

On the Welfare Costs of Business Cycles in the 20th Century*

João Victor Issler Afonso Arinos de Mello Franco

Osmani Teixeira de Carvalho Guillén

Graduate School of Economics - EPGE

Getulio Vargas Foundation

Praia de Botafogo 190 s. 1100

Rio de Janeiro, RJ 22253-900

Brazil

E-mail: jissler@fgv.br, amfranco@fgv.br, and osmani@fgvmail.br

JEL Codes: E32; C32; C53.

Feb., 2003.

Abstract

Lucas(1987) has shown a surprising result in business-cycle research, that the welfare cost of business cycles are relatively small. Using standard assumptions on preferences and a reasonable reduced form for consumption, we computed these welfare costs for the pre- and post-WWII era, using three alternative trend-cycle decomposition methods. The post-WWII period is very quiet, with the welfare costs rarely exceeding 1% of consumption. However, for the pre-WWII era this basic result is dramatically altered. For the Beveridge and Nelson decomposition, and reasonable preference parameter and discount values, we get a compensation of about 5% of consumption, which is by all means a sizable welfare cost (about US\$ 1,000.00 a year).

1. Introduction

A main question of economics is whether governments should or should not intervene in markets. Although there are several aspects of this issue, a particularly important one in macroeconomics is the welfare costs of business cycles. The idea is straightforward. The best a macroeconomist can hope to achieve in terms of welfare improvement is eliminating completely the cyclical variation of macroeconomic aggregates. In some sense, this is the equivalent to eliminating systematic risk,

*We gratefully acknowledge the comments of Larry Christiano, Wouter den Haan, Robert F. Engle, Pedro C. Ferreira, Antonio Fiorenco, Clive Granger, Rodolfo Manuelli, Samuel Pessoa and Octavio Tourinho. All errors are ours. We thank CNPq-Brazil and PRONEX for financial support.

even if idiosyncratic risk is still present. Assuming a complete success in this task, Lucas (1987, 3) calculates the amount of extra consumption a rational consumer would require in order to be indifferent between an infinite sequence of consumption under uncertainty and its certainty equivalent – a cycle-free consumption sequence. For 1983 figures, using a reasonable parametric utility function (CES or Power utility function), and post-WWII data, the extra consumption is about \$ 8.50 per person in the U.S., a surprisingly low amount.

This led to the conclusion that eliminating business cycles is undesirable, since the upper-bound of its payoff is extremely low, while the cost of implementing such policy may be high. The power of the argument comes from the simplicity of the calculations, which basically rely on the assumption that consumption is *log-Normally* distributed about a deterministic trend and that preferences are represented by a CES utility function. Specifically, it is not assumed a full structural model describing how the economy works, in the sense that consumption is described using only actual data and a reduced form. The only structure imposed is how preferences are represented.

Several papers have been written since Lucas first presented these results and subsequent work have either changed the basic environment of the problem or relaxed the basic assumptions made to compute the welfare costs of business cycles. For example, Imrohroglu (1989) and Atkeson and Phelan (1995) recalculated welfare costs using a structural model in an artificial-economy environment, where the hypothesis of complete markets was relaxed. Obstfeld(1994), Van Wincoop(1994), Pemberton(1996), Dolmas(1998) and Tallarini(2000) have either changed preferences or relaxed expected utility maximization¹. In some of these papers, the welfare costs of business cycles can reach up to 25% of per-capita consumption in all dates and states of nature.

Of course, these extreme results can be criticized on different grounds. Incomplete markets are a fact of life in real economies. Even in good times, some people will suffer from the fact that they cannot insure against undiversifiable risk, and there is not much macroeconomic policy can do about that. Regarding preferences, Otrok(2000) notes that “it is trivial to make the welfare cost of business cycle as large as one wants by simply choosing an appropriate form for preferences.” This happens because, when time separability of the utility function is lost, consumers treat economic fluctuations as changes in growth, which, of course, have large welfare effects.

In this paper we depart from the original exercise of Lucas in two different ways, while keeping the CES specification for the utility function. First, we chose to model the trend in (the log of) consumption as a stochastic processes process containing a unit root². This choice relies on a sizable amount of econometric evidence available on consumption, or, alternatively, on the amount of authors that have used such a specification, e.g., Hall(1978), Nelson and Plosser(1982), Camp-

¹Obstfeld and Dolmas have also changed the representation of (the log of) consumption to include a unit root in the autoregressive polynomial.

²Lucas(1987, pp. 22-23, footnote 1) explicitly considers the possibility that the trend in consumption is stochastic as in Nelson and Plosser(1982).

bell(1987), King et al.(1991), Cochrane(1994), and Issler and Vahid(2001), *inter alia*. This is a potentially interesting departure, since, in this case, the unconditional variance of consumption is infinite, which can lead to a high payoff for eliminating consumption variability. Our second departure from Lucas is changing the sample from which to compute the moments of consumption. The whole of the literature chose to work with post-WWII data. However, for this period, actual consumption is already a result of counter-cyclical policies, and potentially smoother than what it otherwise have been in their absence. Hence, the previous literature measured the benefits of a *marginal* effort of business-cycle smoothing, and not the welfare costs of business cycles observed in economies with no counter-cyclical policy. To compute the latter we also use pre-WWII data, being able to discuss “The Welfare Costs of Business Cycles in the 20th Century.”

Following the work of Beveridge and Nelson(1981), and its generalization in Stock and Watson(1988), the secular trend in (the log of) consumption is modelled as a random walk and its cyclical component reasonably approximated by an *ARMA* process. As a curiosity, and to make our results comparable to previous work, we also modelled the trend as either a deterministic linear process (with and without a break) or following a slowly evolving secular process captured by the Hodrick and Prescott(1997) filter.

Our paper has two original contributions. First, we take an econometric approach, asking the question of how to best model the reduced form for consumption prior to any welfare calculations. This, coupled with the CES specification for preferences, is enough to compute the welfare cost of business cycles. Indeed, as noted by Lucas(1987, p. 30): “It is worth re-emphasizing that these calculations rest on assumptions about preferences *only*, and not about any particular mechanism – equilibrium and disequilibrium – assumed to generate business cycles.” While our focus is on the reduced form, other authors have recently focused on structural forms. For example, Tallarini(2000) and Otrok(2000) choose preference parameters respectively to match financial data and observed fluctuations in a general equilibrium model of business cycles, despite the fact that several of those structural models have failed to pass either econometric specification tests or being counter-factual; see Hansen and Singleton(1982, 1984), Epstein and Zin(1991) and Mehra and Prescott(1985). Our second original contribution is using pre- and post-WWII data to compute welfare costs. As is well known, business cycles in these two periods are very different; see Romer(1986, 1989, 1999) and Watson(1994). This allows making the distinction between welfare costs and *marginal* welfare costs of business cycles. Using long-span data has the additional benefit of taking into account the existence of rare recurrent events with a large impact on welfare, which may be absent when short-span data is used.

