ABSTRACT

SHOCK PERSISTENCE AND STOCHASTIC TRENDS
IN AUSTRALIAN AGGREGATE OUTPUT AND CONSUMPTION

by

Luigi Ermini

Working Paper No. 91-16
August 1989
Revised July 1991

The distinction between transitory and permanent shocks bears important implications for economic analysis in general, and for forecasting and policy issues in particular. This distinction is at the center of the debate on which class of models is best suited to represent economic variables: stationary models around a deterministic trend, or stationary models around a stochastic trend. The debate is here focussed on the Australian case. It is found that both aggregate output and consumption are characterized by stochastic trends, but without a transitory component. This corresponds to a measure of shock persistence equal to one for both variables. For the specific case of aggregate output, this result may be interpreted as indicating the absence of business cycles.
SHOCK PERSISTENCE AND STOCHASTIC TRENDS
IN AUSTRALIAN AGGREGATE OUTPUT AND CONSUMPTION

by

Luigi Ermini

Working Paper No. 91-16
August 1989
Revised July 1991

Abstract
The distinction between transitory and permanent shocks bears important implications for economic analysis in general, and for forecasting and policy issues in particular. This distinction is at the center of the debate on which class of models is best suited to represent economic variables: stationary models around a deterministic trend, or stationary models around a stochastic trend. The debate is here focussed on the Australian case. It is found that both aggregate output and consumption are characterized by stochastic trends, but without a transitory component. This corresponds to a measure of shock persistence equal to one for both variables. For the specific case of aggregate output, this result may be interpreted as indicating the absence of business cycles.

Acknowledgments
Some of the empirical results reported in this paper were presented during a series of lectures on recent developments of time-series analysis given at the department of econometrics of the University of Sydney in the spring 1989. I thank Denzil Fiebig and two anonymous referees for very helpful and pertinent suggestions.
SHOCK PERSISTENCE AND STOCHASTIC TRENDS
IN AUSTRALIAN AGGREGATE OUTPUT AND CONSUMPTION

Luigi Ermini

Department of Economics
University of Hawaii, U.S.A.

August 1989
Revised July 1991

Abstract
The distinction between trend analysis in general, and for
the debate on which class
around a deterministic tre
on the Australian case. It is stochstic trends, but without a tran
persistence equal to one for both variables. For the case of aggregate output, this result may be
interpreted as indicating the absence of business cycles.

Acknowledgments Some of the empirical results reported in this paper were presented during a series
of lectures on recent developments of time-series analysis given at the department of econometrics of
the University of Sydney in the spring 1989. I thank Denzil Fiebig and two anonymous referees for
very helpful and pertinent suggestions.

JEL Classification: no. 211
Keywords: business cycles, stochastic trends, persistence
1. Introduction

Within the modern view, that economies are "driven" by unanticipated shocks, or innovations, an important question is to establish whether the effects of such shocks fade away after a reasonably short period of time (transitory shocks), or persist throughout the future life of the economy (persistent shocks). The distinction is clearly relevant for economic analysis, particularly for forecasting and policy issues. A transitory shock will affect only the short-run behavior of the relevant variable; a persistent shock will affect the long-run behavior as well. Similarly, if the effect is transitory, a shock-generating policy - whether fiscal or monetary - will only have a short term benefit; otherwise, it will affect the economy permanently.

The distinction between transitory and permanent shocks provides the econometrician with a richer range of models available to represent economic variables. Following the conventional practice of decomposing economic variables into a permanent trend component and a transitory stationary component, this distinction is at the center of the current debate on whether economic variables are best represented by stationary processes around a deterministic trend or by stationary processes around a stochastic trend. In the former case, shocks are transitory, as they pertain only to the stationary component of the variable. For example, in case of aggregate output, shock-generating policies would be only policies of output stabilization around the natural rate of growth. In the latter case, shocks are also permanent, as they pertain in various degrees to both the transitory and the permanent component of the relevant variable. Regarding again aggregate output, a shock-generating policy could affect the business cycle - that is, the transitory component of aggregate output - and simultaneously alter the rate of growth. Not surprisingly, the debate about which class of models best represents economic variables is far from settled. For a review, see Watson [1986], Clark [1987], King, Plosser, Stock and Watson [1987], Campbell and Mankiw [1987a, 1987b], Cochrane [1988], and Stock and Watson [1988].

The purpose of this paper is to focus the debate on the Australian case, by concentrating on the measure of shock persistence in gross domestic product (GDP) and private consumption (expenditures on non-durable goods and services). Studies on the dynamic properties of Australian income and consumption are surprisingly rare. For the specific case of consumption, the few reported studies have concentrated only on the question of whether or not Australian consumption follows the permanent income hypothesis (Johnson [1983], McKibbin and Richards [1988]). The main result of this
paper is that both GDP and consumption seem to follow a random walk with deterministic drift. That is, they can be characterized by a stochastic trend with no transitory component. With due caution, this result seems to imply that Australian GDP has no business cycles: external shocks only affect its permanent, or trend, component. Two opposite interpretations could be given to this conclusion: (a) policy shocks do not affect the short-term outlook of the Australian economy, but only its long-term growth (hence, the current debate on government's effectiveness in the short-term may be misplaced); (b) government policies are extremely effective in the short-run, to the point of reducing transitory movements of GDP to a white noise process. Although these results may be viewed inconclusive by some, their implication on the existence of business cycles raises some fundamental questions about policy effectiveness, which have not been fully addressed by the literature.

