# UNIVARIATE DETRENDING METHODS WITH STOCHASTIC TRENDS

### Mark W. WATSON*

*Harvard University and NBER, Cambridge, MA 02138, USA*

This paper discusses detrending economic time series, when the trend is modelled as a stochastic process. It considers unobserved components models in which the observed series is decomposed into a trend (a random walk with drift) and a residual stationary component. Optimal detrending methods are discussed, as well as problems associated with using these detrended data in regression models. The methods are applied to three time series: GNP, disposable income, and consumption expenditures. The detrended data are used to test a version of the Life Cycle consumption model.

## 1. Introduction

Most macroeconomic time series exhibit a clear tendency to grow over time and can be characterized as 'trending'. The statistical theory underlying most modern time series analysis relies on the assumption of covariance stationarity, an assumption that is clearly violated by most macroeconomic time series. In applied econometric work it is usually assumed that this statistical theory can be applied to deviations of the observed time series from their trend value. Since it is often the case that these deviations or economic fluctuations are of primary interest, modern time series techniques are often applied to 'detrended' economic time series.

Much recent work has been devoted to issues involving trends or long-run components in economic series. Some of this work has devoted itself to the proper characterization of 'trends' in economic data. A notable contribution on this topic is the paper by Nelson and Plosser (1982) which considers the use of deterministic and stochastic trends. Other work [e.g., Nelson and Kang (1981, 1984)] has been concerned with the econometric consequences of misspecification in the model for the trend component. Still more work has addressed the issue of detecting long-run relations [e.g. Geweke (1983)] or incorporating long-run relationships in short-run dynamic relations [e.g., the work on 'error correction models' begun in Davidson et al. (1978) and the

co-integrated processes introduced in Granger (1983) and discussed further in Engle and Granger (1984)].

Research concerning the proper characterization of long-run trend behavior in economic time series is important for a variety of reasons. First, the work of Nelson and Kang and others shows that misspecification in the model for the trend can seriously affect the estimated dynamics in an econometric model. Proper estimation of these dynamic relations is important if they are to be used to test modern theories of economic behavior. These theories put very tight restrictions on the dynamic interrelations between economic variables. Misspecification of trend components will often lead the analyst to incorrect inferences concerning the validity of these theories. The proper specification of long-run relations is also critical for long-term forecasting.

This paper makes two points. The first point concerns the theory of stochastic detrending, and the second concerns its empirical implementation. We first propose a method for removing stochastic trends from economic time series. The method is similar to the one originally proposed by Beveridge and Nelson (1981), but differs in one important respect. The problem is cast in an unobserved components framework, in which the trend is modelled as a random walk with drift, and the residual 'cyclical' component is modelled as a covariance stationary process. This framework allows us to discuss and construct optimal detrending methods. One of the optimal methods corresponds to the (negative of the) Beveridge and Nelson transitory component. Our theoretical work shows that, in principle, our unobserved components (UC)[1] representation of the process describing the data will give exactly the same results as the usual ARIMA representation. That is, both models imply exactly the same long-run behavior of the data.

The second point of this paper is that *estimated* ARIMA and UC models imply very different long-run behavior of economic time series. For example, we estimate UC and ARIMA models for post-war US GNP. Both models yield essentially identical values of the likelihood function and short-run forecasts. However, estimated long-run behavior of the models is quite different. The ARIMA model implies that an innovation of one unit in GNP is expected to eventually increase the level of GNP by 1.68 units, the UC model implies that the same innovation will eventually increase GNP by only 0.57 units. These results suggest that it is a dangerous practice to use the estimated time series models such as ARIMA or UC models to make inferences about long-run characteristics of economic time series.

The paper is organized as follows. The next section specifies the model for the observed series, the trend component, and the residual 'cyclical' compo-

---

[1]Unobserved component models of the kind used in this paper have been advocated in Harvey and Todd (1983) and more recently in Harvey (1985). These two references contain excellent discussions of the similarities and differences between dynamic behavior of models parameterized as parsimonious ARIMA models and those parameterized as UC models.

nent. We begin with a general ARIMA formulation for the observed series and describe how the stochastic process can be 'factored' into the two processes for the components. Identification issues are also addressed in this section. Section 3 discusses methods for estimating and eliminating the stochastic trends. The properties of the detrended series are addressed, and special attention is paid to the use of these series in constructing econometric models. The final two sections contain empirical examples. Three post-war US macroeconomic time series – GNP, disposable income, and consumption of non-durables – are analysed. Section 4 looks at each of these series in isolation, and presents and compares various estimates of the detrended series. Section 5 uses the data for non-durable consumption and disposable income to test the 'consumption follows a random walk' implication of one version of the life cycle consumption model. The final section contains a short summary and some concluding remarks.

## 2. The model

This section will introduce three models that additively decompose an observed time series, $x_t$, into a trend and cyclical component. Each of these models will assume, or imply, that the change in $x_t$ is a covariance stationary process. We begin with the Wold representation for the change in $x_t$, which we write as

$$(1 - B)x_t = \delta + \theta^x(B)e_t^x, \qquad \text{var}(e_t^x) = \sigma_{ex}^2, \tag{2.1}$$

where $\theta^x(B)$ is a polynomial in the backshift operator $B$, and $e_t^x$ is white noise. The assumption that $(1 - B)x_t$ is stationary is appropriate for most non-seasonal macroeconomic time series. Seasonal series often require a differencing operator of the form $(1 - B^s)$ where $s$ is the seasonal span (12 for monthly, 4 for quarterly, etc.). Our assumption is not adequate to handle these series, and therefore, we are assuming that the series are non-seasonal.

The representation for the level of the series $x_t$ that we consider in this paper is

$$x_t = \tau_t + c_t, \tag{2.2}$$

where

$$\tau_t = \delta + \tau_{t-1} + e_t^\tau, \qquad \text{var}(e_t^\tau) = \sigma_{e\tau}^2, \tag{2.3}$$

and $(c_t, e_t^\tau)$ is a jointly stationary process. A variety of specific assumptions about this process will be discussed below. The component $\tau_t$ corresponds to the trend component in the variable $x_t$; the 'detrending' will attempt to

eliminate this component. Eq. (2.3) represents this trend as a random walk
with drift, which can be viewed as a flexible linear trend. The linear trend is a
special case of (2.3) and corresponds to the restriction $\sigma^2_{e\tau} = 0$. The forecast
function of (2.3) is linear with a constant slope of $\delta$ and a level that varies with
the realization of $e^\tau_t$. More general formulations are certainly possible. Harvey
and Todd (1983) consider a model in which the drift term, $\delta_t$, evolves as a
random walk. This allows the slope as well as the level of the forecast function
of $\tau_t$ to vary through time. This formulation implies that $x_t$ must be dif-
ferenced twice to induce stationarity, and therefore is ruled out by our
assumption that $(1 - B)x_t$ is stationary.

To complete the specification of the model we must list our assumptions
concerning the covariance properties of $c_t$ and the cross-covariances between
$c_t$ and $e^\tau_t$. We consider three sets of assumptions. The assumptions will differ
in the way that $c_t$ and $e^\tau_{t-k}$ are correlated. We consider each model in turn:

*Model 1*

In this model the component $c_t$ evolves independently of the $\tau_t$ component,
and follows the process

$$c_t = \theta^c(B)e^c_t, \tag{2.4.1}$$

where $e^c_t$ and $e^\tau_{t-k}$ are uncorrelated for all $k$. The parameters of the UC model
given by (2.2), (2.3), and (2.4.1) are econometrically identifiable. To see this,
equate the representations for $(1 - B)x_t$ corresponding to (2.1) and the UC
model given in (2.2), (2.3), and (2.4.1). This implies

$$\theta^x(B)e^x_t = e^\tau + (1 - B)\theta^c(B)e^c_t,$$

so that

$$|\theta^x(1)|^2\sigma^2_{ex} = \sigma^2_{e\tau}. \tag{2.5}$$

The coefficients in $\theta^c(B)$ can be found by forming the factorization of

$$\theta^x(z)\theta^x(z^{-1})\sigma^2_{ex} - \sigma^2_{e\tau} = (1 - z)(1 - z^{-1})\theta^c(z)\theta^c(z^{-1})\sigma^2_{ec}, \tag{2.6}$$

subject to the usual identifying normalizations [e.g., $\theta^c_0 = 1$ and the roots of
$\theta^c(z)$ are on or outside the unit circle].

