat{\beta}^*) = \frac{1}{T-3} \sum_{t=2}^{T} \hat{u}_t^2 \Big/ \sum_{t=2}^{T} (\Delta \tilde{x}_t)^2 , \tag{2.6}$$

where $\hat{u}_t$ represents the least squares residual at t.[5] Unfortunately, the estimator in (2.6) may be biased if the disturbances are autocorrelated. In this particular case, where the disturbances follow the MA(1) process (2.4)

[3] This MA(1) model implies a first-order autocorrelation $\rho_1 = -\theta/(1+\theta^2) = -0.5$ for the disturbances in (2.4). This is roughly equivalent to a Durbin–Watson statistic of 3.0.

[4] In the more general case where the disturbances in (2.1) follow a stationary autoregressive-moving average process, overdifferencing introduces a unitary root into the moving average structure of the disturbances in (2.3).

[5] The matrix equivalent to (2.6) is $\hat{\sigma}_u^2(\Delta \tilde{x}' \Delta \tilde{x})^{-1}$.

with $\theta=1$, it is straightforward to show that an unbiased estimate of the sampling variance of $\hat{\beta}^*$ conditional on $\{x_t\}$ is

$$\bar{s}^2(\hat{\beta}^*)=\hat{\sigma}_u^2(1-\hat{\rho}_1)\Big/\sum_{t=2}^{T}(\Delta\hat{x}_t)^2, \tag{2.7}$$

where $\hat{\sigma}_u^2$ is an unbiased estimate of σ_u^2 and $\hat{\rho}_1$ is an estimate of the first-order autocorrelation coefficient of the Δx_t series.[6] It is interesting to note that if x_t is a random walk, then Δx_t is serially uncorrelated ($\hat{\rho}_1 \cong 0$), and the usual least squares formula (2.6) yields an unbiased estimate of the sampling variance even though the disturbances are autocorrelated. If the changes in x_t are positively autocorrelated, the usual formula (2.6) overestimates the true sampling variance of $\hat{\beta}^*$. If changes in x_t are negatively autocorrelated, the usual formula underestimates the true sampling variance of $\hat{\beta}^*$.[7]

If the error structure is known, a generalized least squares procedure should be used to take account of the moving average process of the disturbances in (2.4). However, if the model is known to be overdifferenced ($\theta=1$), the generalized least squares procedure is equivalent to simply estimating the model in terms of the levels of the variables y_t and x_t.[8] Unfortunately, rarely is this much known about the error process. Therefore, a more practical approach would be to estimate the moving average parameter θ to see whether it is close to one. This approach is feasible, although there are some statistical problems of estimating the MA parameter when $\theta=1$.[9] Plosser and Schwert (1977) perform Monte Carlo experiments of the joint estimation of β and θ in (2.3) and (2.4). Even though there are certain problems with estimating θ when it is equal to one, the estimates of β are not adversely affected. Furthermore, the efficiency of the resulting estimator for β is improved relative to ordinary least squares estimator in (2.5) which is

[6]Equation (2.7) is obtained from the correct formula for the sampling variance of the least squares estimator, $(\Delta\tilde{x}'\Delta\tilde{x})^{-1}\ \Delta\tilde{x}'\hat{\Sigma}\Delta\tilde{x}\ (\Delta\tilde{x}'\Delta\tilde{x})^{-1}$, where $\hat{\Sigma}$ is an unbiased estimator of the covariance matrix of the regression disturbances. Note that if $\hat{u}_t$ represents the least squares residual, the expression

$$\frac{1}{T-3}\sum_{i=2}^{T}\hat{u}_i^2,$$

does not yield an unbiased estimate of σ_u^2 if the Δx_t series is autocorrelated.

[7]If there is more than one regressor, the sampling variance of $\hat{\beta}^*$ will depend not only on the autocorrelation structure of the individual regressors, but on the covariances among them as well. Consequently, it would be very difficult to predict a priori the direction of the bias contained in the usual least squares formula for the sampling variance of $\hat{\beta}^*$ in these more complicated circumstances.

[8]Maeshiro and Vali (1977) prove this equivalence.

[9]When $\theta=1$ the MA(1) process in (2.4) is noninvertible; it does not have an infinite autoregressive representation. Further discussion can be found in Box and Jenkins (1976) or Plosser and Schwert (1977).

calculated ignoring the moving average process for the disturbances. Therefore, the costs associated with overdifferencing may not be large when care is taken to analyze the properties of regression disturbances.

2.2. *The effects of underdifferencing*

Now consider the effects of underdifferencing. In other words, rather than estimating the relationship between y_t and x_t we consider the regression equation

$$Y_t = \alpha'' + \beta X_t + \eta_t, \qquad t = 0, 1, 2, \ldots, T, \tag{2.8}$$

where

$$\eta_t = \rho \eta_{t-1} + \varepsilon_t, \tag{2.9}$$

and

$$\Delta Y_t = y_t, \quad \Delta X_t = x_t \quad \text{and} \quad \rho = 1.$$

In other words, the correct model is in terms of the changes in Y_t, y_t, and the changes in X_t, x_t, but we estimate (2.8) instead. The identifying characteristic of the *underdifferenced* regression eq. (2.8) is that the disturbance term is *nonstationary*. In this case, the disturbances follow a first-order autoregressive process with $\rho = 1$, and eq. (2.9) describes a sequence of disturbances $\{\eta_t\}$ which follow a random walk.[10]

Suppose that we try to estimate (2.8) using ordinary least squares. The usual formula provides the estimator for β:

$$\hat{\beta}' = \sum_{t=0}^{T} \tilde{X}_t \tilde{Y}_t \Big/ \sum_{t=0}^{T} \tilde{X}_t^2 . \tag{2.10}$$

Because $\{\eta_t\}$ is nonstationary, the sampling distribution of this estimator is not well behaved. For example, since η_t does not have a finite unconditional mean or variance, the distribution of the estimator $\hat{\beta}'$ does not have finite moments.[11] Furthermore, the estimator of β given in (2.10) may be incon-

[10]In the more general case where $\{\varepsilon_t\}$ follows a stationary autoregressive–moving average process, underdifferencing leaves a unitary root in the autoregressive structure of the disturbance which causes the disturbances to be nonstationary.