The paper is organized as follows: Section 2 reviews the theoretical aspects associated with the notion of shock persistence, for three classes of stochastic processes (stationary, integrated, and unobserved component (UC)). Section 3 reports the empirical results, including a test for cointegration between GDP and expenditures on non-durables and services. Section 4 provides some concluding remarks.

2. Measures of persistence

2.1. Stationary processes

It is known from Wold's decomposition theorem that any wide-sense zero-mean stationary process \( \{y_t\} \) can be represented in the form:

\[
y_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}, \quad c_0 = 1,
\]

where \( \epsilon_t \) is a zero-mean white noise series of variance \( \sigma^2_{\epsilon} \); \( \epsilon_t \) is termed the "innovation" or shock in \( y_t \), as it corresponds to that part of \( y_t \) that could not be anticipated on the basis of the (univariate) information set available at time \( t-1 \), \( I_{t-1} = \{\epsilon_{t-j}, j \geq 1\} \); that is, \( \epsilon_t = y_t - E(y_t | I_{t-1}) \), with \( E \) the expectation operator. A compact notation for (1) is \( y_t = C(B) \epsilon_t \), where \( C(B) = \sum_{j=0}^{\infty} c_j B^j \) is a polynomial of infinite order in the backward operator \( B \) (i.e., \( B^j \epsilon_t = \epsilon_{t-j} \)). Stationarity requires \( \sum c_j^2 < \infty \).

The Wold's decomposition theorem essentially states that the current value of a wide-sense stationary process \( y_t \) is made up of a linear combination of all its past shocks, the shock \( \epsilon_{t-k} \) contributing the amount \( c_k \epsilon_{t-k} \). Since stationarity implies (Harvey [1981], Granger and Newbold [1986])
\[
\lim_{j \to \infty} c_j = 0, \quad (2)
\]

the effect of shocks to a wide-sense stationary process fades away with time. Thus, stationarity implies transitory shocks, or, equivalently, implies zero persistence. This property can also be interpreted as "finite memory" (the very distant past does not affect the present), or as "mean reversion": if from time \( t \) on, the process is not subject to further shocks, its value at time \( t+k \) becomes 
\[
y_{t+k} = \sum_{j=0}^{\infty} c_j \, \varepsilon_{t+k-j},
\]
from which, recalling (2), \( \lim_{k \to \infty} y_{t+k} = 0 \). Thus, a stationary process \( y_t \) tends with time to revert towards its unconditional mean value (which is zero).

Consider now the optimal mean-square (univariate) forecast of \( y_{t+k} \), \( f_{t+k} \), based on the univariate information set \( i_t \) (Granger and Newbold [1986]):
\[
f_{t+k} = \sum_{j=0}^{\infty} c_{j+k} \, \varepsilon_{t+j}.
\]

From (2),
\[
\lim_{k \to \infty} f_{t+k} = 0,
\]
that is, the long-run optimal forecast also coincides with the unconditional expectation of the variable. Since under mean-square rule the optimal forecast is the conditional expectation, stationary processes exhibit the property that their conditional mean ultimately reverts to their unconditional mean. An identical conclusion holds if, more generally, the unconditional mean is not zero but a deterministic function of time, \( T_t \), that is 
\[
y_t = T_t + C(B) \varepsilon_t.
\]

The notion of long-run forecast can be used as a criterion to decompose a stochastic process into a permanent and a transitory component: the permanent component coincides with the long-run forecast; the transitory component with its complement (the "detrended" short-run forecast), i.e.
\[
y_t = y^{p}_t + y^{T}_t.
\]

For the class of stationary processes around a deterministic trend, or "trend-stationary" processes, \( y_t = T_t + C(B) \varepsilon_t \), we have
\[
y^{p}_t = T_t \quad (permanent \ component) \quad (4)
\]

---

1 The definition of transitoriness according to (2) is less restrictive than the notion of transitoriness relevant to economic analysis. A shock that fades away in 10 years is technically transitory, according to (2), but essentially permanent to all economic effects.
\[ y_t^r = C(B)e_t \quad \text{(transitory component)} \]

In this case, the permanent component \( y^p_t \) is deterministic by construction, and hence unaffected by the shock \( e_t \).

Aggregate macrovariables such as income and consumption typically exhibit an upward trend. If modelled as trend-stationary processes, their permanent component would be specified as a monotonically increasing deterministic function of time. The detrended series - provided it is stationary - identifies the transitory component, and, for the case of output, is interpreted as describing the business cycle. However, recent literature (particularly, Durlauf and Phillips [1988]) has shown that problems of inference can arise if the detrended series is not stationary.