Eq. (2.6) can be used to show that (2.2)–(2.4.1) place testable restrictions on
the $x_t$ process given in (2.1). To see this, set $z = e^{-i\omega}$, so that the right-hand
side of (2.6) is the spectrum of $(1 - B)c_t$. Since the spectrum is non-negative,
the left-hand side of (2.6) must be non-negative for all $\omega$. This implies that

$\theta^x(e^{-i\omega})\theta^x(e^{i\omega})\sigma_{ex}^2 \geq \sigma_{e\tau}^2$ for all $\omega$, with equality guaranteed at $\omega = 0$ by (2.5). We conclude that the spectrum of $(1 - B)x_t$, $\theta^x(e^{-i\omega})\theta^x(e^{i\omega})\sigma_{ex}^2$, has a global minimum at $\omega = 0$. Only processes with this feature can be represented by Model 1. [As discussed in Nelson and Plosser (1981), this rules out many common processes such as the ARIMA $(1,1,0)$ with positive autoregressive coefficient.] The restrictiveness of this assumption will be investigated empirically for three macroeconomic time series in section 4.

Most of the discussion in the remainder of the paper will be devoted to Model 1. The restrictiveness of the model, however, suggests that other models are needed if the UC model is to be useful in describing all models of the form given in (2.1). Because of this we briefly discuss two other models. These differ from Model 1 in the assumptions they make about the covariance between $c_t$ and $e_{t-k}^\tau$. The first of these is:

## Model 2

The model for $c_t$ is

$$c_t = \theta^\tau(B)e_t^\tau, \qquad (2.4.2)$$

where $\theta^\tau(B)$ is a one-sided polynomial in the backshift operator. In this model, the innovations in the trend and cyclical components are perfectly correlated. The advantage of this model is that there is a one-to-one correspondence between models of the form (2.1) and models characterized by (2.2), (2.3), and (2.4.2). This implies that the parameters of the UC formulation (2.3) and (2.4.2) are econometrically identifiable, and that unlike Model 1, this UC formulation places no constraints on the model (2.1). [A proof of this assertion can be found in an earlier version of this paper, Watson (1985).]

The perfect correlation between the innovations in the components is an assumption that some might find objectionable on *a priori* grounds. Some correlation is needed, however, to give the UC model enough flexibility to capture all of the dynamic behavior possible in the model (2.1). Our final model is a mix of Models 1 and 2, in which the $c_t$ and $\tau_t$ are partially correlated.

## Model 3

The component $c_t$ is represented as

$$c_t = \phi^c(B)e_t^c + \phi^\tau(B)e_t^\tau, \qquad (2.4.3)$$

where $\phi^c(B)$ and $\phi^\tau(B)$ are one-sided polynomials in $B$. This model can be viewed as a mixture of Models 1 and 2. Since both of those models are individually identifiable, Model 3 is not.

## 3. Estimation issues

The models presented in the last section suggest that 'detrending' should be viewed as a method for estimating and removing the component $\tau_t$ from the observed series $x_t$. If we denote the estimated trend by $\hat{\tau}_t$, then the detrended series is given by $\hat{c}_t = x_t - \hat{\tau}_t$. Different detrending methods correspond to different methods for estimating $\tau$. A variety of criteria can be used to choose among competing estimation methods. We will consider linear minimum mean square error (LMSE) estimators constructed using information sets $X_0^h = (x_0, x_1, \ldots, x_h)$. We concentrate on these estimators for a variety of reasons. In addition to the usual reasons, including ease of computation and optimality for quadratic loss, the use of LMSE estimators guarantee certain orthogonality properties involving the estimation errors $\tau_t - \hat{\tau}_t = \hat{c}_t - c_t$. These properties play a key role in the formation of instrumental variable estimators that are discussed below. We concentrate on a univariate information set for computational ease. In general, multivariate methods will produce more accurate estimates. The univariate methods considered in this paper can serve as a benchmark to measure the marginal gains from considering more general, multivariate models.

We will discuss detrending in the context of the general model – Model 3. The results for Model 1 can be found by setting $\phi^\tau(B) = 0$ and $\theta^c(B) = \phi^c(B)$, and the results for Model 2 can be found by setting $\phi^c(B) = 0$ and $\theta^\tau(B) = \phi^\tau(B)$. A convenient starting point for the discussion is the LMSE of using the information set $(\ldots, x_{-1}, x_0, x_1, \ldots)$. In this case the standard Wiener filter for stationary processes [see, e.g., Whittle (1963)] can be extended to this non-stationary case [see Bell (1984)] to yield

$$\hat{\tau}_t = V(B)x_t = \sum v_i x_{t-i}, \tag{3.1}$$

where the coefficients in the two-sided polynomial $V(B)$ can be found from

$$V(x) = \sigma_{e\tau}^2 \left[ 1 + (1 - z^{-1})\phi^\tau(z^{-1}) \right] \left[ \theta^x(z)\theta^x(z^{-1})\sigma_{ex}^2 \right]^{-1}. \tag{3.2}$$

We denote $E(w_t | X_0^h)$ by $w_{t/h}$ for any variable $w$ (where E is used to denote the projection operator). From eq. (3.1),

$$\tau_{t/h} = E\left(\hat{\tau}_t | X_0^h\right) = \sum v_i E\left(x_{t-i} | X_0^h\right), \tag{3.3}$$

so that the estimates of the trend using the information set $X_0^h$ can be formed from the Wiener filter with unknown values of $x$ replaced by forecasts or backcasts constructed from the set $X_0^h$.

The form of the filter $V(B)$ given in (3.2) makes it clear that the LMSE estimate depends on $\phi^\tau(z)$. Different $V(B)$ polynomials will be associated with

Models 1, 2, and 3, so that different LMSE estimates of $c_t$ will be constructed. Eq. (3.2) makes it clear, however, that the difference arises from the way in which future data is used in the construction of $\tau_{t/h}$. All models produce the identical values of $\tau_{t/h}$ for $h \le t$. This is easily demonstrated. From (2.2),

$$x_{t/h} = \tau_{t/h} + c_{t/h}.$$

For $h \le t$,

$$\tau_{t+k/h} = \tau_{t/h} + k\delta \quad \text{for} \quad k = 1, 2, \ldots,$$

so that

$$x_{t+k/h} - k\delta = \tau_{t/h} + c_{t+k/h}.$$

Since all of the models imply that $c_t$ is stationary (with mean zero),

$$\lim(k \to \infty) c_{t+k/h} = 0,$$

so that

$$\lim(k \to \infty)\left[x_{t+k/h} - k\delta\right] = \tau_{t/h}.$$

This result shows that, in principle, the estimates $c_{t/t}$ (which will be called the filtered estimates) can be formed without access to any specialized software. To calculate the filtered estimates one merely constructs an ARIMA model for $x_t$ and then forecasts the series (less the deterministic increases $k\delta$) into the distant future. This forecast corresponds to the filtered estimate, $\tau_{t/t}$, and $c_{t/t} = x_t - \tau_{t/t}$. This estimate of a permanent component was first suggested by Beveridge and Nelson (1981) in their permanent/transitory decomposition of economic time series. They define their transitory component as $\tau_{t/t} - x_t = -c_{t/t}$, in the notation above. This discussion shows that their estimate of the permanent component corresponds to an optimal one-sided estimator for the trend in the models under consideration in this paper.