[11]Since the level of a random walk is the accumulation of all past shocks,

$$\eta_t = \sum_{i=0}^{\infty} \varepsilon_{t-i},$$

the unconditional distribution of η_t does not have finite moments. Alternatively, the distribution of $\hat{\beta}'$ given the 'initial condition', η_0, is not independent of η_0 for any sample size.

sistent. A necessary condition for the consistency of $\hat{\beta}'$ is that:

$$\operatorname{plim}\left[\frac{1}{T-1}\sum_{t=0}^{T}\hat{\eta}_t^2 \Big/ \sum_{t=0}^{T}\tilde{X}_t^2\right]=0, \tag{2.11}$$

where $\hat{\eta}_t$ is the ordinary least squares residual from (2.8). If η_t follows a random walk as in (2.9), the numerator of (2.11) goes to infinity as $T\to\infty$. Nevertheless, (2.11) could still converge to zero if the denominator goes to infinity faster than the numerator as the sample size increases.[12] If this does not occur, the estimator $\hat{\beta}'$ is inconsistent. Consequently, under certain conditions estimating regression relationships in the presence of nonstationary errors could result in estimates that are seriously misleading.

A logical question to ask is under what conditions is $\hat{\beta}'$ consistent? One simple case occurs when X_t is a linear time trend. Under these circumstances $\sum_{t=0}^{T}\tilde{X}_t^2$ goes to infinity with T^3. Therefore, it is of higher order than T and the expression (2.11) converges to zero as $T\to\infty$.

An alternative method of estimating (2.8) is to apply a Cochrane–Orcutt (1949) transformation, or estimate the parameters α'', β and ρ jointly using a maximum likelihood procedure. One way of doing the joint estimation would be to consider (2.8) as a transfer function model of the type discussed in Box and Jenkins (1976). If the resulting estimate of ρ is close to one, as it should be in this case, differencing would be indicated, leading to the correct model in (2.1).

Thus, the costs associated with underdifferencing can be serious since the error term does not have a stationary distribution. Nevertheless, careful analysis of the properties of the error terms should lead to the correct specification of the model.

3. Some economic examples

3.1. The quantity theory of money

In order to illustrate the points brought out in the previous section, it is useful to consider some simple examples. Our first example is motivated by the classical quantity theory of money. In its most simple form, the model can be written:

$$y=v+m, \tag{3.1}$$

where y, v, and m denote the logs of nominal income, velocity and money,

[12]More precisely, the sequence $1/(T-1)\sum_{t=0}^{T}\hat{\eta}_t^2$ is of order T; therefore, the power of T that is required to bring the probability limit of the expression $T^{-p}\cdot\sum_{t=0}^{T}\tilde{X}_t^2$ to a nonzero constant must be greater than one, $p>1$.

respectively. The classical quantity theory implies that m and v are uncorrelated,[13] and regressions of y_t on m_t,

$$y_t = \alpha + \beta m_t + \varepsilon_t, \tag{3.2}$$

should yield a slope coefficient near unity, $\beta \cong 1$.

If the slope coefficient is unity, the constant term in this regression, α, is the average value of the log of velocity over the sample period, and the disturbance, ε_t, is the deviation of the log of velocity from this average value in each period. In this case, the error structure associated with this regression has the same stochastic characteristics as the log of velocity.

The quantity equation as stated in (3.1) can be rewritten as

$$\Delta y = \Delta v + \Delta m, \tag{3.3}$$

where Δy, Δv and Δm denote the changes in the logs of nominal income, velocity, and money, that is, their rates of growth. As before, the quantity theory would imply that regressions of Δy_t on Δm_t,

$$\Delta y_t = \alpha' + \beta \Delta m_t + u_t, \tag{3.4}$$

should yield a slope coefficient near unity. If $\beta = 1$, the intercept α in (3.4) could be interpreted as the average rate of growth of velocity, and the disturbances have the same stochastic properties as Δv_t.

In this context it makes no difference whether the relationship is measured between y and m, or between Δy and Δm, as long as the stochastic properties of the error term are taken into account. Gould and Nelson (1974) argue that annual velocity data from 1869 to 1960 are well approximated by a random walk. Thus, if the stochastic properties of the error term are ignored and the error term ε_t is nonstationary, then regressions estimated between y and m may produce misleading results.

Using the data on annual income and money provided in Friedman and Meiselman (1963), we estimate the contemporaneous relationship between the log of income, y_t, and the log of money, m_t, several different ways. The results are summarized in table 1. The first column in table 1 gives the results of the levels regression of y_t on m_t. Note that the coefficient on m_t appears to be precisely estimated and significantly different from unity at the 1% level when the usual test statistics are applied. However, there appears to be significant residual autocorrelation as evidenced by a first-order serial correlation coefficient (r_1) of 0.82. This corresponds to a Durbin–Watson statistic of approximately 0.36.

[13]Some empirical validation of the independence of changes in velocity with respect to changes in money can be found in Gould et al. (1978).

Table 1
Regression models of the log of income on the log of money: 1897–1958[a]

	Regression model				
Variable	(1) Levels y	(2) Time trend y	(3) Cochrane–Orcutt y	(4) First differences Δy	(5) Second differences $\Delta^2 y$
Constant	2.257 (0.161)	36.80 (6.67)	−1.078 (1.561)	−0.016 (0.012)	0.0003 (0.012)
m	0.842 (0.015)	1.168 (0.064)	1.127 (0.122)		
Δm				1.141 (0.126)	
$\Delta^2 m$					1.193 (0.194)
t		−0.020 (0.004)			
$\hat{\rho}_1$[b]			0.956 (0.029)		
$\bar{R}^2$	0.980	0.986	0.994	0.575	0.384
$\hat{\sigma}_e^2$	0.0163	0.0114	0.0048	0.0049	0.0088
r_1	0.82 (0.13)	0.74 (0.13)	0.12 (0.13)	0.12 (0.13)	−0.36 (0.13)

[a]Standard errors are in parentheses; $\bar{R}^2$ is the coefficient of determination, adjusted for degrees of freedom; $\hat{\sigma}_e^2$ is the residual variance; and r_1 is the first-order autocorrelation coefficient of the residuals with its large sample standard error in parentheses.
[b]First-order autoregressive parameter for the error structure. The standard errors for the parameters in this regression are based on asymptotic distribution theory.