The paper is divided as follows. Section 2 provides a theoretical and statistical framework to evaluate the welfare costs of business cycles. Section 3 provides the estimates that are used in calculating them. Section 4 provides the calculations results, and Section 5 concludes. There is

also an Appendix providing the econometric background necessary to implement the calculations carried out in the paper.

2. The Problem

Lucas (1987) proposed the following way to evaluate the welfare gains of cycle smoothing. Suppose that consumption (c_t) is *log-Normally* distributed about a deterministic trend:

$$c_t = \alpha_0 (1 + \alpha_1)^t \exp\left(-\frac{1}{2}\sigma_z^2\right) z_t, \quad (2.1)$$

where $\ln(z_t) \sim N(0, \sigma_z^2)$. Cycle-free consumption will be the sequence $\{c_t^*\}_{t=0}^\infty$, where $c_t^* = E(c_t) = \alpha_0 (1 + \alpha_1)^t \exp(-\frac{1}{2}\sigma_z^2) E(z_t) = \alpha_0 (1 + \alpha_1)^t$. Notice that $\{c_t^*\}_{t=0}^\infty$ is the resulting sequence when we replace the random variable c_t with its unconditional mean. Hence, for any particular time period, c_t represents a mean-preserving spread of c_t^* .

Another intuitive way of thinking about c_t^* is realizing that, if there is no mean correction $\exp(-\frac{1}{2}\sigma_z^2)$ in (2.1), letting $c_t = \alpha_0 (1 + \alpha_1)^t z_t$, then,

$$c_t^* = \lim_{\sigma_z^2 \rightarrow 0} E(c_t) = \lim_{\sigma_z^2 \rightarrow 0} \alpha_0 (1 + \alpha_1)^t \exp\left(-\frac{1}{2}\sigma_z^2\right) = \alpha_0 (1 + \alpha_1)^t.$$

Hence, c_t^* is a degenerate random variable with all the mass of its distribution in $\alpha_0 (1 + \alpha_1)^t$, which is risk free. It then becomes obvious that the term $\exp(-\frac{1}{2}\sigma_z^2)$ in (2.1) is just the mean correction that makes c_t a mean-preserving spread of c_t^* .

Of course, risk averse consumers prefer $\{c_t^*\}_{t=0}^\infty$ to $\{c_t\}_{t=0}^\infty$. Then, to evaluate the welfare gains of cycle smoothing, amounts to calculating λ , which solves the following equation³:

$$E\left(E_0 \sum_{t=0}^{\infty} \beta^t u((1 + \lambda) c_t)\right) = \sum_{t=0}^{\infty} \beta^t u(c_t^*), \quad (2.2)$$

where $E_t(\cdot) = E(\cdot | \Omega_t)$ is the conditional expectation operator of a random variable, using Ω_t as the information set. Then, the welfare gains are expressed as the compensation λ , that consumers would require at all dates and states of nature, that makes them indifferent between the uncertain stream $\{c_t\}_{t=0}^\infty$ and the stream $\{c_t^*\}_{t=0}^\infty$. Notice that uncertainty here comes in the form of stochastic business cycles, since the trend in consumption is deterministic.

To start the discussion of difference-stationary consumption, we first assume that the utility function is in CES class:

$$u(c_t) = \frac{c_t^{1-\phi} - 1}{1 - \phi}, \quad (2.3)$$

³Notice that Lucas (1987) uses the unconditional mean operator instead of the conditional mean operator in (2.2). The same problem can be proposed using the conditional expectation instead. This is exactly how we proceed in this paper.

where $u(c_t)$ approaches $\ln(c_t)$ as $\phi \rightarrow 1$. As shown in Beveridge and Nelson(1981), every linear difference-stationary process can be decomposed as the sum of a deterministic term, a random walk trend, and a stationary cycle (*ARMA* process). This result was generalized by Stock and Watson(1988) in a multivariate framework, and can be shown to apply to a cointegrated vector autoregression (VAR), or vector error-correction model (VECM).

The analogue of (2.1) when consumption is difference stationary is:

$$\begin{aligned} \ln(c_t) &= \ln(\alpha_0) + \ln(1 + \alpha_1) \cdot t - \frac{\omega_t^2}{2} + \sum_{i=1}^t \epsilon_i + \sum_{j=0}^{t-1} \psi_j \mu_{t-j} \\ &= \ln(\alpha_0 \cdot (1 + \alpha_1)^t) - \frac{\omega_t^2}{2} + \ln(X_t) + \ln(Y_t), \end{aligned} \quad (2.4)$$

where $\ln[\alpha_0 (1 + \alpha_1)^t \cdot \exp(-\omega_t^2/2)]$ is the deterministic term, $\ln(X_t) = \sum_{i=1}^t \epsilon_i$ is the random walk component, $\ln(Y_t) = \sum_{j=0}^{t-1} \psi_j \mu_{t-j}$ is the *MA*(∞) representation of the stationary part (cycle), and $\omega_t^2 = \sigma_{11} \cdot t + 2 \cdot \sigma_{12} \sum_{j=0}^{t-1} \psi_j + \sigma_{22} \sum_{j=0}^{t-1} \psi_j^2$ is the conditional variance of $\ln(c_t)$, included in the mean-correction term.

The permanent shock ϵ_t and the transitory shock μ_t are assumed to have a bi-variate Normal distribution as follows:

$$\begin{pmatrix} \epsilon_t \\ \mu_t \end{pmatrix} \sim IN \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right), \quad (2.5)$$

i.e., shocks are independent, thus uncorrelated, across time, but may be contemporaneously correlated if $\sigma_{12} \neq 0$.

Under normality, the structure in (2.5) encompasses several cases of trend-cycle decomposition methods existent in the literature, particularly those based on the Beveridge-Nelson decomposition. For example, the Unobserved Components method discussed by Watson(1986) and the method in King et al.(1991) impose $\sigma_{12} = 0$, which is not required in the method proposed by Vahid and Engle(1993).

Calculating the welfare cost of business cycles for the difference-stationary case requires first a discussion on how to deal with the fact that now uncertainty comes both in the trend and the cyclical component of $\ln(c_t)$. Moreover, since the trend component has a unit root, its unconditional mean and variance are not defined. Notice that, in the exercise proposed by Lucas, all the cyclical variation in $\ln(c_t)$ is eliminated, which is equivalent to eliminating *all* its variability, since the trend is deterministic. Here, this equivalence is lost, because the trend is stochastic as well.