While the optimal filtered estimate $c_{t/t}$ is identified – is not model-dependent – its precision is not identified. That is, the mean square of $(c_t - c_{t/t})$ will depend on the assumed model. If Model 2 is used to describe the decomposition of the data, then $c_{t/t} = c_t$, so that the mean square error is zero. For the other models, $x_t$ is made up of both noises $e^\tau$ and $e^c$ so that it is impossible to perfectly disentangle $\tau_t$ and $c_t$ when only their sum, $x_t$, is observed. Since the Models 1, 2, and 3 are observationally equivalent, the mean square of $(c_t - c_{t/t})$ is not identified.

The remainder of this section will focus on the use of estimated values of $c_t$ in linear regression models. Replacing $c_t$ by $c_{t/h}$ in regression models leads to

problems similar to those in the classic errors-in-variables model. As the examples below will demonstrate, OLS estimates of regression coefficients will quite often be inconsistent.

We will begin our discussion by writing the orthogonal decomposition of $c_t$ as

$$c_t = c_{t/h} + a_{t/h},$$

where $a_{t/h}$ is the signal extraction error that arises from the use of information set $X_0^h$. Ordinary least square regression estimates rely on sample covariances between observable variables, and the consistency of OLS estimates follows from the consistency of these sample covariances. Consider then the covariance between an arbitrary variable $w$ and $c$. From the decomposition of $c$, we have

$$\text{cov}(w_t c_t) = \text{cov}(w_t c_{t/h}) + \text{cov}(w_t a_{t/h}).$$

The $\text{cov}(w_t c_t)$ will be consistently estimated by the sample covariance between $w_t$ and $c_{t/h}$ if $\text{cov}(w_t a_{t/h}) = 0$. In general this will not be true, so that $\text{cov}(w_t c_t) \neq \text{cov}(w_t c_{t/h})$. Recall, however, that $a_{t/h}$ is a projection error, so that it is uncorrelated with linear combinations of data in $X_0^h$. By constructing $w_t$ as a linear combination of the elements in $X_0^h$, we can guarantee that $\text{cov}(w_t c_t) = \text{cov}(w_t c_{t/h})$. We will use this fact in construction of instrumental variable estimators. It will be convenient to discuss a variety of estimation issues in the context of some specific models.

The first model has current and lagged values of $c_t$ as independent variables, so that

$$y_t = \sum_{i=1}^{k} c_{t-i} \beta_i + z_t' \gamma + \xi_t,$$

where $z_t$ is a $j$ vector of unobserved variables, and we will assume (without loss of generality) that $\xi_t$ is white noise. We can rewrite this model in terms of unobserved variables as

$$y_t = \sum_{i=1}^{k} c_{t-i/h} \beta_i + z_t' \gamma + \xi_t + \sum_{i=1}^{k} a_{t-i/h} \beta_i$$

$$= \sum_{i=1}^{k} c_{t-i/h} \beta_i + z_t' \gamma + u_t,$$

where $u_t$ is the composite error term $(\xi_t + \sum_{i=1}^{k} a_{t-i/h} \beta_i)$. The unknown parameters $\beta_1, \beta_2, \ldots, \beta_k$ and $\gamma_1, \ldots, \gamma_j$ will be consistently estimated by OLS if the regressors $c_{t-i/h}$ and $z_t$ are uncorrelated with the error term $u_t$. This may not be true for a variety of reasons.

Two sources of correlation are immediately apparent. First, consider the correlation between $c_{t-i/h}$ and $\xi_t$. In many models it will be reasonable to assume that $\xi_t$ is uncorrelated with current and lagged values of $c_t$, but unreasonable to assume that $\xi_t$ is uncorrelated with future values of $c_t$. (Correlation between $\xi_t$ and $c_{t+i}$ will exist if there is feedback from $y_t$ to $c_t$.) But when $h > t$, $c_{t-i/h}$ will contain future $x_t$'s, and therefore $c_{t-i/h}$ will contain linear combinations of future $c_t$'s. This may induce a correlation between $c_{t-i/h}$ and $\xi_t$ even though $c_{t-i}$ and $\xi_t$ are uncorrelated. The second source of correlation between the regressors and the disturbance arises from the possible correlation between $z_t$ and $a_{t-i/h}$. These variables will be correlated when $z_t$ contains useful information about $c_t$ not contained in $X_0^h$.

These problems can be circumvented by the use of instrumental variables. In particular, the variables in $X_0^t$ can be used as instruments to estimate the model. When constructing instrumental variable estimates, it is useful to make use of the fact that, for $h > t$, $c_{t/t} = E(c_{t/h} | X_0^t)$, so that IV estimates can be formed by regressing $y_t$ on the filtered values, $c_{t/t}$, and $\hat{z}_t$, the fitted values from the regression of $z_t$ onto the set of instruments. Finally, it should be pointed out that the error terms, $u_t$, may be serially correlated so that standard errors may have to be calculated using the procedures outlined in Cumby, Huizinga and Obstfeld (1983) or Hayashi and Sims (1983).

The second regression model that we consider is

$$c_t = \sum_{i=1}^{k} c_{t-i} \beta_i + z_t' \gamma + \xi_t.$$

This model should be interpreted as the true generating equation for $c_t$, so that $\xi_t$ is uncorrelated with all variables dated $t-1$ or earlier. Models 1–3 described in the last section can be viewed as reduced forms of this model, where the $z_t$'s have been solved out as in Zellner and Palm (1976). Writing the model in terms of observables, we have

$$c_{t/h} = \sum_{i=1}^{k} c_{t-i/h} \beta_i + z_t' \gamma + \xi_t + \sum_{i=1}^{k} a_{t-i/h} \beta_i - a_{t/h}$$

$$= \sum_{i=1}^{k} c_{t-i/h} \beta_i + z_t' \gamma + u_t.$$

As in the previous model, the observed regressors $c_{t-i/h}$ and $z_t$ may be correlated with the error term, $u_t$, leading to inconsistency of the OLS estimates. When $h \geq t$, the estimates $c_{t-i/h}$ contain future $c_t$'s and therefore will be correlated with $\xi_t$. The variables $z_t$ can be viewed as 'causes' of $c_t$ and will therefore contain useful information about the $c$'s, which is not contained in the univariate information set $X_0^h$. This will induce a correlation between $z_t$ and $a_{t-i/h}$. Instrumental variables can again be used to estimate the model. The data in $X_0^{t-1}$ are valid instruments.

In the discussion above, we replaced the true values of $c_{t-i}$ by the estimates $c_{t-i/h}$. Since these estimates depend crucially on the model for $c_t$, a useful alternative is to replace them by the estimates $c_{t-i/t-i}$, which do not depend on the model assumed for $c_t$. The see the implications of this procedure, rewrite the first regression example as

$$y_t = \sum_{i=1}^{k} c_{t-i/t-i}\beta_i + z_t'\gamma - \xi_t + \sum_{i=1}^{k} a_{t-i/t-i}\beta_i.$$

This differs from the formulation above in that each $c_{t-i}$ is estimated using a different information set. Since $\text{cov}(c_{t-i/t-i}a_{t-j/t-j}) \neq 0$ for $i > j$, care must be taken in choosing instruments. Since $X_0^t \supset X_0^{t-1} \supset \cdots \supset X_0^{t-k}$, data in $X_0^{t-k}$ are valid instruments, and this set can be used.

Finally, it is important to keep in mind that the inconsistency in OLS estimates will depend on the magnitude of the error in the estimate of $c_t$. When $c_t$ is estimated very precisely, the inconsistencies from OLS have no practical importance.

## 4. Univariate examples

In this section we will analyze three US macroeconomic time series – real GNP, real disposable income, and real consumption of non-durables. We begin by applying the univariate detrending methods outlined in the last section to the logs of these series, using quarterly data from 1949 through 1984. The estimated univariate models and corresponding trend and cyclical components will be discussed in this section. In the next section, we will investigate the relation between the cyclical components of disposable income and consumption using regression methods.