Suppose we hypothesize that both y_t and m_t contain deterministic time trends, but they are independent of one another except for the common trend. This implies that the strong relationship obtained by regressing y_t on m_t is spurious due to the omission of an important variable, time. An alternative rationale for this specification might be that the time trend variable proxies for some omitted variable (possibly velocity or technological progress) that has a deterministic trend. Such a hypothesis suggests the inclusion of a time trend variable in the regression of y_t on m_t,

$$y_t = \alpha + \beta m_t + \gamma t + \varepsilon_t. \tag{3.5}$$

The estimate of this time trend model in table 1 indicates that including the time trend variable increases the coefficient on m_t substantially and the estimated standard error of the coefficient is larger. Based on usual test statistics it seems that $\hat{\beta}$ is significantly greater than unity at conventional significance levels. However, it is important to note that including the time trend does not eliminate the residual autocorrelation. The first-order serial correlation coefficient is 0.74 (a Durbin–Watson statistic of approximately 0.52).

A third approach to estimating the relationship between y_t and m_t without differencing the data is to assume that the residuals follow a first-order autoregressive process and jointly estimate the parameters. The third column of table 1 presents estimates of the Box–Jenkins transfer function model,

$$y_t = \alpha + \beta m_t + \eta_t, \tag{3.6a}$$

$$\eta_t = \rho_1 \eta_{t-1} + \varepsilon_t, \tag{3.6b}$$

which is analogous to the Cochrane–Orcutt (1949) estimation procedure. The estimate of the autoregressive coefficient ρ_1 is 0.956 with an estimated asymptotic standard error of 0.029, so ρ_1 seems very close to unity.[14] The 'Cochrane–Orcutt' estimate of the coefficient of m_t is closer to the 'time trend' estimate than the 'levels' estimate given by the regression of y_t on m_t. As one would expect, the standard error of this estimate is also larger and indicates that the coefficient is not significantly different from unity at the 5% level of significance. Finally, the residuals no longer display any significant autocorrelation (at lag one).

Since $\hat{\rho}_1$ is so close to one in the Cochrane–Orcutt regression, the first differences regression of Δy_t on Δm_t in (3.4) should yield similar results. As seen in column 4 of table 1, the estimated coefficients, the estimated standard errors, and the variance of the residuals from these two models are almost identical. The constant term in the 'first differences' regression corresponds to the coefficient on the time trend variable in the time trend regression.[15] While the estimates of this trend parameter are similar in the time trend and first differences models, the estimated standard error of this parameter is much larger for the first differences regression. This reflects, at least in part,

[14]If $\rho_1 = 1$, $\hat{\rho}_1$ is downward biased and the asymptotic standard error tends to understate the sampling variability of $\hat{\rho}_1$ in finite samples. Zellner and Plosser (1977) discuss these issues in the case of univariate autoregressive models. Also, it should be noted that $\bar{R}^2$ in the Cochrane–Orcutt regression is the proportion of the sample variance of y_t explained, *not* the proportion of the sample variance of $(y_t - \hat{\rho}_1 y_{t-1})$ explained by the regressor. Thus, if $\hat{\rho}_1$ is close to one, $\bar{R}^2$ will be close to one, even if the regression coefficient is zero.

[15]Consider the first difference of equation (3.5):

$$\begin{aligned} y_t - y_{t-1} &= (\alpha + \beta m_t + \gamma t + \varepsilon_t) - (\alpha + \beta m_{t-1} + \gamma(t-1) + \varepsilon_{t-1}) \\ &= (\alpha - \alpha) + \beta(m_t - m_{t-1}) + \gamma(t - t + 1) + (\varepsilon_t - \varepsilon_{t-1}), \\ \Delta y_t &= \gamma + \beta \Delta m_t + \Delta \varepsilon_t, \end{aligned}$$

so γ is the constant term in the first differences regression.

the residual autocorrelation in the time trend regression which causes the estimated standard errors to be biased downward for that model. Thus, there is no evidence of a significant negative trend in velocity over time when the standard errors are estimated correctly.[16]

Note that the $\bar{R}^2$ statistic for the differences regression is lower than for the levels regressions. While many would be disappointed by this relative 'lack of fit' in the differences regression, such disappointment is not justified. If the regressand y_t is nonstationary, the sample variance of the levels of y_t is likely to be much larger than the sample variance of the differences Δy_t. Therefore, for the same residual variance, $\bar{R}^2$ is larger in the levels than the differences.[17]

The residual variance σ_e^2 is the relevant basis for comparing the levels regression with the differences regression. Note that the error term in the levels regression, ε_t, is related to the error term in the first differences regression, $u_t = \varepsilon_t - \varepsilon_{t-1}$. If the error term for the levels regression is serially uncorrelated, the residual variance for the differences regression is twice as large as the levels regression, $\sigma_u^2 = 2\sigma_\varepsilon^2$. On the other hand, if the error term for the levels regression, ε_t, follows a random walk, $\varepsilon_t = \sum_{i=0}^{\infty} u_{t-i}$, the error term for the differences regression, u_t, is serially uncorrelated, and the residual variance is much larger for the levels regression than for the differences regression.

In order to show that differencing should not matter in a properly specified linear regression model, suppose that we difference again and regress $\Delta^2 y_t$ on $\Delta^2 m_t$. This 'second differences' regression is an over-differenced model if the first differences model is well specified, so it should have moving average errors. Inspection of the final column in table 1 bears these suspicions out. Note that the estimated coefficient of $\Delta^2 m_t$ is much closer to the values presented in columns 3 and 4 than those in column 1. This supports the arguments we presented in the previous section. That is, even though the relationship is overdifferenced, as evidenced by the negative autocorrelation in the residuals and the increased residual variance, the estimates of the regression coefficient and its standard error are more in accord with the seemingly well-specified models (columns 3 and 4) than are the results from the underdifferenced model (column 1).