2.2. Integrated processes

A natural extension of the previous class of processes is obtained by hypothesizing that the trend component is stochastic as well. Central to these stochastic-trend models is the class of "integrated processes". A process \( y_t \) is said to be integrated of order \( d \), \( I(d) \), if it must be differenced \( d \) times in order to become wide-sense stationary. That is, if

\[ (1-B)^d y_t = m + C(B) e_t, \quad (5) \]

with \( e_t \) a zero-mean white noise, \( m \) a constant and \( C(B) \) a polynomial (possibly of infinite order) in the backward operator \( B \), satisfying the stationarity condition. Conventionally, a stationary process is indicated as \( I(0) \), as it does not require any differencing to become stationary \((d=0)\).

As many macrovariables appear to be \( I(1) \) (at least with US data; see, for example, Nelson and Plosser [1982]), this section will concentrate on \( I(1) \) processes, i.e., processes whose first differences \( \Delta y_t = (1-B)y_t \) are stationary:

\[ (1-B)y_t = m + C(B) e_t, \quad (6) \]

Again, \( e_t \) is the innovation of \( y_t \), i.e. \( e_t = y_t - E(y_t | I_{t-1}) \). \( I(1) \) processes are also called difference-stationary.

\( I(1) \) processes exhibit a set of completely different properties from stationary processes. For instance, both their unconditional variance and unconditional mean increase linearly with time at the rate \( \sigma^2 \varepsilon \) and \( m \) respectively (see (7) and (8) below). Regarding shock persistence, the effect of a shock \( \varepsilon_t \) on \( I(1) \) processes persists forever, as opposed to the effect on \( I(0) \) processes which fades away with time. To illustrate, consider the simple case of a random walk with drift, obtained from (6) for \( C(B) = 1 \):
\[ y_t - y_{t-1} = m + \varepsilon_t . \]

On solving this difference equation, with zero initial value without loss of generality, one gets

\[ y_t = m(t-t_o) + \sum_{j=0}^{t-t_o} \varepsilon_{t-j} , \tag{7} \]

which shows that, in addition to a linear time trend, the value of the process is made up of the undiscounted accumulation of all the past shocks. Each shock \( \varepsilon_{t-j} \) contributes its full value, and not its discounted value \( c_j \varepsilon_{t-j} \), with \( c_j \to \infty \) as \( j \to \infty \), as for \( I(0) \) processes. Thus, shocks to a random walk persist forever. Moreover, as the fraction of what each shock contributes to the process at any time is equal to one, one is a natural choice for a measure of random walk persistence.

For the general \( I(1) \) process (6) we have:

\[ y_t = m(t-t_o) + \sum_{j=0}^{t-t_o} C_j \sum_{s=0}^{j} \varepsilon_{t-s} , \tag{8} \]

from which the contribution of each shock \( \varepsilon_{t-j} \) to the current value of the process, \( y_t \), is equal to \( \sum_{s=0}^{j} c_s \), and therefore its long-run persistence is equal to \( \sum_{s=0}^{\infty} c_s = C(1) \).

In fact, \( C(1) \) can be used as a measure of persistence of any class of models. For a random walk, as \( C(B) = 1 \) for all \( B \) by construction, the persistence of one, \( C(1) = 1 \), obtains. For stationary processes, \( C(1) = 0 \) (see, for example, Granger and Newbold [1986]), thus yielding the zero measure of persistence already outlined in the previous section. For a general \( I(1) \) process, \( C(1) \) could be greater or less than one. For example, US monthly consumption is well fitted by the (univariate) first-order moving average model \( \Delta c_t = \varepsilon_t - 0.211 \varepsilon_{t-1} \) (Ermini [1989]), which has a persistence of 0.789: on average, every dollar of unexpected monthly consumption expenditure signals a permanent monthly increase of 79 cents in consumption expenditures for the entire remaining life.

As shocks to integrated processes persist forever, integrated processes are not mean reverting (but their differences are). Further, as their optimal (mean-square, univariate) forecast is

\[ f_{t+k} = y_t + \sum_{j=0}^{\infty} c_{j+k} \varepsilon_{t-j} , \tag{9} \]

the long-run forecast tends toward the current value of the series, and not toward its unconditional mean.
Regarding the decomposition of $I(1)$ processes into a permanent and a transitory component, the permanent component is stochastic, as opposed to the stationary case. For example, consider the case of a first-order integrated, first-order moving average process, or IMA(1,1):

$$(1-B)y_t = (1+cB)\varepsilon_t .$$

From (8), the effect of $\varepsilon_t$ on $y_{t+k}$ is readily seen to be equal to $\varepsilon_t$ for $k = 0$, and equal to $(1+c)\varepsilon_t$ for any $k$ greater than zero. Thus, the shock persistence of the IMA(1,1) follows, with the exception of the current period, the same pattern of the random walk $(1-B)y_t = (1+c)\varepsilon_t$. To emphasize this feature, (10) can be rewritten as