Two univariate models have been estimated for each series. The first is the usual ARIMA model. The second is an unobserved components (UC) model suggested by the independent trend/cycle decomposition in Model 1. For each time series, we will present and compare the models and their corresponding

trend/cycle decomposition. The analysis in the last section indicated that, *given the Wold decomposition* of the observed series, the optimal one-sided estimates of the components could be formed using any of the observationally equivalent representations of the data. In principle then, it shouldn't matter whether we form the one-sided estimates from the ARIMA model or from the UC model. The results below show that, in practice, it matters a great deal which representation is used. This apparent contradiction arises from the fact that the Wold representation for the data is not known. The ARIMA model and the UC model correspond to different parsimonious approximations to the Wold representation. Given a finite amount of data is is very difficult to discriminate between these alternative representations for the data sets that we consider.

## 4.1. GNP

The estimated autocorrelations for log GNP suggested that the data were non-stationary (the first estimated autocorrelation was 0.98). The correlogram for the change in the series suggested that an ARIMA $(1,1,0)$ model was appropriate. [An ARIMA$(0,1,1)$ was also possible and yields very similar results.] The estimated model was

$$(1 - B)x_t = \underset{(0.001)}{0.005} + \underset{(0.077)}{0.406} (1 - B)x_{t-1},$$

$$SE = 0.0103, \quad L(1) = 0.73, \quad L(3) = 6.3,$$

$$Q(23) = 14.9, \quad LLF = 292.07.$$

*SE* is the estimated standard error, $L(1)$ and $L(3)$ are LM statistics for serial correlation in the error term of order 1 and 3, respectively, and $Q(df)$ is the Box–Pierce statistic of the residuals. Under the null hypothesis of no serial correlation, the LM test statistics are distributed as $X_1^2$ and $X_3^2$, respectively. The final statistic reported, *LLF*, is the value of the log likelihood function.

Interestingly, this estimated model suggests that the spectrum for $(1 - B)x_t$ has a global maximum at the zero frequency, so that decomposition of $(1 - B)x_t$ into an independent random walk and stationary component is not possible. (Recall that this decomposition required that the spectrum had a global minimum at the zero frequency.) This means, as was pointed out in Nelson and Plosser (1981), that the trend/cyclical decomposition in Model 1 is inappropriate. Nevertheless, we estimated a model of the form given into

Model 1.[2] The results were

$$x_t = \tau_t - c_t,$$

$$(1 - B)\tau_t = \underset{(0.001)}{0.008}, \qquad\qquad \sigma = 0.0057,$$

$$c_t = \underset{(0.121)}{1.501}\, c_{t-1} - \underset{(0.125)}{0.577}\, c_{t-2}, \quad \sigma = 0.0076,$$

$$SE = 0.0099, \quad Q(17) = 10.4, \qquad LLF = 294.42.$$

The values for $\sigma$ next to each equation refer to the standard deviation of the disturbance in that equation. The value for $SE$ is the standard error of the innovation in $x_t$, i.e., the one-step-ahead forecast standard error. It is comparable to the $SE$ reported for the ARIMA model.[3]

Both the ARIMA model and the UC model are special cases of the general Wold representation for $(1 - B)x_t$ given in (2.1). In the Wold representation the moving average polynomial $\theta^x(B)$ is, in general, an infinite-degree polynomial. Both the ARIMA and UC models can be viewed as ways of approximating this infinite-degree polynomial. The ARIMA model uses a ratio of finite-order polynomials, while the UC model uses a (restricted) sum of ratios of finite-order polynomials. Presumably neither of these approximations is entirely correct. One may be better at approximating certain characteristics of the $\theta^x(B)$ polynomial and the other may provide a better approximation to other characteristics. With this in mind we will now discuss some of the similarities and differences in the estimated ARIMA and UC models.

The UC model performs slightly better than the ARIMA model in terms of (within-sample) one-step-ahead forecasting, or equivalently, in terms of the value of the likelihood function. Indeed, both models imply very similar behavior for the short-run behavior of the series. To see this, notice that the

---

[2] The UC models were estimated by maximum likelihood methods. Details concerning the methods can be found in Watson and Engle (1983). For those familiar with varying parameter regression (VPR) models, notice that the UC model described in the text can be viewed as a VPR model with serially correlated errors and a time-varying intercept. This allowed us to use previously written VPR computer programs to carry out the estimation.

[3] There a variety of interesting features of the estimated UC model that one might want to test. For example, the sum of the autoregressive coefficients for the cyclical component is 0.92, which implies substantial persistence in $c_t$. Roots this close to unity can be troublesome, and one might want to test the hypothesis that the AR process for $c_t$ contains a unit root (and hence the $c_t$ process was misspecified). In addition, conditional on no unit roots in the $c_t$ process, a zero variance for $e_t^\tau$ implies that the trend is deterministic rather than stochastic. This too would be an interesting hypothesis to test. Unfortunately, standard tests cannot be carried out, as the usual asymptotic normal approximations for test statistics are not valid for the reasons discussed in Fuller (1976) and elsewhere. Construction of test statistics for these sorts of hypotheses is an interesting area for future research.

UC model implies that

$$(1 - 1.501B + 0.577B^2)(1 - B)x_t$$

$$= (1 - 1.501B + 0.577B^2)e_t^\tau + (1 - B)e_t^c. \qquad (4.1)$$

By Granger's Lemma the right-hand side of (4.1) can be represented as a MA(2) and solving for the implied coefficients yields

$$(1 - 1.501B = 0.577B^2)(1 - B)x_t$$

$$= (1 - 1.144B + 0.189B^2)a_t, \qquad \sigma_a = 0.0099.$$

The autoregressive representation for the model is

$$(1 - 1.144B + 0.189B^2)^{-1}(1 - 1.501B + 0.577B^2)(1 - B)x_t = a_t.$$

Carrying out the polynomial division we have, approximately,

$$(1 - 0.36B - 0.05B^2)(1 - B)x_t \approx a_t,$$

which is very close to the estimated ARIMA model.

While the short-run behavior of the UC and ARIMA models are very similar, their long-run properties are quite different. This shows itself in a variety of ways. The spectra implied by the models are quite different at the low frequencies (so that sums of the implied moving average coefficients $\theta_i^x$ are quite different), the models produce markedly different long-run forecasts and give quite different one-sided estimates of the trend and cyclical components. Fig. 1a compares the two implied spectra. The ARIMA model implies a spectra with a maximum at the zero frequency, while the UC model implies a model with a minimum at that frequency. Unfortunately, with only 142 data points there are very few periodogram points corresponding to the low frequencies, so that direct frequency domain estimation methods help little in discriminating between the models.

Many readers may be more comfortable comparing moving average coefficients rather than spectra. In fig.1b, we have plotted the coefficients on the moving average processes for $(1 - B)x_t$ implied by the two models. The plots are very similar for the first few lags; they differ for longer lags. The ARIMA $(1, 1, 0)$ model has moving average coefficients that follow $\theta_i^x = (0.406)^i$, and therefore die off exponentially, but are always positive. The moving average coefficients in the UC model become negative at lag 4 and remain negative until lag 24. The differences in these moving average representations can lead to significantly different conclusions about the long-run behavior of $x_t$. In the
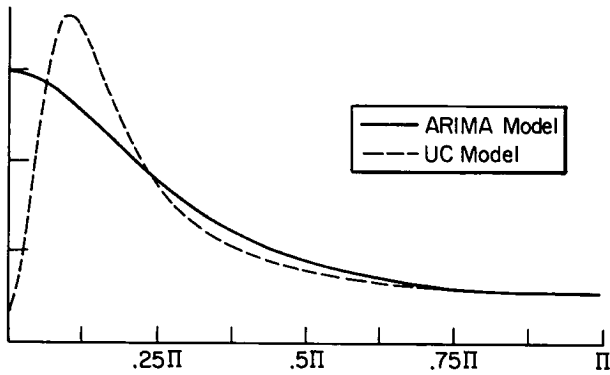
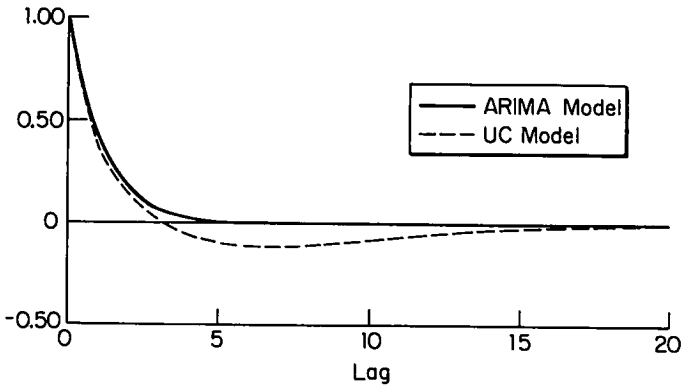Fig. 1a. Spectra for change in log GNP.