3.2 The quantity theory of sunspots

Now let us consider another model which hypothesizes that income is related to sunspots. Although most economists would dismiss such a theory out of hand, variants on this theme have been taken quite seriously by some,

[16]This is analogous to the findings of Gould and Nelson (1974) that there is no significant negative drift in the random walk model for the log of velocity.

[17]Pierce (1975) discusses some related limitations of conventional goodness-of-fit measures, and Pierce (1977b) proposes alternative measures of goodness-of-fit.

Table 2
Regression models of the log of income on the log of accumulated sunspots: 1897–1958.[a]

	Regression model				
Variable	(1) Levels y	(2) Time trend y	(3) Cochrane–Orcutt y	(4) First differences Δy	(5) Second differences $\Delta^2 y$
Constant	6.305 (0.284)	66.66 (9.35)	19.10 (38.07)	0.043 (0.017)	−0.002 (0.016)
s	0.717 (0.042)	0.140 (0.080)	0.124 (0.146)		
Δs				0.140 (0.124)	
$\Delta^2 s$					−0.049 (0.202)
t		0.040 (0.005)			
$\hat{\rho}_1$[b]			0.995 (0.025)		
$\bar{R}^2$	0.827	0.914	0.986	0.005	−0.016
$\hat{\sigma}_e^2$	0.1435	0.0718	0.0118	0.0116	0.0144
r_1	0.90 (0.13)	0.91 (0.13)	0.38 (0.13)	0.37 (0.13)	−0.21 (0.13)

Note: For footnotes see table 1.

including Jevons (1884). We might refer to this theory as the 'quantity theory of sunspots'. In particular, we are concerned with the relationship between the log nominal income, y_t, and the log of accumulated sunspots, s_t.[18] One way of proceeding would be to estimate the levels regression of y_t on s_t. In column 1 of table 2, we present the results of this calculation. The results of this levels regression are striking at first glance. The coefficient of s_t has a t-statistic of 17.1 and is significantly different from zero at the 1% level using the usual test procedure. The adjusted R^2 is 0.827, which suggests that the log of accumulated sunspots explains over 82% of the variance in the log of nominal income. Closer inspection, however, reveals that the residuals in this regression are highly autocorrelated ($r_1 = 0.90$). In the second column we include a time trend variable and find that the coefficient of s_t falls substantially, but it is significant at the 10% level based on a conventional t-test. The time trend coefficient is positive and has a very large t-statistic, indicating that the trend in both income and accumulated sunspots may be

[18]Another way of stating this theory is that the rate of growth of income is affected by the level of sunspots. Therefore, the level of income is determined by total sunspots since 1897. The data are monthly averages for the year and are taken from Jacobs (1960).

an important omitted variable in the levels regression; however, the residuals from the time trend model are highly autocorrelated.

The Cochrane–Orcutt model estimates are given in column 3. Note that $\hat{\rho}_1$ is very near 1 and that the coefficient of s_t is reduced, and no longer different from zero at the 10% level of significance.[19] Similar conclusions are reached about the effect of accumulated sunspots on income from the first differences regression of Δy_t on Δs_t (column 4), or the second differences regression of $\Delta^2 y_t$ on $\Delta^2 s_t$ (column 5), where the latter model is probably an over-differenced model. Therefore, it is once again apparent that conclusions drawn from a model with residuals that behave like a random walk are often far more misleading than conclusions drawn from a model with residuals which display the characteristics of overdifferencing.

3.3. *The quantity theory versus the income expenditure theory*

In order to emphasize the importance of these points, we present one more set of results. These examples are motivated by the simple models analyzed in Friedman and Meiselman (1963). Our purpose here is not to defend or attack the methodology or data used by Friedman and Meiselman. Rather, we wish to reiterate that the qualitative as well as quantitative results of empirical work can be influenced by the error structure, and that over-differencing is not a serious problem if the linear regression model is correctly specified.

For the purpose of this illustration, we accept the simple dichotomy of the quantity theory and the income–expenditure theory as put forward by Friedman and Meiselman. One aspect of their analysis involves regressing consumption on money and consumption on autonomous expenditures. They report the results of these regressions for the period 1897–1958, as well as others for different subperiods, and conclude that the quantity theory model describes the time series behavior of aggregate consumption more accurately than a simple autonomous expenditure model.

Tables 3 and 4 present estimates of the regression of the log of consumption, c_t, on the log of money, m_t, and the regression of the log of consumption on the log of autonomous expenditures, a_t, respectively. The data provided in Friedman and Meiselman (1963) are the basis for these regressions, but the natural logarithms of the data are used in order to reduce heteroscedasticity in the residuals. In the context of the simple models set forth by Friedman and Meiselman, the simple quantity theory suggests that the coefficient relating c_t and m_t should be near unity. Similarly, the

[19]It is also interesting to note that some first-order residual autocorrelation remains in this model. This suggests the presence of additional first-order moving average properties as yet unaccounted for. Including a moving average parameter does not alter the basic results. The coefficient of s_t remains virtually the same and insignificant, but the variance of the residuals is slightly reduced.

Table 3

Regression models of the log of consumption expenditures on the log of money: 1897–1958.[a]

	Regression model				
Variable	(1) Levels c	(2) Time trend c	(3) Cochrane–Orcutt c	(4) First differences Δc	(5) Second differences $\Delta^2 c$
Constant	2.138 (0.128)	15.75 (6.12)	1.571 (0.702)	0.002 (0.008)	0.001 (0.008)
m	0.843 (0.012)	0.971 (0.059)	0.896 (0.062)		
Δm				0.833 (0.083)	
$\Delta^2 m$					0.862 (0.123)
t		−0.008 (0.003)			
$\hat{\rho}_1$[b]			0.903 (0.060)		
$\bar{R}^2$	0.988	0.988	0.998	0.623	0.449
$\hat{\sigma}_e^2$	0.0103	0.0096	0.0021	0.0022	0.0035
r_1	0.86 (0.13)	0.83 (0.13)	0.22 (0.13)	0.17 (0.13)	−0.33 (0.13)