$$(1-B)y_t = (1+c)\varepsilon_t - (1-B)c\varepsilon_t ,$$

which yields a decomposition of the process $y_t$ into a permanent and a transitory component, as in (3), with

$$(1-B)y^p_t = (1+c)\varepsilon_t ,$$

$$y^T_t = -c\varepsilon_t .$$

The permanent component of the IMA(1,1) process is a random walk, and its transitory component is a stationary process (indeed, a white noise). Recalling (9), the decomposition (12) reflects the criterion outlined in the previous section, whereby the permanent component of the IMA(1,1) coincides with its long-run forecast (i.e., $y_t + c\varepsilon_t$), and the transitory component with the detrended or short-run forecast (i.e., $-c\varepsilon_t = y_t - (y_t + c\varepsilon_t)$).

This type of decomposition, whereby the permanent component is always a random walk, is known as the Beveridge-Nelson decomposition (Beveridge and Nelson [1981]). More generally, as for any polynomial $C(B)$ there exists a polynomial $C^*(B)$ such that $C(B) = C(1) + (1-B) C^*(B)$, this type of decomposition takes on the form

$$(1-B)y^p_t = m + C(1)\varepsilon_t ,$$

$$y^T_t = C^*(B)\varepsilon_t .$$

in which the permanent component captures, through $C(1)$, the measure of shock persistence of the overall process $y_t$. For sake of generality, the permanent component may also contain a deterministic part $m$.

This type of decomposition qualifies the class of stationary processes around a stochastic trend as stationarity around a random walk. Note, however, that, as any polynomial $C(B)$ can always be written as
\[ C(B) = C_1(B) + (1-B)C_2(B) \] (14)

with arbitrary choice of \( C_1(B) \) and with \( C_2(B) = [C(B) - C_1(B)]/(1-B) \), there exists an infinite number of ways of decomposing stochastic processes. Which one to choose is still an open question in the literature.

2.3. Unobserved component (UC) processes

In the decomposition (13), both permanent and transitory components are "driven" by the same innovation \( \varepsilon_t \). A natural extension of the stochastic-trend approach is to introduce the possibility of two different innovations driving the two components:

\[ (1-B)y^p_t = m + P(B)v_t \]

\[ y^T_t = T(B)\eta_t \] (15)

with \( v_t \) and \( \eta_t \) white noise processes, in general contemporaneously correlated. As both \( v_t \) and \( \eta_t \) are unobservable, this class of models is referred to as "unobserved component processes". The relationship between the shock \( \varepsilon_t \) to the overall process \( y_t \) and the two component-specific shocks is given by:

\[ C(B)\varepsilon_t = P(B)v_t + (1-B)T(B)\eta_t \] (16)

The Beveridge-Nelson decomposition (13) obtains when \( v_t = \eta_t, P(B) = C(1) \), and \( T(B) = C^*(B) \). When \( v_t = \eta_t \), but \( P(B) \neq C(1) \), decompositions like (14) are obtained. The opposite case of no correlation, \( Ev, \eta = 0 \) is the case most often considered in the literature, and it corresponds to a restricted version of the general model (6). It is shown in Watson [1986], in fact, that the condition \( Ev, \eta = 0 \) imposes the restriction that the spectral density of \( \Delta y_t \) reaches a global minimum at the zero frequency. Hence, under this condition, not all polynomials \( C(B) \) are suited to represent \( y_t \) as in (6). Regarding persistence, Campbell and Mankiw [1987b] show that for UC models with uncorrelated components the measure of persistence \( C(1) \) is at most equal to one.

The measure of persistence \( C(1) \) for UC models with uncorrelated components can be obtained from knowledge of \( P(B) \) and \( T(B) \), and from using Granger's Lemma (Granger and Newbold [1986], p. 29) to solve for the implied moving average coefficients in (16). For example, consider the case of a random walk corrupted by an uncorrelated white noise measurement error, \( y_t = y^*_t + \eta_t \), where \( (1-B)y^*_t = m + v_t \). In the notation of this section, this corresponds to:

\[ y^p_t = y^*_t \]

\[ y^T_t = \eta_t \] (17)
\[ E \nu_t \eta_t = 0. \]

\[ P(B) = 1 \]

\[ T(B) = 1 \]

From (16) it follows that

\[ C(B) \varepsilon_t = v_t + (1-B) \eta_t = (1+cB) \varepsilon_t, \]

that is, the first differences of the measured variable follow an IMA(1,1) process, with autocovariances

\[ R_{\varepsilon \varepsilon}(0) = (1+c^2) \sigma^2_{\varepsilon} = \sigma^2_v + 2\sigma^2_{\eta} \]

\[ R_{\varepsilon \eta}(1) = c \sigma^2_{\varepsilon} = -\sigma^2_{\eta}. \]