Fig. 1b. Moving average coefficients for change in log GNP.

ARIMA model, for example, the sum of the moving average coefficients is 1.68, while in the UC model the sum of the moving average coefficients is 0.59. This means that using the ARIMA model, a one-unit innovation in $x$ will eventually increase log GNP by 1.68, while in the UC model the same innovation is predicted to give rise to a 0.57 increase in GNP. Hypotheses concerning the effects of innovations on permanent income, defined as the discounted expected future sum of $x_t$, are also quite different. The ARIMA model predicts an impact nearly three times as large, and therefore has fundamentally different implications for permanent income. [The relationship between the moving average coefficients in measured income and the permanent income hypothesis is discussed in Deaton (1985).]
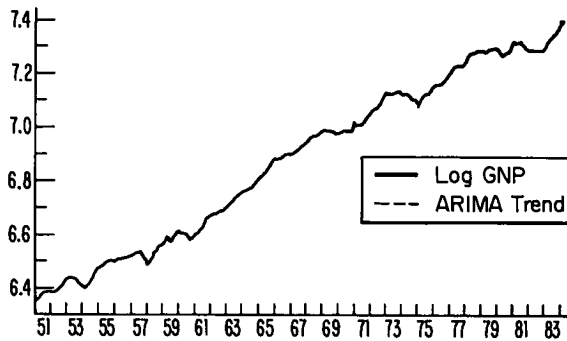
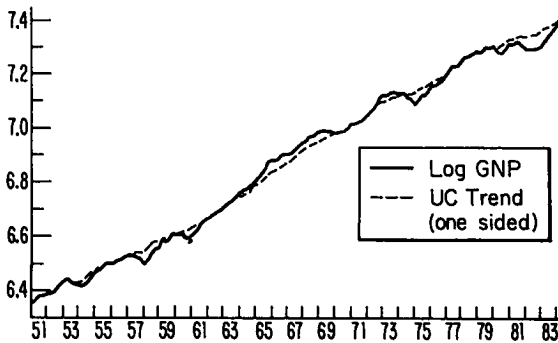Fig. 2a. Trend decomposition for the log of GNP.



Fig. 2b. Trend decomposition for the log of GNP.

In figs. 2a and 2b, we present the actual series and the optimal one-sided estimate of the trend using the two models. The trend in the ARIMA model is very close to the actual series, whereas the trend in the UC model smoothes the series considerably. In fig. 2d we compare the (one-sided) estimates of the cyclical component. Here again the differences are substantial. The estimates constructed from the ARIMA model are difficult to interpret; interestingly, the estimates corresponding to the UC model correspond closely to conventional chronologies of post-war cyclical behavior. This correspondence can be seen in fig. 2e, where we have plotted the UC one-sided estimates, and the shade and peak to trough business cycle periods as calculated by the NBER.

The one-sided estimates of the trend correspond to the long-run forecasts of the series. Since these estimates differ markedly between the ARIMA and UC models, the long-run forecasts are also quite different. One way to discriminate between the competing specifications is to compare their long-run forecasting
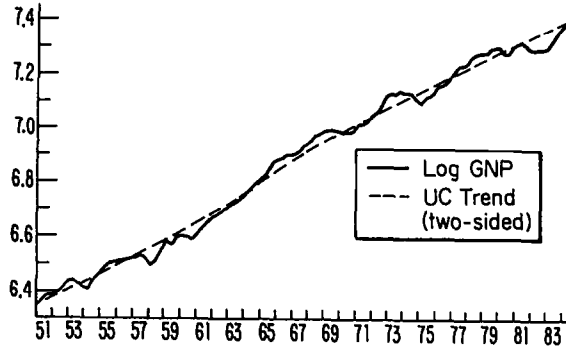
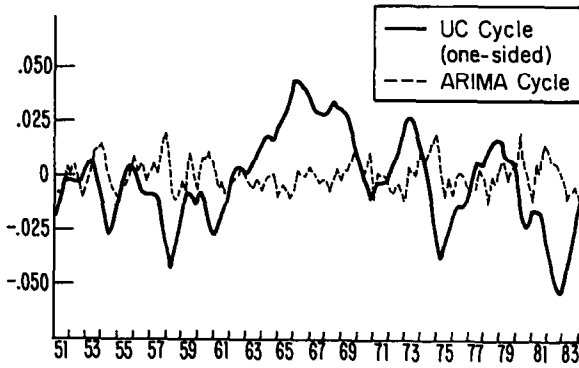Fig. 2c. Trend decomposition for the log of GNP.



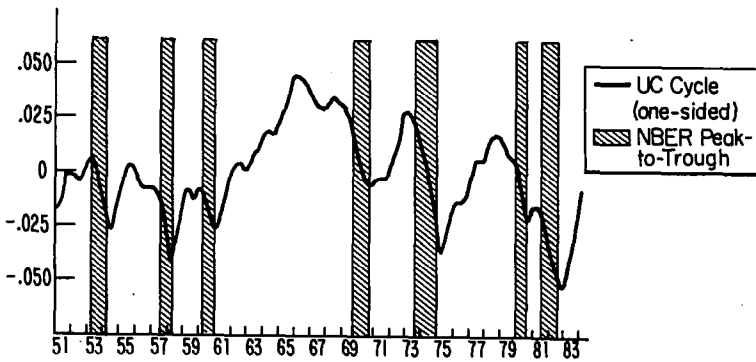Fig. 2d. Cycle decomposition for the log of GNP.



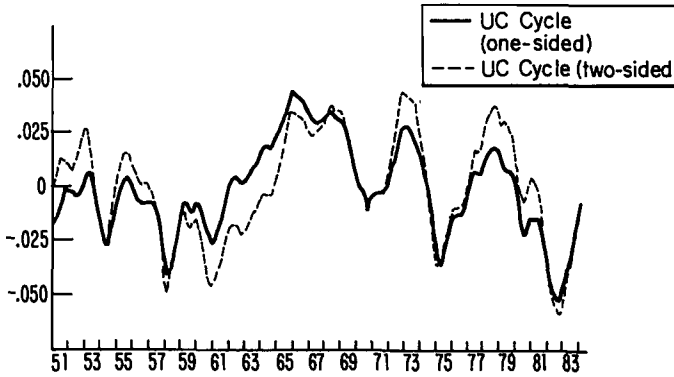Fig. 2e. Cycle decomposition for the log of GNP.

Fig. 2f. Cycle decomposition for the log of GNP.

ability. Within the sample, the UC model produces more accurate forecasts for all forecast horizons up to forty quarters, but within-sample comparisons can be misleading. A true out-of-sample experiment would be much more convincing.

As we pointed out in the last section, conditional on the estimated model for $x_t$, the one-sided LMSE estimates of $c_t$ are unique, i.e., do not depend on the choice of Model 1, 2, or 3. In addition, if we accept Model 2 as the appropriate decomposition, then conditional on the parameters of the ARIMA model, the one-sided estimates are exact, i.e., they have a root mean square error of zero. If, however, we assume that Model 1 or 3 is the correct representation for the cyclical component, then more precise estimates can be constructed. These estimates use future as well as past data to improve the one-sided LMSE estimates. Using the estimated UC model, which assumes that Model 1 is the correct representation for $c_t$, we have constructed the optimal two-sided estimates. The estimate of the trend is plotted in fig. 2c, and fig. 2e compares the optimal one-sided and two-sided estimates of the trend component.