Note: For footnotes see table 1.

simple income–expenditure model suggests that the coefficient relating c_t and a_t should also be near unity.[20]

The first column of table 3 presents the results of the levels regression of c_t on m_t. The estimated coefficient is close to one (0.843), but significantly different from one based on usual significance tests. However, as in the previous examples, the residuals from this regression are highly autocorrelated. The first column in table 4 contains the estimate of the levels regression of c_t on a_t, representing the income–expenditure theory. In this case the estimated coefficient (1.085) seems statistically indistinguishable from unity at the usual significance levels. One might interpret these results as suggesting that both theories appear to be reasonably well supported by the data, although the estimated residual variance ($\hat{\sigma}_e^2$) is lower, and hence the

[20]Although Friedman and Meiselman use the simple linear regression formulation in relating both consumption with money and consumption with autonomous expenditures, one can justify the log linear form if it is assumed that average and marginal velocities are identical for the quantity theory and that average and marginal multipliers are identical for the income–expenditure theory. Under these circumstances the log linear form arisesquite naturally with unitary regression coefficients implied for both models.

Table 4

Regression models of the log of consumption expenditures on the log of autonomous expenditures: 1897–1958.[a]

	Regression model				
Variable	(1) Levels c	(2) Time trend c	(3) Cochrane–Orcutt c	(4) First differences Δc	(5) Second differences $\Delta^2 c$
Constant	0.374 (1.49)	−75.80 (2.94)	16.92 (10.75)	0.051 (0.008)	0.4×10^{-4} (0.004)
a	1.085 (0.152)	0.280 (0.053)	0.140 (0.032)		
Δa				0.139 (0.032)	
Δa^2					0.088 (0.029)
t		0.044 (0.002)			
$\hat{\rho}_1$[b]			0.993 (0.010)		
$\bar{R}^2$	0.449	0.955	0.955	0.234	0.123
$\hat{\sigma}_e^2$	0.4550	0.0368	0.0044	0.0044	0.0056
r_1	0.89 (0.13)	0.91 (0.13)	0.33 (0.13)	0.33 (0.13)	−0.32 (0.13)

Note: For footnotes see table 1.

$\bar{R}^2$ is higher, for the quantity theory model. However, residual autocorrelation is present in both models and may be distorting the analysis.

The estimates of the time trend models in column 2 of tables 3 and 4 seem to indicate that an exogenous time trend in the log of consumption affects the estimates of the levels regressions in column 1 of both tables. The coefficients of the trend variable are significant by usual standards in both time trend regressions, and the coefficients on m_t and a_t change substantially when the trend variable is added to the levels regression. The coefficient of m_t in column 2 of table 3 increases to become quite close to unity (0.971), and the coefficient of a_t in column 2 of table 4 drops substantially below unity (0.280). Although this evidence seems to favor the quantity theory versus the income–expenditure theory, the residuals from both time trend models are highly autocorrelated indicating the need for further analysis.

If we take the residual autocorrelation into account some rather striking changes occur. Consider the quantity theory Cochrane–Orcutt regression in column 3 of table 3. The coefficient of m_t changes only slightly relative to the levels regression in column 1, but the estimate of the standard error of this parameter is larger, so it is no longer significantly different from one at the

5% level. This result can be contrasted with the estimate of the income–expenditure Cochrane–Orcutt regression in column 3 of table 4. The estimate of the autoregressive parameter, $\hat{\rho}_1$, is 0.993, which strongly suggests that differencing is appropriate. Furthermore, and more important for an economic interpretation of the results, the coefficient of autonomous expenditures falls from 1.085 for the levels regression to 0.140. Although the latter figure is different from zero at the 5% level of significance, it is much less than one. Thus, the Cochrane–Orcutt regression further reinforces the suspicions about the income–expenditure regression that were raised by the time trend regression. The implication is that a 1% change in the rate of growth of autonomous expenditures is associated with a 0.14% change in the rate of growth of consumption in the same year on average.

The first differences regressions are similar to the Cochrane–Orcutt regressions in both tables 3 and 4. The slope coefficient estimates, their standard errors, the residual variances and the residual autocorrelations are all very close for these two estimation techniques, as would be expected since $\hat{\rho}_1$ is close to unity in the Cochrane–Orcutt regressions. It is interesting to remember that the constant term in the first differences regressions corresponds to the time trend parameter in the time trend regressions, but since the first differences models have serially uncorrelated residuals we can use the estimated standard errors of the coefficients to construct proper tests of the exogenous time trend hypothesis. The estimate of the constant in column 4 of table 3 is not significantly different from zero, indicating that there is no significant exogenous time trend in the log of consumption apart from the quantity theory model. On the other hand, the estimate of the constant in column 4 of table 4 is significantly positive and quite close to the estimate of the time trend coefficient in column 2, indicating that there is a significant exogenous time trend in the log of consumption beyond what is predicted by the income–expenditure theory.

Finally, our inferences about the economic significance of the models are not affected if we consider the second differences regression, which represents the case of overdifferencing if the first differences regression is the correctly specified model. Again, overdifferencing is indicated in column 5 of both tables 3 and 4. The residual variances increase for the second differences regressions (column 5) relative to the first differences regressions (column 4), and there is negative residual autocorrelation at lag 1 in both second differences regressions.

We hasten to point out that we do not regard these results as conclusive evidence for accepting or rejecting either the quantity theory or the income–expenditure theory.[21] Our purpose is simply to demonstrate how misleading

[21] As mentioned previously, we do not wish to become involved in the methodological issues surrounding the validity of the Friedman-Meiselman approach to model comparison [cf. Geisel (1975)]. In addition, because of the illustrative nature of our objectives, we have taken some

the parameter estimates can be when the effects of underdifferencing (nonstationary disturbances) are ignored.

The examples in this section make it clear that differencing does not affect the values of the regression coefficients in a correctly specified linear regression model. By examining the behavior of regression residuals it is possible to detect underdifferencing, or overdifferencing, if it occurs. In addition, ignoring the effects of underdifferencing can be a far more costly error than ignoring the effects of overdifferencing.