To deal with this issue, we follow Obstfeld(1994) in considering not the unconditional expectation operator $E(\cdot)$, but the conditional expectation operator $E_0(\cdot)$ in (2.2), where now c_t^* is redefined as $c_t^* = E_0(c_t)$. Therefore, we are assuming that it is possible to offer the consumer an uncertain

consumption stream c_t^* (with no trend and cyclical variation) based on information available at the outset of the problem. Of course, the alternative for the consumer is to face c_t , which has a conditional variance that depends on ω_t^2 . Because consumption has now a unit root, $\omega_t^2 \rightarrow \infty$, as $t \rightarrow \infty$ (although $\omega_t^2 < \infty$ for all t finite). Hence, uncertainty can get relatively large as the horizon increases⁴.

As in Obstfeld, the problem we propose solving here is:

$$E_0 \sum_{t=0}^{\infty} \beta^t u((1+\lambda)c_t) = \sum_{t=0}^{\infty} \beta^t u(E_0(c_t)). \quad (2.6)$$

Under (2.3), (2.4) and (2.5), and using the properties of the moments of log-normal distributions, we can calculate (2.6). Apart from an irrelevant constant term, its left-hand side is given by:

$$E_0 \sum_{t=0}^{\infty} \beta^t u((1+\lambda)c_t) = \frac{[\alpha_0(1+\lambda)]^{1-\phi}}{1-\phi} \sum_{t=0}^{\infty} \left[\beta(1+\alpha_1)^{1-\phi} \right]^t \exp \left[-\frac{(1-\phi)\phi\omega_t^2}{2} \right]. \quad (2.7)$$

Notice that, (2.7) converges if $\beta(1+\alpha_1)^{1-\phi} \exp \left[-\frac{(1-\phi)\phi\sigma_{11}}{2} \right] < 1$.

Calculating the conditional mean of c_t yields $c_t^* = E_0(c_t) = \alpha_0(1+\alpha_1)^t \exp \left(-\frac{\omega_t^2}{2} \right) \times E_0(X_t Y_t) = \alpha_0(1+\alpha_1)^t$. Hence, apart from an irrelevant constant term, the right-hand side of (2.6) is:

$$\sum_{i=0}^{\infty} \beta^t u(c_t^*) = \frac{\alpha_0^{1-\phi}}{1-\phi} \sum_{t=0}^{\infty} \left[\beta(1+\alpha_1)^{1-\phi} \right]^t, \quad (2.8)$$

which converges if $\beta(1+\alpha_1)^{1-\phi} < 1$.

Given the parameters defining the processes $\{c_t^*\}_{t=0}^{\infty}$ and $\{c_t\}_{t=0}^{\infty}$, $\lambda(\phi, \beta)$ is:

$$\lambda(\phi, \beta) = \left\{ \frac{\sum_{t=0}^{\infty} \left[\beta(1+\alpha_1)^{1-\phi} \right]^t}{\sum_{t=0}^{\infty} \left[\beta(1+\alpha_1)^{1-\phi} \right]^t \exp \left[-\frac{(1-\phi)\phi\omega_t^2}{2} \right]} \right\}^{1/(1-\phi)} - 1. \quad (2.9)$$

If in ω_t^2 in (2.9), we replace $\sigma_{12} \sum_{j=0}^{t-1} \psi_j$ and $\sigma_{22} \sum_{j=0}^{t-1} \psi_j^2$ by their respective unconditional counterparts, $\tilde{\sigma}_{12} = \sigma_{12} \sum_{j=0}^{\infty} \psi_j$ and $\tilde{\sigma}_{22} = \sigma_{22} \sum_{j=0}^{\infty} \psi_j^2$ (which may be a reasonable approximation even for relatively small t , and a very good approximation for large t), making $\omega_t^2 = \sigma_{11} \cdot t + 2\tilde{\sigma}_{12} + \tilde{\sigma}_{22}$, and assuming that the conditions for (2.7) and (2.8) to converge holds, (2.9) converges to:

$$\lambda(\phi, \beta) = \exp \left[\frac{\phi(2\tilde{\sigma}_{12} + \tilde{\sigma}_{22})}{2} \right] \left\{ \frac{\left(1 - \beta(1+\alpha_1)^{1-\phi} \exp \left[-\frac{(1-\phi)\phi\sigma_{11}}{2} \right] \right)}{\left(1 - \beta(1+\alpha_1)^{1-\phi} \right)} \right\}^{1/(1-\phi)} - 1, \quad (2.10)$$

⁴Discounting, however, can balance this increasing uncertainty.

which shows the way we chose to estimate $\lambda(\phi, \beta)$ in this paper⁵, for $\phi \neq 1$; a similar formula applies for the cases when $\phi = 1$ ⁶.

We now turn to other possible ways of modelling the trend component. If the trend is modelled as a deterministic function of time (with or without a break), as in (2.1), then the analysis is done as originally proposed by Lucas(1987). In spite of the fact that Lucas has proposed the analysis as in (2.2) above, he actually implemented it in a different way (see Lucas(1987, footnote 2, p. 23)), removing the trend in consumption using the filtering procedure proposed in Hodrick and Prescott(1997). The filter is two sided, i.e., uses past and future consumption values to get the slowly-moving trend. In principle, the trend removed using such a procedure should be treated as a random variable. However, for simplicity, Lucas treated the trend as deterministic, which we also do here. Hence, when using the Hodrick and Prescott trend, our results should be viewed as a lower-bound for the gains of cycle smoothing. To implement the calculations in this case, we computed the deterministic growth rate present in the Hodrick and Prescott trend, treating the cyclical component as in (2.6) above. Hence, $c_t = \alpha_0 (1 + \alpha'_1)^t \exp(-\sigma_{z'}^2/2)z'_t$, $\ln(z'_t) \sim N(0, \sigma_{z'}^2)$, and $c_t^* = \alpha'_0 (1 + \alpha'_1)^t$, where α'_0 and α'_1 are now the deterministic components associated with the Hodrick-Prescott trend, and z'_t is the residual cyclical component associated with it.