Note that the first-lag autocovariance imposes the condition \( c < 0 \). The solution to the equation \((1+c^2) \sigma^2_{\eta} - c(\sigma^2_v + 2\sigma^2_{\eta}) = 0\) is

\[ c = -\left(1 + \frac{1}{2} SN\right) \pm \sqrt{SN \left(1 + \frac{SN}{4}\right)}, \]

where \( SN = \sigma^2_v/\sigma^2_{\eta} \) is the non-negative signal-to-noise ratio. For \( SN \geq 0 \), (20) implies \(-1 \leq c \leq 0\). Correspondingly,

\[ C(1) = 1 + c = \sqrt{SN \left(1 + \frac{SN}{4}\right)} - \frac{1}{2} SN \]

which varies between 0 and 1. When \( SN = 0 \), which entails \( v_t = 0 \), the permanent component of \( y_t \) is zero, and the process is stationary (\( y_t = \eta_t \)); its measure of persistence is thus \( C(1) = 0 \). When \( SN = \infty \), which entails \( \eta_t = 0 \), the transitory component is zero, and the process is a random walk (\( \Delta y_t = v_t \)), uncorrupted by measurement errors; in this case, its measure of persistence is \( C(1) = 1 \).

The range of possible models can be further expanded by introducing higher-order integrated processes, \( I(d) \) with \( d > 1 \). For example, the model presented in Clark [1987], in which the drift of the permanent component in (14) or (15) also follows a random walk, corresponds to an \( I(2) \) integrated process. However, as output and consumption clearly appear to be \( I(1) \) and not \( I(2) \) (see Section 3), this expanded class of models will not be considered in this paper.

Measures of persistence based on \( C(1) \) cannot be compared across different classes of models, for example between an UC model with uncorrelated components and an unrestricted \( I(1) \), as they pertain to different shocks (\( v_t \) for UC processes, and \( \varepsilon_t \) for unrestricted I(1) processes). To overcome this problem, an alternative measure of shock persistence, \( V \), is defined as (Cochrane [1988]):
\[ V = \lim_{k \to \infty} \frac{1}{k+1} \frac{\text{var}(y_{t+k-1}-y_t)}{\text{var}(y_{t+1}-y_t)}. \]  

(21)

As this measure is equivalent to the explained variation of \( \Delta y_t \) due to its permanent component (see also Campbell and Mankiw [1987b]):

\[ V = \frac{\text{var} \Delta y^p_t}{\text{var} \Delta y_t}, \]  

(22)

the relationship between the measure \( C(1) \) and the measure \( V \) is:

\[ V = C^2(1) (1-R^2). \]  

(23)

Under the Beveridge-Nelson decomposition (13), recalling that \( \Delta y^p_t = C(1) \varepsilon_t \), and letting \( 1-R^2 = \sigma^2_d/\text{var} \Delta y_t \), be the unexplained variation of \( \Delta y_t \), (23) is readily obtained from (22). Similarly, in the case of UC models with uncorrelated components, it can be shown that \( \text{var}(\Delta y^p_t) = C^2(1) \sigma^2 \varepsilon \), where \( C(B) \) is now the polynomial in (16). Thus, again (23) follows from (22).

3. The empirical results

The empirical investigation addresses three issues. First, we establish, by means of a unit-root test, whether Australian GDP and total private consumption belong to the class of trend-stationary processes or of difference-stationary processes, i.e. whether they are \( I(0) \) or \( I(1) \) (or perhaps, of higher order). Second, following the analysis of the previous section, we calculate their shock persistence. As the empirical research on aggregate output and consumption is typically based at times on the log values of the series and at times on their linear values, these first two issues are investigated both in log and linear values. The third issue, then, is to establish which of the two representations - in log values or in linear values - is more suited to describe the dynamic properties of GDP and consumption.

3.1 Unit-root tests

The main tool to distinguish an \( I(1) \) process from an \( I(0) \) process is the Dickey-Fuller unit-root test, available in two forms (Dickey and Fuller [1979, 1981]): (i) the simple Dickey-Fuller test (DF), and (ii) the augmented Dickey-Fuller (ADF). DF is essentially a \( t \)-test with non-standard distribution on the coefficient \( \rho \) in the regression:

\[ x_t = \alpha_0 + \alpha_1 t + \rho x_{t-1} + \mu_t. \]  

(24)

\footnote{In place of a formal proof, consider again the example described in (17). We have: \( C^2(1) \sigma^2 \varepsilon = (1+c)^2 \sigma^2 \varepsilon = (1+c^2) \sigma^2 \varepsilon + 2c \sigma^2 \eta \), but from (19), \( (1+c^2) \sigma^2 \varepsilon = \sigma^2 \varepsilon + 2\sigma^2 \eta \), and \( c \sigma^2 \varepsilon = -\sigma^2 \eta \); thus, \( C^2(1) \sigma^2 \varepsilon = \sigma^2 \varepsilon = \text{var}(\Delta y^p_t) \).}
The null $H_0: \rho = 1$ is rejected against the alternative $H_a: \rho < 1$ for high absolute values of the $t$-statistic (the test is a left-tail test). The non-standard distributions for the DF test are found in Dickey and Fuller [1981], and are also available for the case $\alpha_i = 0$. The DF is more powerful if $u_i$ is close to a white noise process. If $u_i$ exhibits serial correlation, the ADF is more powerful. The latter is an $F$-test associated with the regression

$$\Delta x_t = \alpha_0 + \alpha_1 t + \alpha_2 x_{t-1} + \sum_{j=1}^{p} c_j \Delta x_{t-j} + \nu_t . \quad (25)$$