4. Specification issues and differencing

Thus far, we have focused on the effects of differencing on the error structure in a correctly specified linear regression model. In this section we turn our attention to the effects of differencing on some particular types of specification errors.

4.1. Measurement error

Many arguments against differencing are related to the measurement error problem. These arguments seem to be based on the notion that economic relationships are more or less deterministic, but we only observe the relevant variables with added noise or measurement error. Therefore, it is argued, the 'systematic' portions of the variables are of critical importance and the effect of differencing is to eliminate these systematic movements, or at least reduce the importance of such movements relative to the measurement error. The arguments can be put more precisely by considering a textbook case of the 'errors-in-variables' model. Suppose the regression model under consideration is

$$y_t = \beta x_t^* + \varepsilon_t, \tag{4.1}$$

but the exogenous variable is measured with error so that the model which can be estimated is

$$y_t = \beta x_t + w_t, \tag{4.2}$$

liberties with the data on autonomous expenditures provided in Friedman and Meiselman. In particular, because their time series on autonomous expenditures contains several negative values, we add a constant to each observation so that the logs of the data can be used. This obviously distorts the analysis of the income–expenditure model and, therefore, we make no pretense that we are conducting a viable comparison of the two theories. However, it is interesting to note that if one carries out the income–expenditure regressions using the raw data, ignoring heteroscedasticity, the qualitative results are the same. That is, the estimated coefficient on autonomous expenditures (the estimated multiplier) changes from 5.16, with a standard error of 0.578, in the levels regression to 0.122, with a standard error of 0.166, in the first differences regression.

where

$$x_t = x_t^* + u_t, \tag{4.3}$$

and

$$w_t = \varepsilon_t - \beta u_t, \tag{4.4}$$

If ε_t and u_t are each serially independent, identically distributed random variables with $\text{cov}(\varepsilon_t, u_t) = 0$, then it is well known that the probability limit of the least squares estimator obtained from equation (4.2) is

$$\text{plim}\, \hat{\beta} = \beta \left[\frac{1}{1 + (\sigma_u^2 / \sigma_{x*}^2)} \right], \tag{4.5}$$

where we assume that x_t^* is stationary, so that σ_{x*}^2 exists. As long as the variance of u_t is nonzero, this estimator is an inconsistent estimator for β.

Now suppose the differences of y_t and x_t are used to estimate β,

$$\Delta y_t = \beta \Delta x_t + \Delta w_t. \tag{4.6}$$

It is straightforward to show that the probability limit of the least squares estimator obtained from (4.6) is:

$$\text{plim}\, \hat{\beta} = \beta \left[\frac{1}{1 + (\sigma_u^2 / \sigma_{x*}^2 (1 - \rho_1))} \right], \tag{4.7}$$

where ρ_1 is the first-order serial correlation coefficient for the unobserved regressor x_t^*. This estimator is also inconsistent, and if the unobserved regressor is positively autocorrelated the probability limit of the slope coefficient is closer to the true value in the levels regression than in the differences regression. However, if $\rho_1 < 0$ the opposite is true. These results illustrate that differencing can increase the amount of inconsistency of least squares estimators under some, but not all conditions.[22]

It is important to make two additional observations. First, the least squares estimators are inconsistent for *both* the differenced and undifferenced regressions. The real problem here is not differencing, it is measurement

[22]Another example where differencing does not necessarily increase the inconsistency of least squares estimators can be found in the multiple regression case, where the covariances among the regressors will also be relevant, so that it would be difficult to predict a priori whether the levels or the differences would be more inconsistent. Jacobs (1976) notes that there is one case where differencing leads to inconsistent estimators while the levels estimator is consistent. If there is only one regressor which follows a linear time trend, but it is measured with error, the levels estimator is consistent but the differences estimator is not.

error, and the need for finding a consistent, efficient estimator for either one of the regressions.

The second point involves the nature of the measurement error. In the previous example it is assumed that the measurement error in each period is independent of the other periods (that is, serially uncorrelated). If u_t is autocorrelated, the effect of differencing on the inconsistency of the differences regression relative to the levels regression is even more ambiguous. For example, if the first-order autocorrelation coefficient for the measurement error is ρ_1^*, then (4.7) becomes

$$\operatorname{plim} \hat{\beta} = \beta\left[\frac{1}{1+(\sigma_u^2(1-\rho_1^*)/\sigma_{x^*}^2(1-\rho_1)}\right]. \tag{4.8}$$

and depending on the relative magnitudes of ρ_1^* and ρ_1, the inconsistency of the differences regression may be greater than or less than for the levels regression.

In fact, one might even hypothesize that the measurement error associated with x_t^* follows a random walk. For example, suppose that data on the capital stock in any time period are actually calculated from investment figures (that is, changes in the capital stock). If there is serially independent measurement error in the investment series, the measurement error in the capital stock data would follow a random walk. Under these circumstances, differencing a linear regression which contains the capital stock as the independent variable might be preferable in the sense that the resulting estimator is 'less inconsistent' than the levels regression estimator if $\rho_1^* > \rho_1$.

Based on these considerations, we think that arguments which rely on measurement error as a basis for suggesting that differencing is undesirable are irrelevant. Such arguments critically depend on the nature of the data and the form of the measurement error. If the measurement error is accumulated due to the way the data is compiled, (as in the capital stock example), differencing might be preferred. However, since the effect of differencing on measurement error cannot be known in general, efforts should be centered on obtaining consistent estimators and not on the question of differencing.

4.2. Long-run vs. short-run

Many economic relationships are expected to hold in the 'long run', and not necessarily in the 'short run'. This notion can be expressed in the frequency domain by saying that these economic relationships are low frequency relationships. It is well known [cf., Anderson (1971) pp. 411–414] that the effect of differencing is to emphasize high frequencies and filter out low frequencies. Therefore, differencing might be considered counter-productive for estimating long-run or low frequency relationships.