3. Reduced Form and Long-Run Constraints

A full discussion of the econometric models employed here can be found in Beveridge and Nelson(1981), Stock and Watson(1988), Engle and Granger(1987), Campbell(1987), Campbell and Deaton(1989), and Proietti(1997). We start by assuming that $y_t = (\ln(c_t), \ln(I_t))'$ is a 2×1 vector containing the logarithms of consumption and disposable income. We also assume that both series individually contain a unit-root, and are generated by a p -th order vector autoregression (VAR):

$$\begin{aligned} y_t - \pi_1 y_{t-1} - \pi_2 y_{t-2} - \dots - \pi_p y_{t-p} &= \varepsilon_t, \\ (I_n - \pi_1 L - \pi_2 L^2 - \dots - \pi_p L^p) y_t &= \varepsilon_t, \text{ or,} \\ \Pi(L) y_t &= \varepsilon_t. \end{aligned}$$

and we decompose $\Pi(L) = I_n - \pi_1 L - \pi_2 L^2 - \dots - \pi_p L^p$ as:

$$\Pi(L) = -\Pi(1) L^p + (1 - L) \Gamma(L),$$

⁵ Jumping straight to our results, we have observed that, for all values of (ϕ, β) we considered here $\beta(1 + \alpha_1)^{1-\phi} < 1$. However, it was not always the case that $\beta(1 + \alpha_1)^{1-\phi} \exp\left[-\frac{(1-\phi)\phi\sigma_{11}}{2}\right] < 1$, since the term $\exp\left[-\frac{(1-\phi)\phi\sigma_{11}}{2}\right]$ was always greater than unity, and sometimes large enough as to prevent the convergence condition to hold.

⁶ When (the log of) consumption is a martingale as in Obstfeld(1994), (2.10) collapses to equation (5) in his paper; see p. 1474.

leading to the vector error-correction model (VECM):

$$\begin{aligned}\Delta y_t - \Gamma_1 \Delta y_{t-1} - \Gamma_2 \Delta y_{t-2} - \cdots - \Gamma_{p-1} \Delta y_{t-p+1} - \Pi(1) y_{t-p} &= \varepsilon_t, \text{ or,} \\ \Delta y_t - \Gamma_1 \Delta y_{t-1} - \Gamma_2 \Delta y_{t-2} - \cdots - \Gamma_{p-1} \Delta y_{t-p+1} - \gamma \alpha' y_{t-p} &= \varepsilon_t,\end{aligned}$$

where $\Gamma_j = -I_n + \sum_{i=1}^j \pi_i$, $j = 1, 2, \dots, p-1$.

Cointegration between the logarithms of consumption and income arises naturally using the theory of permanent-income. There, consumption can be viewed as proportional to the expected present discounted value of all income stream. Hence, the expected present value of consumption and income are equal, and both series are proportional in the long run. Indeed, using a logarithm approximation to the equations implied by permanent-income theory, Campbell(1987) and Campbell and Deaton(1989) show that the logarithm of the saving ratio:

$$\begin{aligned}s_t &= \ln(S_t/I_t) = \ln((I_t - C_t)/I_t) = \ln(1 - C_t/I_t) \\ &\approx -\ln(C_t/I_t) = \ln(I_t) - \ln(C_t),\end{aligned}\tag{3.1}$$

can be approximated by a linear function of instantaneous expected future-income growth $E_t(\Delta \ln I_{t+s})$:

$$s_t \approx -\sum_{s=1}^{\infty} \rho^s E_t(\Delta \ln I_{t+s}),\tag{3.2}$$

where ρ is the one-period discount factor for future income growth.

Combining (3.1) and (3.2) shows that:

$$\ln(I_t) - \ln(C_t) \approx -\sum_{s=1}^{\infty} \rho^s E_t(\Delta \ln I_{t+s}),\tag{3.3}$$

i.e, that the logarithm difference between disposable income and consumption is stationary. Hence, the logarithms of consumption and disposable income must cointegrate if $\ln(I_t)$ is an integrated series. Moreover, the cointegrating vector will be $\alpha = (-1, 1)'$. Hence, the final reduced form to be estimated, after appropriate testing is:

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \gamma(-1, 1) y_{t-p} + \varepsilon_t\tag{3.4}$$

We turn now to the discussion of how to extract trends and cycles from (3.4). Jumping straight to our results, we found that the system (??) is well described by a *VECM*(1), which can be put in state-space form, as discussed in Proietti(1997):

$$\begin{aligned}\Delta y_{t+1} &= Z f_{t+1} \\ f_{t+1} &= T f_t + Z' \varepsilon_{t+1},\end{aligned}\tag{3.5}$$

where,

$$f_{t+1} = \begin{bmatrix} \Delta y_{t+1} \\ \Delta y_t \\ \alpha' y_{t-1} \end{bmatrix}, \quad T = \begin{bmatrix} \Gamma_1 - \gamma \alpha' - \gamma \\ I_2 & 0 & 0 \\ 0 & \alpha' & 1 \end{bmatrix}$$

with the associated VECM being,

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \gamma \alpha' y_{t-2} + \varepsilon_t, \quad \text{and,} \\ Z = [I_2 \ \mathbf{0} \ \mathbf{0}].$$

From the work of Beveridge and Nelson(1981), and Stock and Watson(1988), ignoring initial conditions and deterministic components, the series in y_t can be decomposed into a trend (μ_t) and a cyclical component (ψ_t), as follows:

$$y_t = \mu_t + \psi_t,$$

where,

$$\mu_t = y_t + \lim_{l \rightarrow \infty} \sum_{i=1}^l E_t [\Delta y_{t+i}], \quad \text{and,} \\ \psi_t = - \lim_{l \rightarrow \infty} \sum_{i=1}^l E_t [\Delta y_{t+i}].$$

It is straightforward to show that μ_t is a multivariate random-walk. Using the state-space representation (3.5), we can compute the limits above. The cyclical and trend components will be, respectively:

$$\psi_t = -Z [I_m - T]^{-1} T f_t, \\ \mu_t = y_t - \psi_t. \tag{3.6}$$

or, using formulas (6) and (7) in Proietti(1997),

$$\psi_t = -K \Gamma^*(L) \Delta y_t + P y_t, \tag{3.7}$$

and

$$\mu_t = K \sum_{i=1}^t \varepsilon_i, \tag{3.8}$$

where K and P are projection matrices; see the Appendix.

We can also use (3.5) to forecast trend and cyclical components at any horizon into the future. The forecast of ψ_{t+s} , given information up to t , is:

$$\widehat{\psi}_{t+s|t} = E_t [\psi_{t+s}] = -K \Gamma^*(L) Z T f_{t+s-1} + P y_t + P Z \left(\sum_{i=1}^S T^i \right) f_t,$$

and the forecast of μ_{t+s} , given information up to t , is:

$$\widehat{\mu}_{t+s|t} = \mu_t,$$

since the best forecast of a random walk $t + s$ periods ahead is simply its value today.