The null $H_3$: $\alpha_1 = \alpha_2 = 0$, or $H_2$: $\alpha_0 = \alpha_1 = \alpha_2 = 0$, are tested against the unrestricted model (25). The non-standard $F$-distributions for the ADF statistic $\Phi_2$ (corresponding to $H_2$) and $\Phi_3$ (corresponding to $H_3$) are also reported in Dickey and Fuller [1981].

Choosing $p = 4$ in (25), and with a sample size of 119 which entails 5%-critical values $\Phi_2^* = 4.88$ and $\Phi_3^* = 6.49$, the ADF statistics for quarterly GDP, $y_t$, and total private consumption, $c_t$, are

<table>
<thead>
<tr>
<th></th>
<th>$\Phi_2$</th>
<th>$\Phi_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_t$</td>
<td>10.14</td>
<td>2.02</td>
</tr>
<tr>
<td>$\ln y_t$</td>
<td>10.01</td>
<td>1.56</td>
</tr>
<tr>
<td>$c_t$</td>
<td>20.11</td>
<td>6.28</td>
</tr>
<tr>
<td>$\ln c_t$</td>
<td>17.12</td>
<td>1.81</td>
</tr>
</tbody>
</table>

from which, as $H_2$ is rejected for all four series but not $H_3$, one can conclude that at the 5% level the four series exhibit unit root with drift. An identical conclusion is obtained with simple DF tests (results not reported). Moreover, the presence of a second unit root is strongly rejected, thus supporting the conjecture that all series are $I(1)$.

However, this conclusion may be incorrect if the relevant series exhibit a significant structural break within the sample. As shown in Perron [1989], in fact, in case of structural breaks the DF and ADF tests are biased towards non-rejection of the unit root hypothesis. Using different non-standard distributions derived for the case when $\alpha_0$ and/or $\alpha_1$ change once within the sample, for example, Perron rejects the unit-root hypothesis for US quarterly real GNP. As both $\ln y_t$ and $\ln c_t$ exhibit a change of slope (see figures 1 and 2), with break-point in 1973 - the same found by Perron with US data - corresponding to the first oil-price shock, a Perron-modified unit-root test is thus performed to account for this structural break.

---

3 Both series are taken from the ABS databank (release June 89), and are from 59-III to 89-I, seasonally adjusted, at constant dollars 84/85.
When the change affects the slope but not the level of the series - which is the relevant case for Australian GDP and consumption - Perron test is a non-standard $t$-test for the hypothesis $\alpha = 1$ against $\alpha < 1$, where $\alpha$ is the coefficient in the regression

$$\bar{x}_t = \alpha \bar{x}_{t-1} + \sum_{j=1}^{p} c_j \Delta \bar{x}_{t-j} + e_t,$$

and where $\bar{x}_t$ is the detrended series from

$$x_t = \alpha_0 + \alpha_1 t + \alpha_2 DT^* + \bar{x}_t,$$

with $DT^* = t - t_b$, and $t_b = 1973$ is the break-point.

This Perron test fails to reject the unit-root hypothesis for all the four series. For example, in case of $\ln y$, the $t$-statistic for $\alpha = 1$ is -1.32, against a 5%-critical value of -3.96 (see table V.B in Perron [1989], for time of break relative to total sample size equal to 0.6). Thus, the previous conclusion that all four series are $I(1)$ with drift receives further support.

Given this conclusion, it is interesting to test GDP and consumption for cointegration, that is for the existence of a stable (stationary) long-run equilibrium relationship. Following Engle and Granger [1987], the two cointegrating equations were estimated ($t$-statistic in parentheses):

$$c_t = 181 + 0.59 y_t + u_t,$$

$$\text{(0.93) (123)}$$

$$\ln c_t = -0.4 + 0.989 \ln y_t + \nu_t,$$

$$\text{(-5.83) (149.8)}$$

The low values of the Durbin-Watson statistic (0.247 for the first regression and 0.26 for the second one) reject the hypothesis of cointegration in both cases at the 1% level (see Table II and III of Engle and Granger [1987]). The same conclusion of non-cointegration is obtained with the augmented Dickey-Fuller test described in Engle and Granger: the $t$-statistic of the coefficient $\phi$ in $\Delta u_t = - \phi u_{t-1} + \sum_{j=1}^{d} \beta_j \Delta u_{t-j}$ is -1.22, well below the critical values up to the 10% level (same tables); the $t$-statistic for the coefficient $\phi$ in a similar equation for $\nu_t$ is -1.28. Note that non-cointegration between Australian aggregate output and consumption was also found by McKibbin and Richards [1988], using a different data set.