This type of argument has been used more often in recent years as economists have taken to temporal disaggregation in their search for larger samples. For example, quarterly or monthly data, rather than annual data, are often used to estimate econometric models. One disturbing result of this temporal disaggregation has been that when these economic relationships are investigated using the changes in the variables of interest it is sometimes difficult to ascertain any identifiable relationship between the variables.[23] On the other hand, when the same regression is estimated between the levels of the variables, the 'relationship' appears to be significant and parameter estimates conform to the model builder's expectations. Unfortunately, these levels regressions are often associated with highly autocorrelated errors (which are usually dismissed as being bothersome, but not important enough to affect the magnitude of relationship that has been estimated). Based on the examples presented in section 3, it should be clear that such levels regressions require far more scrutiny.

The problem of estimating 'long-run' relationships involves the temporal specification of the econometric model. The temporal specification of an econometric model refers to the interval over which the data are measured. Ideally, we would want to use the same time unit for which we believe the economic model being tested is valid. For example, suppose we are interested in estimating the relationship between the inflation rate and the rate of growth of the money supply. It is possible that there is intra-year variation in prices which is unrelated to the behavior of the money supply (due to seasonal demand or supply conditions, for example), but that from year to year (or decade to decade) the rate of inflation is roughly proportional to the rate of growth of the money supply. In this case, the data observed monthly would not correspond to the concepts associated with the annual observations. In fact, this is one of the primary motivations behind the use of 'seasonally adjusted' data, although we do not advocate the use of such ad hoc techniques to solve the problem of temporal specification.

Another way to view this problem is in terms of 'permanent' and 'transitory' components of observed economic variables. For example, Friedman's (1957) permanent income hypothesis says that permanent consumption should be related to permanent income, but the transitory components of these variables need not be related. If permanent income grows over time, but transitory income and consumption are stationary random variables, the estimated relationship between observed consumption and observed income may be stronger between annual levels of these variables than between the monthly changes of the variables. In fact, this can be viewed as an example of the 'errors-in-variables' problem discussed above. However, this problem arises because we do not use direct measures of permanent income and consumption to estimate the regression relationship.

[23] Pierce (1977a) provides some examples of this phenomenon.

Based on the examples given above, we believe that many of the problems encountered in attempts to estimate economic relationships using more frequent time series observations arise not from differencing, but from the incorrect specification of the dynamic properties of the model. Although many models incorporate dynamic properties (such as stock adjustment models), they are often ad hoc in their formulation. Consequently, we find it unfortunate that while there is a large literature on the effects of temporal aggregation [cf. Zellner and Montmarquette (1971)] where, for example, the effects of estimating a monthly relationship using quarterly data are investigated, there is relatively little work investigating the problems that might be associated with excessive temporal disaggregation.[24]

Thus, we think that consideration of the temporal specification of the model should not be confused with the question of differencing. If data are measured over intervals which are relevant for the economic theory which underlies the econometric model, it should not matter whether the levels or the differences are used to estimate the regression parameters as long as the disturbances are treated properly. In situations where the levels of the data yield qualitatively different results than the differences, careful consideration should be given to the probability of specification errors in *both* forms of the model.

5. Conclusions

Our purpose in this paper has been to discuss the use of the difference transformation as it relates to econometric model building. We have shown that the parameters of a correctly specified linear regression model can be estimated between the levels of the variables, or the changes of the variables, or the second differences of the variables. We argue that the problem of nonstationary disturbances (possibly in the levels regression) are far more serious than the problems caused by excessive differencing (in the second differences regression, for example). In the underdifferencing case, where the disturbances are nonstationary, regression parameter estimators do not have moments and may be inconsistent. On the other hand, in the overdifferencing case regression parameter estimators are unbiased and consistent, although they are not as efficient as the estimators for the correctly specified model.

We illustrate our arguments using annual data on income, consumption, money, and autonomous expenditures from 1897–1958 obtained from Friedman and Meiselman (1963). We also use annual data on sunspot activity as a possible explanatory variable. Several different techniques are used to estimate the relationships between the logarithms of these variables.

[24]The literature on 'band-spectral regression' could be applied in this case where only low frequency bands are used to estimate the regression function. Engle (1974) discusses the theory of this technique.

In all cases, the regression between the levels of the variables has substantial residual autocorrelation and the exogenous variables are positively autocorrelated, so the estimated standard errors of the regression parameters are substantially downward biased. Including a time trend variable as an additional regressor in the levels regression sometimes affects the value of the regression parameter of interest, but it never eliminates the residual autocorrelation. A Cochrane–Orcutt procedure is used to jointly estimate the regression parameter and an autoregressive parameter for the levels regression residuals, and in all of our examples the autoregressive parameter for the residuals is very close to one (indicating nonstationarity). Thus, the Cochrane–Orcutt procedure suggests that the first differences regression is correctly specified in all of our examples. Indeed, the regression equations between the first differences of the variables yield well behaved residuals in almost all cases.

In the examples where the change in the log of money is the regressor, the estimates of the slope coefficient are similar for the levels, time trend, Cochrane–Orcutt, first differences, second differences regression models. However, when the change in the logs of autonomous expenditures or sunspots is used as the regressor, the estimate of the slope coefficient from the levels regressions is substantially different from the estimates provided by the other methods. Thus, we have illustrated the potential for obtaining 'spurious regressions' between the levels of variables when underdifferencing is a problem, and we have shown that differencing a correctly specified model does not change the nature of the results.

In section 4 we discuss the effects of specification errors on the relationship between levels regressions and differences regressions. We show that, in general, measurement errors in the regressors cause estimation problems for both the levels regression and the differences regression. The problem in this case is to find a consistent, efficient estimator for both the levels and the differences regressions. Similarly, problems associated with the temporal specification of an econometric model require careful consideration of how the data should be measured, and issues related to differencing are peripheral.

Based on this analysis, it is tempting to suggest that differencing could be used as a crude test of model specification. If a model is correctly specified in the levels of the variables, the differences regression should corroborate the levels regression. On the other hand, if errors-in-variables, omitted variables, or time aggregation problems exist in the levels regression, the differences regression may yield substantially different results. Such a finding should lead to a careful analysis of the model's specification and a search for corrective measures, since it is likely that neither the levels nor the differences regressions yield correct results. Thus, we conclude that a careful analysis of econometric models estimated from time series data should include many of the procedures outlined above. The information provided by simultaneous analysis of levels and differences regressions, and the properties of the errors from each, can provide important insights to a careful model builder.