To fully characterize the elements in (2.10), we need to compute the variance and the covariance of forecasts of trend and cyclical components. Recall that the conditional expectation of a log-Normal random variable is just a function of the mean and variance of the normal distribution associated with it. Hence, to compute the variances of these forecasts, we have just to apply standard results of state-space representations. It is straightforward to show that:

$$E_t \left[(\mu_{t+s} - \widehat{\mu}_{t+s|t}) (\mu_{t+s} - \widehat{\mu}_{t+s|t})' \right] = s \cdot KQK',$$

where $E_t [\varepsilon_{t+i} \varepsilon_{t+i}'] = Q$, and that,

$$E_t \left[(\psi_{t+s} - \widehat{\psi}_{t+s|t}) (\psi_{t+s} - \widehat{\psi}_{t+s|t})' \right] = VQV' + P \left(\sum_{i=1}^{S-1} W(i)QW(i)' \right) P'$$

and

$$E_t \left[(\mu_{t+s} - \widehat{\mu}_{t+s|t}) (\psi_{t+s} - \widehat{\psi}_{t+s|t})' \right] = KQV' + K \left[\sum_{i=1}^{S-1} QW(i)' \right] P',$$

where $V = [P - K\Gamma^*(1)]$, as computed in the Appendix.

Based on these last three covariance matrices the correlations between trend and cyclical components of the data can be fully characterized, recalling that $y_t = (\ln(c_t), \ln(I_t))'$. Hence, to get the corresponding element of means, variances, and covariances associated with $\ln(c_t)$, one has simply to choose the appropriate elements of these vectors and matrices.

4. Empirical Results

Data for consumption of non-durables and services were obtained from DRI from 1929 through 2000. Data for consumption of perishables and services from 1901 to 1929 were obtained from Kuznets(1957) in real terms, and then chained with DRI data, resulting in a long-span series for consumption of non-durables and services from 1901-2000. Data for real GNP were also extracted from DRI from 1929 through 2000 and from Kuznets(1957) from 1929 through 2000. Data on population were extracted from Kuznets and DRI, and then chained. Figure 1 presents the data on consumption and income per-capita for the whole period 1901-2000. The peculiar features are first the magnitude of the great depression in both consumption and income behavior, and second the fact that pre-WWII data present much more volatility than post-WWII data.

We fitted a bi-variate vector autoregression for the logs of consumption and income. Lag length selection indicated that a VAR(2) with an unrestricted constant term was an appropriate description

of the dynamic system. This was true not only in terms of minimizing information criteria but also because this specification did not fail diagnostic testing.

Table 1 presents results of the cointegration test using Johansen's(1988, 1991) technique. The *Trace Statistics* for the null of no cointegration and of at most one cointegrating vector were respectively 16.43 and 0.18. At 5% significance, we concluded that there is one cointegrating vector, given by $(-1.000, 1.005)'$. Conditioning on the existence of one cointegrating vector, we tested the restriction that it was equal to $(-1, 1)'$. Testing this hypothesis, using the likelihood-ratio test in Johansen(1991), yielded a p-value of 0.831, not rejecting the null at usual levels of significance. An interesting by-product of cointegration analysis is testing the significance of the error-correction term in each regression of the system. The *t-statistics* associated with this test are -0.07 and 3.16, for the regression involving consumption and income respectively. Hence, the error-correction term affects income but not consumption, and the latter is long-run *weakly exogenous* in the sense of Engle, Hendry, and Richard(1983) and Johansen(1992).

Given the restricted VECM found in the empirical analysis, we implemented the multivariate Beveridge and Nelson(1981) decomposition in the form suggested by Proietti(1997). Figure 2 shows the result of this exercise. The consumption series and the trend are very close throughout the whole period, reflecting the fact that agents do update their beliefs about future income, and that the permanent-income theory is probably a reasonable approximation to consumption behavior; see Cochrane(1994) *inter alia*. It is obvious that the cyclical component of consumption varies much more in the pre-WWII era than after that period; see also Figures 3-5 for a similar pattern under alternative decompositions.

Table 2 displays the description of the data in terms of the parameters estimates associated with the reduced form of (the log of) consumption under alternative models for the trend in it. Estimates are obtained for three distinct periods: pre-WWII data – 1901-1941, post-WWII data – 1947-2000, and the whole sample – 1901-2000. It is obvious that uncertainty in the pre-WWII period is much larger than in the post-WWII period. For example, using the Beveridge and Nelson decomposition, the variance of the trend and cyclical innovations are more than one order of magnitude larger in the pre-WWII era than in the post-WWII era. Moreover, growth rates were about 50% larger in the post-WWII era compared with the pre-WWII era. These results, of course, will impact the final computations of welfare measures.

The estimates of the welfare cost of business cycles in the 20th Century are presented in Table 3. We start the discussion using the post-WWII period, keeping at hand the Lucas benchmark values for comparison. First, it is comforting to note that, when the HP filter is used, we were able to reproduce his benchmark numbers almost exactly for the 1947-2000 sample, although we use annual data with a larger span and he used quarterly data up to 1983. Second, the linear time trend model produced numbers that are about five times higher than those produced using the HP filter. This shows that interchanging these procedures is not innocuous in this context, since the variance of

the cyclical components increases dramatically when the linear time trend is used. Third, using the Beveridge and Nelson decomposition for post-WWII data gives qualitatively the same results of using the linear time trend model. For the former, a reasonable estimate of the welfare costs of business cycles will be about 0.25% of consumption (associated with $\phi = 2$ and $\beta = 0.971$), which is about US\$ 50.00 a year, in all dates and states of nature, for the year 2000. For the latter, the welfare costs of business cycles will also about 0.4% of consumption (associated with $\phi = 2$), which is about US\$ 80.00 a year. These numbers are much higher than the US\$ 8.50 found by Lucas, but are not a sizable amount either. Hence, qualitatively, these results are the same.

We now turn our attention to the analysis of the pre-WWII period (1901-1941). First, when the linear time trend model produced numbers that are about five times bigger than those produced using the HP filter. For a reasonable degree of risk aversion ($\phi = 2$) we get welfare costs of about 1% of consumption, which translates into US\$ 200.00 a year in current value. Second, using the Beveridge and Nelson decomposition for post-WWII now gives completely different results than those using the other two methods. For higher degrees of risk aversion ($\phi > 15$) there is no compensation that will make the consumer indifferent between the uncertain stream $\{c_t\}_{t=0}^{\infty}$ and the stream $\{c_t^*\}_{t=0}^{\infty}$. This happens because the variance of the innovation of the trend component is too high. For reasonable preference parameter and discount values ($\beta = 0.971, \phi = 2$) we get a compensation of about 5% of consumption, which is by all means a sizable welfare cost (about US\$ 1,000.00 a year). This last estimate is about half of the estimates in Dolmas(1998), Tallarini(2000) and Van Wincoop(1994), where preferences exhibit non-separability across time. However, our results are obtained with the same preference setup as in Lucas, where the only difference is the way the trend in consumption is defined, and the sample period used to compute the moments of the data.