Conditional on the parameter estimates, it is also possible to obtain estimates of the precision of the estimates $c_{t/t}$ and $c_{t/T}$ (the two-sided estimate). Using the estimates from the UC model, the root mean square of $(c_t - c_{t/t}) = 0.020$. The root mean square of $(c_t - c_{t/T})$ depends on the amount of future data available. For $t$ near the middle of the sample, the root mean square of $(c_t - c_{t/T}) = 0.017$. Both estimates are reasonably precise. Using the orthogonal decomposition of $c_t = c_{t/h} + a_{t/h}$, we see that $\mathrm{var}(c_t) = \mathrm{var}(c_{t/h}) + \mathrm{var}(a_{t/h})$, so that a unit-free measure of precision is $R^2(h) = \mathrm{var}(c_{t/h})/\mathrm{var}(c_t)$, which shows the proportion of the variance of $c_t$ explained by $c_{t/h}$. For this model $R^2(t) = 0.54$ and $R^2(T) = 0.71$ (for data near the center of the sample).

## 4.2. Disposable income

The estimated ARIMA model for disposable income was

$$(1 - B)x_t = \underset{(0.001)}{0.011} - \underset{(0.080)}{0.210} (1 - B)x_{t-4},$$

$$SE = 0.010, \quad L(1) = 0.01, \quad L(3) = 1.7,$$

$$Q(23) = 22.9, \quad LLF = 297.1.$$

The corresponding UC model was

$$x_t = \tau_t + c_t,$$

$$(1 - B)\tau_t = \underset{(0.001)}{0.009}, \quad \sigma = 0.0057,$$

$$c_t = \underset{(0.073)}{1.029} c_{t-1} - \underset{(0.094)}{0.024} c_{t-2} + \underset{(0.084)}{0.051} c_{t-3} - \underset{(0.058)}{0.152} c_{t-4}$$

$$+ \underset{(0.017)}{0.055} c_{t-5}, \quad \sigma = 0.0076,$$

$$SE = 0.009, \quad Q(14) = 10.4, \quad LLF = 299.6.$$

We can compare the two models for disposable income, using the same procedures discussed in the comparison of the models for GNP. First, the UC model produces a larger likelihood, but at the cost of six additional parameters. While the models produce similar likelihood values and hence have similar short-run forecasting ability, their long-run forecasts are markedly
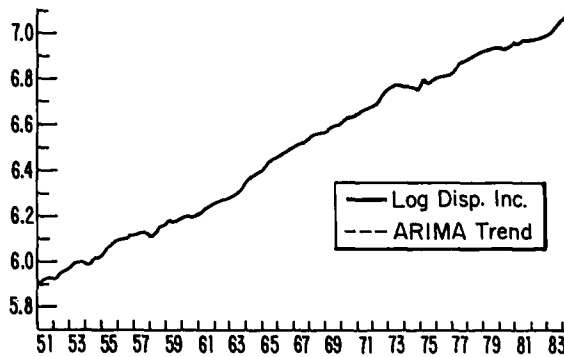


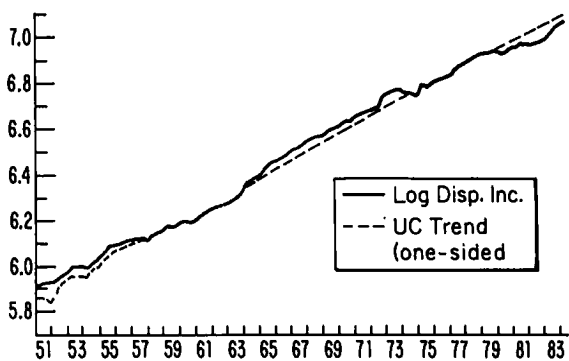Fig. 3a. Trend decomposition for the log of disposable income.

Fig. 3b. Trend decomposition for the log of disposable income.
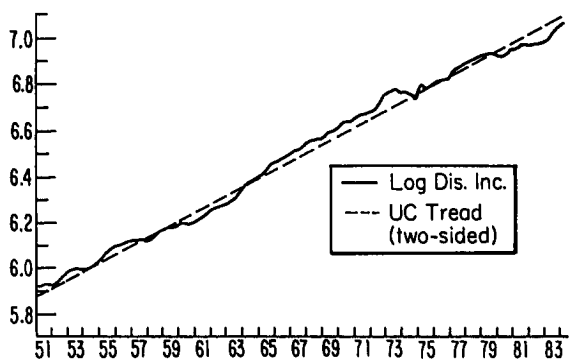


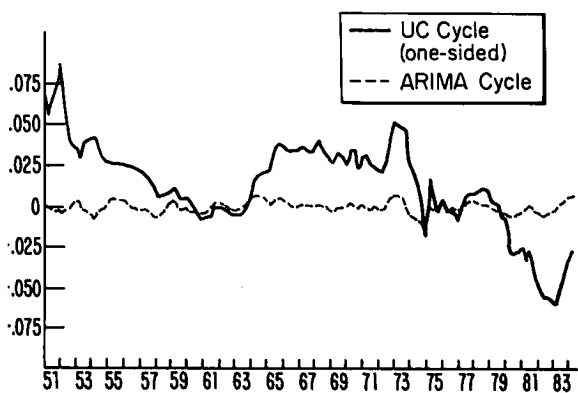Fig. 3c. Trend decomposition for the log of disposable income.



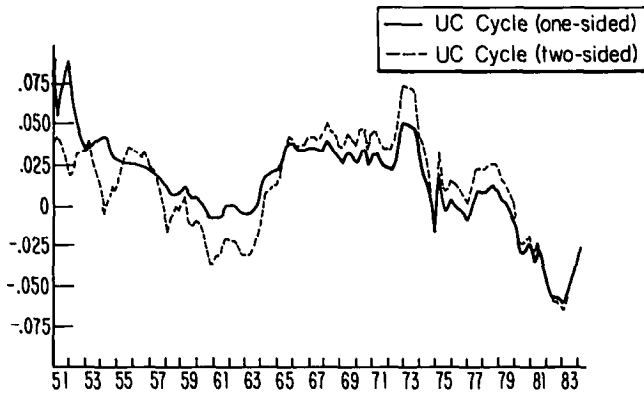Fig. 3d. Cycle decomposition for the log of disposable income.

Fig. 3e. Cycle decomposition for the log of disposable income.

different. These long-run forecasts – the one-sided trend estimates – are compared in figs. 3a and 3b. In fig. 3e, we compare the optimal one-sided and two-sided estimates of the cycle, where the two-sided estimates are calculated assuming that the UC model is correct. The plots are similar for data after 1954, but differ from 1951 to 1954. The one-sided estimates are rather volatile during this early period, reflecting the small number of data points used in their construction. For one-sided estimates constructed with a moderately large amount of data the root mean square of $(c_t - c_{t/t}) = 0.019$. The root mean square of $(c_t - c_{t/T})$ is only slightly smaller.[4] The corresponding $R^2(t)$ is 0.68.

## 4.3. Non-durable consumption

The ARIMA model for non-durable consumption is a random walk. The estimated model and associated statistics are

$$(1 - B)x_t = \underset{(0.0007)}{0.0065},$$

$$SE = 0.0086, \quad L(4) = 1.04, \quad Q(24) = 21.7, \quad LLF = 314.0.$$

[4] Those familiar with recursive signal extraction methods will recognize the mean square of $(c_t - c_{t/t})$ is a time-varying quantity that, for the model under consideration, will converge to a fixed quantity as $t$ grows large. [See Burridge and Wallis (1983).] Since one root of the AR process for $c_t$ is close to the unit circle, and is therefore close to the roots of the AR polynomial for $\tau_t$, the convergence of the mean square of $(c_t - c_{t/t})$ is very slow. The value of the root mean square of $(c_t - c_{t/t})$ reported in the text corresponds to the value for 1984:2. The recursive algorithm, the Kalman filter, was initialized with a vague prior, so that the root mean square errors for earlier dates are larger. For example, the value for 1966 is 0.028. The root mean square of $(c_t - c_{t/T})$ is approximately 0.019 over the entire sample period.