3.2 Calculating shock persistence

Having determined that the four series are $I(1)$, the next step is to estimate the drift $m$ and the polynomial $C(B)$ of the parametrization (6). In general, the true polynomial $C(B)$ is of very high order, possibly infinite. For sake of parameter parsimony, it
can be truncated to a polynomial $D(B)$ of order $q$, with $q$ usually small; this would make (6) an integrated-moving average parametrization, or IMA(1,$q$). Or it can be approximated by the ratio $1/A(B)$, with $A(B)$ also of low order $p$; this would make (6) an integrated-autoregressive parametrization, or ARI($p$,1). Or it can be approximated by the ratio $D(B)/A(B)$, which would make (6) an integrated autoregressive-moving average parametrization, or ARIMA($p$,1,$q$). Following the strategy of identification, estimation and diagnostics outlined, for example, in Granger and Newbold [1986], for all the four series the random walk parametrization ($C(B) = 1$) could not be rejected at the 5% level against the set of all nesting ARIMA($p$,1,$q$) alternatives, with $p + q \leq 5$. Moreover, the residuals from the random walk parametrizations passed the appropriate white noise tests.

Although these results should not be taken too literally (a different strategy of model selection, or a different level of confidence could lead to different parametrizations), yet they seem to indicate that the measure of persistence for the four series is very close to one, and that their dynamic properties are satisfactorily replicated by random walks with drifts. One consequence of this conclusion is that the estimation of alternative parametrizations of the type described in section 2.3, such as the unobserved-component models, is unnecessary: the implied restriction of a global minimum at the zero frequency is satisfied by the condition $C(B) = 1$ for all $B$. A more important consequence is that, as random walks are characterized by stochastic trends with no transitory component, this conclusion seems to indicate the absence of business cycles in aggregate output.

The maximum-likelihood estimates of the four random walks are:

\[
\Delta y_t = 366 + e_t, \quad \text{(residuals s.e. = 508.5)}
\]

\[
\Delta \ln y_t = 0.0098 + e_t', \quad \text{(residuals s.e. = 0.013)}
\]

\[
\Delta c_t = 205 + e_t, \quad \text{(residuals s.e. = 200)}
\]

\[
\Delta \ln c_t = 0.0095 + e_t', \quad \text{(residuals s.e. = 0.0085)}
\]

Note that, in line with the life-cycle permanent-income hypothesis of consumer behavior, Australian consumption exhibits lower volatility than income.

3.3 Log values vs. linear values

Although the main result of the paper - both GDP and consumption follow random walks - appears to be robust to whether the series are estimated in log or linear values, both representations cannot be simultaneously true. For example, if the true
data generating mechanism for GDP is the log-normal random walk with drift, with
time-invariant distribution $\Delta \ln y_t \sim N(m, \sigma^2)$ and with initial condition $\mu$, then, as shown
in Ermini and Granger [1991], the linear series $y_t = e^{\ln y_t}$ is also normally distributed,
but with time-varying moments:

$$E(\Delta y_t) = \Psi e^{\lambda t}$$
$$\text{var}(\Delta y_t) = \Phi e^{2\lambda t + \sigma^2 t},$$

with $\lambda = m + \frac{1}{2}\sigma^2$, $\Psi = e^{\rho(1-e^{-\lambda})}$, and $\Phi = e^{2\rho(1-2e^{-(\lambda+\sigma^2)}) + e^{-(2\lambda+\sigma^2)}}$. It follows that, if the
log random walks of (29) and (30) are true, then the corresponding linear random
walks of (28) and (29) are necessarily mis-specified, and vice versa.

The time-varying relations (31) can be used to establish which one of the two
models is "true", via a mis-specification encompassing test described in Ermini and
Hendry [1991]: if the linear random walk indeed exhibits time-varying first and second
moments according to (31), then the log model encompasses the linear model in the
sense that the former correctly predicts the features of the latter, and the latter does not
add any new insight about the underlying unknown data generating process. Regarding
the unconditional mean, this encompassing test entails that in the regression

$$\Delta y_t = \gamma + \rho[e^{\lambda t}] + \eta_t$$

the exponential regressor $e^{\lambda t}$ (constructed with the estimates of the log model) should
be significant, and the constant $\gamma$ insignificant. Without correcting for possible
heteroskedasticity, and substituting from (29) and (30) the values of $(m, \sigma^2) = (0.0098, 0.000169)$ and $(0.0095, 0.0000723)$ for GDP and consumption respectively, the follow-
ing estimates were obtained ($t$-statistic in parentheses):

$$\Delta \hat{y}_t = 143 + 116 \ e^{\lambda t}$$  
(0.97)  (1.59)

$$\Delta \hat{c}_t = 124 + 42.7 \ e^{\lambda t}.$$  
(2.06)  (1.4)