Results for the whole period 1901-2000 are indeed a convex combination of those of pre- and post-WWII eras. For reasonable preference parameter and discount values ($\beta = 0.971, \phi = 2$) we get a compensation of about 2% of consumption, with the Beveridge and Nelson trend, of about 0.6% of consumption, with the linear trend, and of about 0.1% of consumption, with the HP trend.

From the discussion above, focusing on the Beveridge and Nelson decomposition, we can conclude the following. First, the *marginal* welfare cost of business cycles is small. Hence, it makes little sense to deepen current counter-cyclical policies. This conclusion is based on results obtained for the post-WWII period – 1947-2000, where consumption behavior was (partially) a result of the current counter-cyclical policies. Second, from the point of view of a pre-WWII consumer, the welfare costs of business cycles were sizable. Indeed, for reasonable parameter values ($\beta = 0.971, \phi = 2$) they were about 5% of consumption in all dates and states of nature. Therefore, from her(is) point of view, it made sense to implement counter-cyclical policies. This conclusion is based on results obtained for the pre-WWII period – 1901-1941, where consumption behavior was (partially) a result of the lack of counter-cyclical policies. Last but not least, a comparison between the welfare costs of business cycles in the pre-WWII and post-WWII period can give some idea of the effectiveness of counter-

cyclical policies implemented in the latter period. Since welfare costs decreased from about 5% to about 0.3% of consumption, and if this reduction can be credited to the existence of counter-cyclical policies, then the latter have been proven effective in increasing the welfare of rational consumers.

5. Conclusions

A major issue in macroeconomics is whether or not governments or central banks should pursue counter-cyclical policies. As central banks are nowadays more prone to implement inflation targeting policies, and these policies aim to reduce the variance of the cyclical component in output, this question is really contemporary.

Using only standard assumptions on preferences and a reasonable reduced form for consumption and income, we computed the welfare cost of business cycles for the 20th Century using three alternative trend-cycle decomposition methods, but focusing on the results of the Beveridge and Nelson(1981) decomposition. Our results show that the post-WWII era is a very quiet one, with the welfare costs rarely exceeding 1% of consumption. Although the benchmark values computed by Lucas are relatively modest compared to this value, our basic conclusion is that deepening counter-cyclical policies is futile.

However, if we focus on the pre-WWII era, this basic conclusion is altered. The linear time trend model produced numbers that are about five times higher than those produced using the HP filter. For a reasonable degree of risk aversion ($\phi = 3$) we get welfare costs of about 2% of consumption. This translates into US\$ 400.00 a year in current value. The Beveridge and Nelson decomposition, for pre-WWII, gives completely different results than those using the other two methods. For higher degrees of risk aversion ($\phi > 15$) there is no compensation that will make the consumer indifferent between the uncertain stream $\{c_t\}_{t=0}^{\infty}$ and the stream $\{c_t^*\}_{t=0}^{\infty}$. This happens because the variance of the innovation of the trend component is too high. For reasonable preference parameter and discount values ($\beta = 0.971, \phi = 2$) we get a compensation of about 5% of consumption, which is by all means a sizable welfare cost (about US\$ 1,000.00 a year). This last estimate is about half of the estimates in Dolmas(1998), Tallarini(2000) and Van Wincoop(1994), where preferences exhibit non-separability across time. However, our results are obtained with the same preference setup as in Lucas, where the only difference is the way the trend in consumption is defined.

From the discussion above, focusing on the Beveridge and Nelson decomposition, we can conclude the following. First, the *marginal* welfare cost of business cycles is small, which goes against deepening current counter-cyclical policies. However, from the point of view of a pre-WWII consumer, the welfare costs of business cycles were sizable. Second, since welfare costs of business cycles decreased from about 5% to about 0.3% of consumption, from the pre-WWII to the post-WWII era, if this reduction can be credited to counter-cyclical policies, then the latter have been proven very effective.

References

- Atkeson, A. and Phelan, C., 1995, "Reconsidering the Cost of Business Cycles with Incomplete Markets", *NBER Macroeconomics Annual*, 187-207, with discussions.
- Beveridge, S. and Nelson, C.R., 1981, "A New Approach to Decomposition of Economic Time Series into a Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, 7, 151-174.
- Campbell, J. 1987, "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica*, vol. 55(6), pp. 1249-73.
- Campbell, John Y. and Deaton, Angus 1989, "Why is Consumption So Smooth?" *The Review of Economic Studies* 56:357-374.
- Campbell, J.Y., 1999, "Asset Prices, Consumption, and the Business Cycle", *Handbook of Macroeconomics*, vol. 1, edited by J.B. Taylor and M. Woodford. Amsterdam: Elsevier, 1231-1303.
- Campbell, J.Y. and Cochrane, J.H., 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior", *Journal of Political Economy*, 107 (2), 205-251.
- Cochrane, J.H., 1994, "Permanent and Transitory Components of GNP and Stock Prices," *Quarterly Journal of Economics*, 30, 241-265.
- Christiano, L.J., and Eichenbaum, M., 1989, "Unit Roots in GNP: Do We Know and Do We Care?" *Carnegie-Rochester Conference Series on Public Policy*.
- Dolmas, J., 1998, "Risk Preferences and the Welfare Cost of Business Cycles", *Review of Economic Dynamics*, 1, 646-676.
- Engle, R.F. and Granger C.W.J., 1987, "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, 251-276.
- Engle, R.F., Hendry, D.F., and Richard, J.F., 1983, "Exogeneity", *Econometrica*, vol. 55, pp. 277-304.
- Epstein, L. and Zin, S., 1991, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation," *Journal of Political Economy*, 99, 263-286.
- Hall, R.E., 1978, "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971-987.