References

Anderson, T.W., 1971, The statistical analysis of time series (Wiley, New York).

Ando, A. and F. Modigliani, 1965, The relative stability of monetary velocity and the investment multiplier, American Economic Review 55, 693–728.

Box, G.E.P. and G.M. Jenkins, 1976, Time series analysis forecasting and control (rev. ed.) (Holden-Day, San Francisco, CA).

Chetty, V.K., 1969, On measuring the nearness of near-money, American Economic Review 59, 270–281.

Chow, G.C., 1966, On the long-run and short-run demand for money, Journal of Political Economy 74, 111–131.

Cochrane, D. and G.H. Orcutt, 1949, Application of least squares regression to relationships containing autocorrelated error terms, Journal of the American Statistical Association 44, 32–61.

Engle, R.F., 1974, Band spectral regression, International Economic Review 15, 1–11.

Fama, E.F., 1975, Short-term interest rates as predictors of inflation, American Economic Review 65, 269–282.

Friedman, M., 1957, A theory of the consumption function (Princeton University Press, Princeton, NJ).

Friedman, M. and D. Meiselman, 1963, The relative stability of velocity and the investment multiplier, in: Commission on money and credit stabilization policies (Prentice-Hall, Englewood Cliffs, NJ) 165–268.

Frisch, R. and F.V. Waugh, 1933, Partial time regressions as compared with individual trends, Econometrica 1, 387–401.

Geisel, M.S., 1973, Bayesian comparisons of simple macroeconomic models, Journal of Money, Credit and Banking 5, 751–772.

Gould, J.P. and C.R. Nelson, 1974, The stochastic structure of the velocity of money, American Economic Review 64, 405–418.

Gould, J.P., M.H. Miller, C.R. Nelson and C.W. Upton, 1978, The stochastic properties of velocity and the quantity theory of money, Journal of Monetary Economics 4, 229–248.

Granger, C.W.J. and P. Newbold, 1974, Spurious regressions in econometrics, Journal of Econometrics 2, 111–120.

Hamburger, H.J. and L.A. Kochin, 1972, Money and stock prices: The channels of influence, Journal of Finance 27, 231–249.

Homa, K.E. and D.M. Jaffee, 1971, The supply of money and common stock prices, Journal of Finance 26, 1045–1066.

Jacobs, G., 1960, Sunspots and the solar influence upon high frequency radio communications, Manuscript, Department of Electrical Engineering, University of Maryland.

Jacobs, R., 1976, Data errors and data differences, Working paper no. 82, Department of Economics, University of California at Los Angeles.

Jevons, W.S., 1884, Commercial crises and sunspots, in: H.S. Foxwell, ed., Investigation in currency and finance (MacMillan, London) 221–243.

Kendall, M.G. and A. Stuart, 1968, The advanced theory of statistics, vol. 3 (2nd ed.) (Hafner, New York).

Keran, M.W., 1971, Expectations, money and the stock market, Federal Reserve Bank of St. Louis Monthly Review 53, no. 1.

Klein, B., 1974, Competitive interest payments on bank deposits and the long-run demand for money, American Economic Review 64, 931–949.

Macaulay, F., 1938, The movements of interest rates, bond yields and stock prices in the U.S. since 1856 (National Bureau of Economic Research, New York).

Maeshiro, A. and S. Vali, 1977, A note on the case of overdifferencing of a regression model, Manuscript, Economics Department, University of Pittsburgh.

Meiselman, D., 1963, Bond yields and the price level: The Gibson paradox regained, in: Deane Carson, ed., Banking and monetary studies (Richard D. Irwin, Homewood, IL) 112–133.
Miller, M.H., 1972, Discussion, Journal of Finance 27, 294–298.
Modigliani, F. and R. Shiller, 1973, Inflation, rational expectations and the term structure of interest rates, Economica 40, 12–43.
Modigliani, F. and R. Sutch, 1966, Innovations in interest rate policy, American Economic Review 56, 178–189.
Modigliani, F. and R. Sutch, 1967, Debt management and the term structure of interest rates, Journal of Political Economy 75, 569–589.
Morishima, M. and M. Saito, 1964, A dynamic analysis of the American economy, 1902–1952, International Economic Review 5, 125–164.
Pierce, D.A., 1975, On trend and autocorrelation, Communications in Statistics 4, 163–175.
Pierce, D.A., 1977a, Relationships – and the lack thereof – between economic time series, with special reference to money, and interest rates, Journal of the American Statistical Association 72, 11–22.
Pierce, D.A., 1977b, R^2 measures for time series, Special studies paper no. 93, Federal Reserve Board, Washington, DC.
Plosser, C.I. and G.W. Schwert, 1977, Estimation of a noninvertible moving average process: The case of overdifferencing, Journal of Econometrics 6, 199–224.
Rowley, J.C.R. and D.A. Wilson, 1973, Quarterly models of wage determination: Sôme new efficient estimates, American Economic Review 63, 380–389.
Shiller, R.J. and J.J. Siegel, 1977, The Gibson paradox and historical movements in real interest rates, Journal of Political Economy 85, 891–908.
Tintner, G., 1940, The variate difference method, Cowles Commission Monograph no. 5 (Cowles, Bloomington, IN).
Vinod, H.D., 1976, Effect of ARMA errors on the significance tests for regression coefficients, Journal of the American Statistical Association 71, 929–933.
Yohe, W.P. and D.S. Karnosky, 1969, Interest rates and price level changes, 1952–1969, Federal Reserve Bank of St. Louis Monthly Review 51, no. 12.
Yule, G.U., 1926, Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series, Journal of the Royal Statistical Society 89, 1–69.
Zellner, A. and C. Montmarquette, 1971, A study of some aspects of temporal aggregation problems in econometric analysis, Review of Economics and Statistics 53, 335–342.
Zellner, A. and C.I. Plosser, 1977, On discriminating between random walk and closely related stochastic processes using Bayesian and non-Bayesian procedures, Manuscript.