The estimated UC model was

$$x_t = \tau_t + c_t,$$

$$(1 - B)\tau_t = \underset{(0.0003)}{0.0067}, \qquad \sigma = 0.0018,$$

$$c_t = \underset{(0.036)}{0.940} c_{t-1}, \qquad \sigma = 0.0082,$$

$$SE = 0.0085, \quad Q(20) = 14.9, \quad LLF = 314.3.$$

The models are clearly very close to one another in terms of their one-step-ahead forecast ability and likelihood values. If we set the variance of the cyclical component to zero, the UC model implies that $x_t$ is a random walk, so that the random walk model is nested within the UC model. It is therefore possible, in principle, to test the competing models, using a likelihood ratio test. Unfortunately, the test is complicated by the fact that the AR coefficient in the model for $c_t$ is not identified under the random walk hypothesis, but it is identified in the more general model. This complicates the distribution of the likelihood ratio statistic; it will not have the usual asymptotic distribution. This problem has been discussed in detail in Watson and Engle (1985) and Davies (1977). They show that the correct (asymptotic) critical value for carrying out the test (using the square root of the likelihood ratio statistic) is bounded below by the critical value for the standard normal distribution. In this example, the square root of the likelihood ratio statistic is 0.77, which implies a lower bound of the (asymptotic) prob-value of 0.27. This suggests that the random walk hypothesis cannot be rejected at levels of 27% or less.

The examples in this section tell a consistent story. The short-run forecasting performance of the ARIMA and UC models are very similar. At longer forecast horizons, the forecasts from the models differ markedly. This difference in the long-run properties of the estimated models, leads to very different estimates of the underlying trend components and cyclical components.

## 5. Regression examples

In this section we investigate the relationship between non-durable consumption expenditures and the cyclical component of disposable income. We

test the proposition that the change in consumption from period $t$ to $t + 1$ is uncorrelated with the cyclical component of disposable income dated $t$ or earlier. The empirical validity of this proposition, first tested in Hall (1978), has been the subject of ongoing controversy. [See, for example, the papers by Flavin (1981), and Mankiw and Shapiro (1985).] The model that we consider implies that the change in log consumption from period $t$ to period $t + 1$ should be unpredictable using information available at time $t$. We test this implication using the cyclical component of disposable income at time $t$. The importance of using a stationary variable (like the cyclical component of income) rather than a non-stationary variable (like the level of income) has been pointed out by Mankiw and Shapiro (1985). We begin by motivating the empirical specification that is used in this paper.

Assume that a consumer is choosing consumption to maximize a time-separable utility function, subject to an intertemporal budget constraint, i.e., the consumer solves

$$\max_{\{C_i\}} \mathrm{E}\left( \sum_{i=0} (1 + \delta)^{-i} u(C_{t+i}) | I_t \right), \tag{5.1}$$

subject to

$$\mathrm{E}\left( \sum_{i=0} (1 + r)^{-i} C_{t+i} | I_t \right) \le W_t,$$

where $W_t$ is wealth at time $t$ (which includes the expected discounted value of future earning), $\delta$ is a time-invariant subjective discount factor, $I_t$ is the information set at time $t$, and $r$ is the constant one-period interest rate. The first-order conditions for utility maximization imply

$$\mathrm{E}(Z_{t+1} | I_t) = (1 + \delta)(1 + r)^{-1}, \tag{5.2}$$

where $Z_{t+1} = u'(C_{t+1})/u'(C_t)$ is the marginal rate of substitution for consumption between periods $t$ and $t + 1$. An empirical specification follows from an assumption concerning the functional form of the utility function and the probability distribution for $Z_t$. Here, we follow Hansen and Singleton (1983). Let $z_t = \log Z_t$, and assume that $z_{t+1} | I_t \sim \mathrm{N}(\mu_t, \sigma^2)$. The log normality of $Z_{t+1}$ implies that $\mathrm{E}(Z_{t+1} | I_t) = \exp(\mu_t + \sigma^2/2)$. But (5.2) implies that $\mathrm{E}(Z_{t+1} | I_t)$ is constant, so that $\mu_t = \mu$ for all $t$. If we now assume that $u(C)$ is of the constant relative risk aversion form, so that $u(C) = (1 - \beta)^{-1} C^{1-\beta}$, then $z_t = -\beta(c_{t+1} - c_t)$, with $c_t = \log C_t$. When $c_t$ is in the information set $I_t$, this

implies

$$\mathrm{E}(c_{t+1}|I_t) = c_t + \alpha,$$

with $\alpha \approx \beta^{-1}[(\sigma^2/2) + r - \delta]$, so that

$$c_{t+1} = c_t + \alpha + e_{t+1}, \tag{5.3}$$

where $e_{t+1}$ is uncorrelated with information available at time $t$. We will test this proposition by investigating the correlation between $(1 - B)c_t$ and lagged values of the cyclical component of disposable income.

The analysis of the last section casts some light on the hypothesis embodied in (5.3). There we showed that the hypothesis that $c_t$ was a univariate random walk was consistent with the data. We calculated a '$t$-statistic' associated with this hypothesis that had a value of 0.77. In this section we ask whether $(1 - B)c_t$ can be predicted by linear combinations of disposable income. The models that we will estimate in this section all have the form

$$(1 - B)c_t = \alpha + \rho(B)y_{t-1}^c + e_t, \tag{5.4}$$

where $y_t^c$ is the cyclical component in disposable income, and $\rho(B)$ is a one-sided polynomial in $B$. Under the assumption of the life cycle consumption model outlined above, $\rho(B) = 0$ for any choice of a (one-sided) polynomial. Since $y_t^c$ is not directly observed, the model (5.4) cannot be estimated by OLS. We will estimate the model using various proxies for the unobserved $y_t^c$ data. The sample period is 1954:1 to 1984:2.[5]

Before proceeding to a test of this hypothesis using the data described in the last section, one issue concerning the data should be addressed. When this random walk hypothesis is tested using macro data, the consumption and income figures are usually deflated by population before the analysis begins. This expresses all variables in per-capita terms, so that the data are loosely consistent with a representative consumer notion. We have chosen not to follow this course. While we agree that this transformation is useful in principle, in practice it can lead to serious problems. These problems arise from the errors in the quarterly population series. While the underlying trend in the population estimate is probaly close to the trend in the true series, the quarter-to-quarter changes in the estimates series is, most likely, almost entirely noise. Indeed, over 50% of the sample variation in quarterly post-war population growth rates can be attributed to three large outliers in the data

---

[5] We have started the sample period in 1954 to eliminate the observations in which the estimates $y_{t/t}^c$ are very imprecise. See the discussion in footnote 3.

Table 1

Consumption and income regressions.

|  | Equation 1[a] | Equation 2[b] | Equation 3[c] |
|---|---|---|---|
| Constant | 0.007 | 0.007 | 0.008 |
|  | (0.001) | (0.001) | (0.001) |
| $y^c_{t-1}$ | 0.201 | 0.203 | 0.115 |
|  | (0.101) | (0.075) | (0.131) |
| $y^c_{t-2}$ | −0.264 | −0.289 | −0.201 |
|  | (0.151) | (0.111) | (0.169) |
| $y^c_{t-3}$ | 0.231 | 0.108 | 0.479 |
|  | (0.128) | (0.111) | (0.261) |
| $y^c_{t-2}$ | −0.192 | −0.074 | −0.445 |
|  | (0.075) | (0.076) | (0.218) |
| DW | 1.6 | 1.8 | 1.7 |
| F | 2.7 | 3.8 | 2.3 |

[a]$y^c_{t/t}$ used as a proxy for $u^c_t$; coefficients estimated by OLS.
[b]$y^c_{t/T}$ used as a proxy for $y^c_t$; coefficients estimated by OLS.
[c]$y^c_{t/T}$ used as a proxy for $y^c_t$; coefficients estimated by IV.

series.[6] Since our specification is in quarterly changes (for $c_t$) or deviations from a stochastic trends (for $y^c_t$), this increase in the noise-to-signal ratio will lead to serious inconsistencies. Rather than introduce this additional noisy series into the analysis, we will use the raw log differences of the non-durable consumption data and the stochastically detrended estimates of the log of disposable income.