- Hansen, L. and Singleton, K., 1982, "Generalized Instrumental Variables Estimation of Nonlinear Expectations Models," *Econometrica*, 50 (5), 1269-1286.
- Hansen, L. and Singleton, K., 1984, Errata of "Generalized Instrumental Variables Estimation of Nonlinear Expectations Models," *Econometrica*, 52 (1), 267-268.
- Hodrick, R.J. and Prescott, E.C., 1997, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, 29, 116.
- Imrohorglu, Ayse, 1989, "Cost of Business Cycles With Indivisibilities and Liquidity Constraints", *Journal of Political Economy*, 97 (6) , 1364-1383.
- Issler, J.V. and Vahid, F., 2001, "Common Cycles and the Importance of Transitory Shocks to Macroeconomic Aggregates," *Journal of Monetary Economics*, 47, 449-475.
- Jarque, C.M. and Bera, A.K., 1987, "A Test of Normality of observations and Regression Residuals," *International Statistical Review*, 55 (2), 163-172.
- Johansen, S., 1988, "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, pp. 231-254.
- Johansen, S., 1991, "Estimation and Hypothesis Testing of Cointegrated Vectors in Gaussian Vector Autoregressions", *Econometrica*, vol. 59-6, pp. 1551-1580.
- Johansen, S., 1992, "Cointegration in Partial Systems and the Efficiency of Single-equation Analysis," *Journal of Econometrics*, vol. 52(3), pp. 389-402.
- King, R.G., Plosser, C.I. and Rebelo, S., 1988, "Production, Growth and Business Cycles. II. New Directions", *Journal of Monetary Economics*, 21, 309-341.
- King, R.G., Plosser, C.I., Stock, J.H. and Watson, M.W., 1991, "Stochastic Trends and Economic Fluctuations", *American Economics Review*, 81, 819-840.
- Kuznets S., 1961, "Capital in the American Economy: Its formation and financing," Princeton University Press.
- Lucas, R., 1987, "*Models of Business Cycles*," Oxford: Blackwell.
- Mehra, R. and Prescott, E., 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145-161.
- Nelson, C.R. and Plosser, C., 1982, "Trends and Random Walks in Macroeconomics Time Series," *Journal of Monetary Economics*, 10, 1045-1066.

- Obstfeld, M., 1994, "Evaluating Risky Consumption Paths: The Role of Intertemporal Substitutability," *European Economic Review*, 38, 1471-1486.
- Otrok, C., 2001, "On Measuring the Welfare Cost of Business Cycles," *Journal of Monetary Economics*, 47, 61-92.
- Pemberton, J., 1996, "Growth Trends, Cyclical Fluctuations, and Welfare with Non-Expected Utility Preferences," *Economic Letters*, 50, 387-392.
- Proietti, T., 1997, "Short-run Dynamics in Cointegrated Systems", *Oxford Bulletin of Economics and Statistics*, 59 (3), 405-422.
- Romer, C.D., 1986, "Is The Stabilization Of The Postwar Economy A Figment Of The Data?," *American Economic Review*, 76 (3), 314-334.
- Romer, C.D., 1990, "The Great Crash And The Onset Of The Great Depression," *Quarterly Journal of Economics*, 105 (3), 597-624.
- Romer, C.D., 1993, "The Nation In Depression," *Journal of Economic Perspectives*, 7 (2), 19-39.
- Romer, C.D., 1989, "The Prewar Business Cycle Reconsidered: New Estimates Of Gross National Product, 1869-1908," *Journal of Political Economy*, 97 (1), 1-37.
- Romer, C.D., 1988, "World War I And The Post War Depression: A Reinterpretation Based On Alternative Estimates Of GNP," *Journal of Monetary Economics*, 22 (1), 91-116.
- Stock, J.H. and Watson, M.W., 1988, "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097-1107.
- Tallarini Jr, T.D., 2000, "Risk-sensitive Real Business Cycles", *Journal of Monetary Economics*, 45, 507-532.
- Stock, J. and Watson, M., 1988, "Testing for Common Trends," *Journal of the American Statistical Association*, 83, pp. 1097-1109.
- Vahid, F. and Engle, R.F., 1993, "Common Trends and common Cycles", *Journal of Applied Econometrics*, 8, 341-360.
- Watson, M.W., 1986, "Univariate Detrending Methods with Stochastic Trends", *Journal of Monetary Economics*, 18, 49-75.
- Watson, M.W., 1994, "Business-Cycle Durations And Postwar Stabilization of The U.S. Economy," *American Economic Review*, 84 (1), 24-46.

Van Wincoop, E., 1994, "Welfare Gains From International Risksharing", *Journal of Monetary Economics*, 34, 175-200.

A. State-Space Representation for Error-Correction Models and the Beveridge and Nelson Decomposition (Proietti(1997))

Proietti(1997) discussed in some length how the Beveridge and Nelson(1981) decomposition can be put in state-space form. Here we adapt some of this discussion. If the series y_t are generated by a vector autoregression (VAR):

$$\begin{aligned} y_t - \pi_1 y_{t-1} - \pi_2 y_{t-2} - \cdots - \pi_p y_{t-p} &= \varepsilon_t, \\ (I_n - \pi_1 L - \pi_2 L^2 - \cdots - \pi_p L^p) y_t &= \varepsilon_t, \text{ or,} \\ \Pi(L) y_t &= \varepsilon_t. \end{aligned}$$

and we decompose $\Pi(L) = I_n - \pi_1 L - \pi_2 L^2 - \cdots - \pi_p L^p$ as:

$$\Pi(L) = -\Pi(1) L^p + (1 - L) \Gamma(L),$$

leading to the vector error-correction model (VECM):

$$\begin{aligned} \Delta y_t - \Gamma_1 \Delta y_{t-1} - \Gamma_2 \Delta y_{t-2} - \cdots - \Gamma_{p-1} \Delta y_{t-p+1} - \Pi(1) y_{t-p} &= \varepsilon_t, \text{ or,} \\ \Delta y_t - \Gamma_1 \Delta y_{t-1} - \Gamma_2 \Delta y_{t-2} - \cdots - \Gamma_{p-1} \Delta y_{t-p+1} - \gamma \alpha' y_{t-p} &= \varepsilon_t, \end{aligned}$$

where $\Gamma_j = -I_n + \sum_{i=1}^j \pi_i$, $j = 1, 2, \dots, p-1$, it is straightforward to put the latter into state space form. To save space, and jumping straight to the series modelled here, we start by assuming that Δy_t is a 2×1 vector containing the instantaneous growth rates of consumption and income, which can be modelled as VECM(1) (or a VAR(2)), where $\alpha' y_t$ is the error-correction vector and γ is the adjustment coefficient vector. The state-space form of the VECM(1) is as follows:

$$\begin{aligned} \Delta y_{t+1} &= Z f_{t+1} \\ f_{t+1} &= T f_t + Z' \varepsilon_{t+1}, \end{aligned} \tag{A.1}$$

where,

$$\begin{aligned} f_{t+1} &= \begin{bmatrix} \Delta y_{t+1} \\ \Delta y_t \\ \alpha' y_{t-1} \end{bmatrix}, \quad T = \begin{bmatrix} \Gamma_1 - \gamma \alpha' - \gamma \\ I_2 & 0 & 0 \\ 0 & \alpha' & 1 \end{bmatrix}, \text{ and,} \\ Z &= [I_2 \ 0 \ 0]. \end{aligned}$$