The hypothesis that the log model encompasses the linear model is rejected at the 5% level. In fact, the significance of the exponential regressor and the insignificance of the constant should be tested after correcting the linear model for the type of heteroskedasticity implied by (31). Although the rate of growth of the variance is extremely small for both cases, the following different estimates were obtained after scaling down the relevant variables:

$$\Delta \hat{y}_t = 125 + 134 \ e^{\lambda t}$$  
(0.9)  (1.36)
\[ \Delta \hat{c}_t = 51.3 + 90.4 \ e^{\mu}, \]

(95) (2.28)

from which a reversal of the previous result is obtained for the case of \( \hat{c}_t \). In summary, for the case of consumption, the correction for heteroskedasticity seems to indicate that the log model encompasses the linear model, so that the former can be considered more appropriate to represent this variable. The same result does not hold for the GDP, for which then the log-vs-linear question remains open. Interestingly, Ermini and Hendry [1991] find that for US GNP the log model encompasses the linear model even without the heteroskedasticity correction.

4. Conclusions

This paper presents three sets of empirical results about Australian aggregate output and consumption. First, it is found that, in line with similar results with US data, these two series appear to be integrated of order one, i.e. to belong to the class of difference-stationary processes. Second, it is found that the best - in the maximum likelihood sense - univariate representation for both series appears to be the random walk model with drift, both in log and linear values. From a forecasting point of view, this result entails - with due caution - that the best mean-square forecasts of any future value of the two series are simply their current values. More importantly, for the case of GDP, this result also entails the absence of business cycles in the Australian economy: policy shocks only affect the permanent (stochastic) component of GDP. Even when this result is not taken too literally, it still raises some important questions regarding the current debate about government’s policy effectiveness in the short run, questions that still remain inadequately addressed in the literature. Third, the paper addresses the question of which representation - in log or in linear values - better describes the statistical properties of the two series. It is found that for the case of consumption the log representation seems to fit the data better than the linear representation; however, the same conclusion could not be reached for GDP.
References


AUSTRALIAN CONSUMPTION AND GROSS DOMESTIC PRODUCT (QTR. SA, '04$)

Sample Period is 1959(3) - 1989(1)

Figure 1
Log Australian Consumption and Gross Domestic Product (QTR. SA, '84$)

Sample Period is 1959(3) - 1989(1)

Figure 2
<table>
<thead>
<tr>
<th>Paper Number</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>91-16</td>
<td>Luigi Ermini</td>
<td>Shock Persistence and Stochastic Trends in Australian Aggregate Output and Consumption</td>
</tr>
<tr>
<td>91-15</td>
<td>Byron Gangnes</td>
<td>Macro and Industry Effects of Japanese Investment in U.S. Auto Production</td>
</tr>
<tr>
<td>91-14</td>
<td>Luigi Ermini</td>
<td>Recent Developments in Regional Econometric Modelling</td>
</tr>
<tr>
<td>91-13</td>
<td>Luigi Ermini, Clive W.J. Granger</td>
<td>Some Generalizations on the Algebra of I(1) Processes</td>
</tr>
<tr>
<td>91-12</td>
<td>Sumner J. La Croix</td>
<td>Property Rights and Institutional Change During Australia’s Gold Rush</td>
</tr>
<tr>
<td>91-11</td>
<td>Luigi Ermini, David F. Hendry</td>
<td>Log Income vs. Linear Income: An Application of the Encompassing Principle</td>
</tr>
<tr>
<td>91-10</td>
<td>Sumner J. La Croix, James Mak, Walter Miklius</td>
<td>Evaluation of Alternative Arrangements for the Provision of Airport Taxi Service</td>
</tr>
<tr>
<td>91-9</td>
<td>Calla Wiener, Zhao Xinghan</td>
<td>Closing the Scissors Gap: Reform and the Interface Between Agriculture and Industry</td>
</tr>
<tr>
<td>91-8</td>
<td>Luigi Ermini</td>
<td>Effects of Transitory Consumption and Temporal Aggregation on Tests of the Permanent Income Hypothesis</td>
</tr>
<tr>
<td>91-7</td>
<td>Sumner J. La Croix</td>
<td>An Economic Analysis of Decentralization In China: The Coevolution of Political Order and Property Rights</td>
</tr>
<tr>
<td>91-6 (R90-29)</td>
<td>Sumner J. La Croix, Carl Bonham</td>
<td>Forensic Forecasting: Fact or Fiction?</td>
</tr>
</tbody>
</table>
91-4  Carl Bonham  Correct Cointegration Tests of the Long Run Relationship Between Nominal Interest and Inflation

(R90-26)

91-3  Edwin T. Fujii  Clifford B. Hawley  Empirical Aspects of Self-Employment


91-1  Luigi Ermini  On the Durability of Non-Durable Goods: Some Evidence from US Time-Series Data

90-30  James Mak  Marcia Sakai  Taxation of Foreign Real Property Investments in the U.S.: A State Perspective

90-28  Calla Wiemer  Grain Prices, Land Rent, and Food Self-Sufficiency in China

90-27  Lou Rose  Land Values and Housing Rents in Urban Japan