The first model we estimate regresses $(1 - B)c_t$ on four lags of $y^c_{t/t}$ (i.e., $y^c_{t-1/t-1}, y^c_{t-2/t-2}, \ldots, y^c_{t-4/t-4}$). Since each of these constructed variables is a linear combination of data at time $t - 1$ or earlier, the population OLS coefficients should be zero. The results of this regression are shown in the first column of table 1. Individually, the coefficients are large and have *t*-statistics ranging from 1.7 to 2.6. The last entry in the table shows the *F*-statistic, which tests that all of the coefficients are equal to zero. It takes on the value 2.7, larger than the 5% critical value.

In the next column, the results are presented for the same model, but with $y^c_{t/T}$, the two-sided estimates, used as proxies for $y^c_t$. The results are similar.

[6] The quarterly population data (published by the Department of Commerce, Bureau of Economic Analysis) are very close to a deterministic trend. The trend is interrupted by three very large quarterly outliers. Over the period 1948:1 to 1980:4, the average quarterly population growth rate was 0.34%, with a sample standard deviation of 0.12%. The data show a 0.85% population growth rate in 1959:1, a 0.79% growth rate in 1971:4, and a −0.23% growth rate in 1972:1. These three data points are responsible for most of the sample variance in post-war population growth rates. They account for 51% of the sample sum of squares. When these data are used to deflate the consumption and income series used in the regressions, the results are dominated by these three data points.

The $F$-statistic is now 3.8, which is significant at any reasonable level. The results in this column, however, are perfectly consistent with the theory. Recall that the two-sided estimates contain future values of disposable income. Since future values of disposable income are not included in $I_{t-1}$, $(1 - B)c_t$ may be correlated with these variables. (Indeed we would expect innovations in consumption to be correlated with innovations in income.) If we proxy the components $y_t^c$ by the smoothed values and estimate the coefficients using data in $I_{t-1}$ as instruments, then the coefficients should not be significantly different from zero. The results of this exercise are shown in the third column. Only the last coefficient is now significant and the $F$-statistic has fallen to 2.3, significant at the 10% but not the 5% levels.

The results of this section suggest that aggregate post-war US data are not consistent with the life cycle model outlined above.


## 6. Concluding remarks

This paper was motivated by the desire for a flexible method to eliminate trends in economic time series. The method that was developed in this paper was predicated on the assumption that deterministic trend models were too rigid and not appropriate for most economic time series. The alternative – modelling economic time series as non-stationary stochastic processes of the ARIMA class – confused long-run and cyclical movements in the series. The useful, fictitious decomposition of a time series into trend and cyclical components could not be used when modelling series as ARIMA processes. The method described in this paper maintains the convenient trend/cycle decomposition, while allowing flexibility in the models for both of the components.

In addition to discussing a new method for 'detrending' economic time series, this paper makes an important empirical point. The paper compares two empirical approximations to the Wold representation for the changes in GNP, disposable income, and non-durable consumption expenditures. These two empirical approximations correspond to ARIMA and UC models. The models imply very similar short-run behavior for the series: the one-step-ahead forecasts from the models are nearly identical. The implied long-run behavior of the models are quite different. The UC models imply that innovations in the process have a much smaller impact on the long-run level of the series than is implied by the ARIMA model. It is very difficult to discriminate between the competing models on statistical grounds: their log-likelihoods are nearly identical. Since both competing models describe the data equally well, we are left with the conclusion that the data are not very informative about the long-run characteristics of the process. While this may seem to an obvious conclusion, one must keep in mind that very different conclusions would be reached using the implied large-sample confidence intervals constructed from

either the UC or the ARIMA models. The difference arises because these large-sample confidence intervals are conditional on the specific parameterization of the Wold representation.

The paper also discussed the use of stochastically detrended data in the construction of econometric models. Here we demonstrated that care has to be taken to avoid inconsistencies arising from complications similar to errors-in-variables. Our empirical example investigated the relation between the change in consumption of non-durables and lags of the cyclical component of disposable income. Here we found a significant relation, which indicates that the simple life cycle model, with its maintained assumptions of constant discount rates, no liquidity constraints, and time-separable utility, is not consistent with aggregate post-war US data.

## References

Bell, W., 1984, Signal extraction for nonstationary time series, Annals of Statistics 12, 646–684.

Beveridge, S. and C.R. Nelson, 1981, 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 7, 151–174.

Burridge, P. and K.F. Wallis, 1983, Signal extraction in nonstationary time series, University of Warwick working paper no. 234.

Cumby, R.E., J. Huizinga and M. Obstfeld, 1983, Two-step two-stage least squares estimation in models with rational expectations, Journal of Econometrics 21, 333–355.

Davidson, J.E.H., David F. Hendry, Frank Srba and Steven Yeo, 1978, Econometric modelling of the aggregate time-series relationship between consumer's expenditure and income in the United Kingdom, Economic Journal 88, 661–692.

Davies, R.B., 1977, Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 64, 247–254.

Deaton, Angus, 1985, Consumer behavior: Tests of the life cycle model, Invited paper delivered at the World Congress of the Econometric Society, Cambridge, MA.

Flavin, M., 1981, The adjustment of consumption to changing expectations about future income, Journal of Political Economy 89, 974–1009.

Fuller, W.A., 1976, Introduction to statistical time series (Wiley, New York).

Geweke, John, 1983, The superneutrality of money in the United States: An interpretation of the evidence, Carnegie-Mellon University discussion paper, May.

Granger, C.W.J., 1983, Co-integrated and error-correcting models, UCSD discussion paper no. 83-13a.

Granger, C.W.J. and R.F. Engle, 1984, Dynamic model specification with equilibrium constraints: Co-integration and error-correction, UCSD discussion paper, Aug.

Hansen, L. and K. Singleton, 1983, Stochastic consumption, risk aversion, and the temporal behavior of asset returns, Journal of Political Economy 91, 249–265.

Hall, R.E., 1978, The stochastic implications of the life-cycle permanent income hypothesis: Theory and evidence, Journal of Political Economy 86, 971–987.

Harvey, A.C., 1985, Trends and cycles in macroeconomic time series, Journal of Business and Economic Statistics 3, 216–227.

Harvey, A.C. and P.H.J. Todd, 1983, Forecasting economic time series with structural and Box–Jenkins models: A case study (with discussion), Journal of Business and Economic Statistics 1, 299–315.

Hayashi, F. and C. Sims, 1983, Efficient estimation of time series models with predetermined, but not exogenous, instruments, Econometrica 51, 783–798.

Mankiw, N.G. and M.D. Shapiro, 1984, Trends, random walks, and tests of the permanent income hypothesis, Journal of Monetary Economics 16, 141–164.

Nelson, C.R. and H. Kang, 1981, Spurious periodicity in inappropriately detrended time series, Econometrica 49, 741–751.

Nelson, C. R. and H. Kang, 1984, Pitfalls in the use of time as an explanatory variable in regression, Journal of Business and Economic Statistics 2, 73–82.

Nelson, C.R. and C.I. Plosser, 1981, Trends and random walks in macroeconomic time series: Some evidence and implications, Journal of Monetary Economics 10, 139–162.

Watson, M.W., 1985, Univariate detrending methods with stochastic trends, H.I.E.R. discussion paper no. 1158.

Watson, M.W. and R.F. Engle, 1983, Alternative algorithms for the estimation of dynamic factor, MIMIC, and varying coefficient regression models, Journal of Econometrics 23, 485–500.

Watson, M.W. and R.F. Engle, 1985, Testing for regression coefficient stability with a stationary AR(1) alternative, Review of Economics and Statistics LXVII, 341–346.

Whittle, P., 1963, Prediction and regulation (English Universities Press, London).

Zellner, A. and F. Palm, 1976, Time-series analysis and simultaneous equation econometric models, Journal of Econometrics 2, 17–54.