A.1. Trends and Cycles and State-Space Representation

The basic idea in Beveridge and Nelson(1981) is that, for unit-root processes with zero drift, the random-walk trend in the series and its long-run forecast will both be the same. Hence, for series in y_t can be decomposed into a trend (μ_t) and a cyclical component (ψ_t), as follows:

$$y_t = \mu_t + \psi_t,$$

where,

$$\begin{aligned}\mu_t &= y_t + \lim_{l \rightarrow \infty} \sum_{i=1}^l E_t [\Delta y_{t+i}], \text{ and,} \\ \psi_t &= - \lim_{l \rightarrow \infty} \sum_{i=1}^l E_t [\Delta y_{t+i}].\end{aligned}$$

Using (A.1), we can compute the limits above. The cyclical component will be:

$$\psi_t = -Z [I_m - T]^{-1} T f_t. \quad (\text{A.2})$$

The trend in y_t can be simply computed as:

$$\mu_t = y_t - \psi_t.$$

A.2. Computing Mean Squared Errors

From Proposition 2 in Proietti(1997),

$$\psi_{t+1} = -(I_N - P)(\Gamma(1) + \gamma\alpha')^{-1}\Gamma^*(L) \Delta y_{t+1} + P y_{t+1}, \quad (\text{A.3})$$

and,

$$\mu_{t+1} = (I_N - P)(\Gamma(1) + \gamma\alpha')^{-1}\Gamma(L)y_{t+1}$$

or,

$$\Delta\mu_{t+1} = (I_N - P)(\Gamma(1) + \gamma\alpha')^{-1}\varepsilon_{t+1}, \quad (\text{A.4})$$

where $P = (\Gamma(1) + \gamma\alpha')^{-1}\gamma[\alpha'(\Gamma(1) + \gamma\alpha')^{-1}\gamma]\alpha'$, and $\Gamma(L) = I_2 - \Gamma_1 L$, which is decomposed as:

$$\begin{aligned}\Gamma(L) &= \Gamma(1) + (1 - L)\Gamma^*(L), \text{ where,} \\ \Gamma^*(L) &= \Gamma_1,\end{aligned}$$

in the present context.

From (A.4) we have,

$$\mu_{t+s} = \mu_t + (I_N - P)(\Gamma(1) + \gamma\alpha')^{-1} \sum_{i=1}^s \varepsilon_{t+i},$$

which implies that $\widehat{\mu}_{t+s|t} = E_t[\mu_{t+s}] = \mu_t$. Denoting $K = (I_N - P)(\Gamma(1) + \gamma\alpha')^{-1}$, and $(\mu_{t+s} - \widehat{\mu}_{t+s|t}) = K \sum_{i=1}^s \varepsilon_{t+i}$ we have,

$$\begin{aligned}E \left[(\mu_{t+s} - \widehat{\mu}_{t+s|t}) (\mu_{t+s} - \widehat{\mu}_{t+s|t})' \right] &= K \left(\sum_{i=1}^s Q_i \right) K' \\ &= s \cdot K Q K',\end{aligned} \quad (\text{A.5})$$

where $E[\varepsilon_{+it}\varepsilon'_{t+i}] = Q_i = Q$. On the other hand, from (A.3),

$$\psi_{t+s} = -K\Gamma^*(L)\Delta y_{t+s} + Py_{t+s},$$

but, $\Delta y_{t+1} = ZTf_t + \varepsilon_{t+1}$, which implies that $\Delta y_{t+s} = ZTf_{t+s-1} + \varepsilon_{t+s}$. However, $y_{t+1} = y_t + ZTf_t + \varepsilon_{t+1}$, which implies that $y_{t+s} = y_t + Z\left(\sum_{i=1}^S T^i\right)f_t + \sum_{i=1}^{S-1}\left(I_N + Z\left\{\sum_{j=1}^{S-i} T^j\right\}Z'\right)\varepsilon_{t+i} + \varepsilon_{t+s}$. Hence, $\widehat{\psi}_{t+s|t} = E_t[\psi_{t+s}] = -K\Gamma^*(1)ZTf_{t+s-1} + Py_t + PZ\left(\sum_{i=1}^S T^i\right)f_t$, which implies that $\left(\psi_{t+s} - \widehat{\psi}_{t+s|t}\right) = [P - K\Gamma^*(1)]\varepsilon_{t+s} + P\sum_{i=1}^{S-1}\left(I_N + Z\left\{\sum_{j=1}^{S-i} T^j\right\}Z'\right)\varepsilon_{t+i}$. Denoting $V = [P - K\Gamma^*(1)]$ and $W(i) = \left(I_N + Z\left\{\sum_{j=1}^{S-i} T^j\right\}Z'\right)$. Thus,

$$E\left[\left(\psi_{t+s} - \widehat{\psi}_{t+s|t}\right)\left(\psi_{t+s} - \widehat{\psi}_{t+s|t}\right)'\right] = VQV' + P\left(\sum_{i=1}^{S-1} W(i)QW(i)'\right)P' \quad (\text{A.6})$$

and

$$E\left[\left(\mu_{t+s} - \widehat{\mu}_{t+s|t}\right)\left(\psi_{t+s} - \widehat{\psi}_{t+s|t}\right)'\right] = KQV' + K\left[\sum_{i=1}^{S-1} QW(i)'\right]P'. \quad (\text{A.7})$$

It is straightforward to extract the estimates of σ_{11} , σ_{22} , and σ_{12} , in (2.5) from equations (A.5), (A.6) and (A.7). One has simply to evaluate them at sample estimates selecting their first diagonal element, which will correspond to consumption moments.

Table 1: Cointegration test – Johansen(1988, 1991) Technique

Cointegrating Vectors under H_0	Eigenvalues	Trace Stat.	5 % Crit. Value	λ_{\max} Stat.	5 % Crit. Value
None	0.150	16.44	15.41	16.26	14.07
At most 1	0.0018	0.18	3.76	0.18	3.76

Estimate of the cointegrating vector is: $(-1, 1.005)$.

$H_0 : \beta' = (-1, 1)$, conditional on $r = 1$, $p - value = 0.831$.

Table 2: Parameters Estimates in Equation (2.10)

Consumption										
	1901-2000			1901-1941			1947-2000			1901-2000
	BN	HP	Linear	BN	HP	Linear	BN	HP	Linear	Linear-break
$\ln(1+\alpha_1)$	0.020738	0.020718	0.020718	0.015310	0.014687	0.014517	0.023543	0.023347	0.023380	0.023217
σ_{11}	0.001031			0.002294			0.000142			
σ_{12}	-0.000085			-0.000198			0.000008			
σ_{22}	0.000271	0.000640	0.005124	0.000541	0.001278	0.008231	0.000049	0.000169	0.002799	0.001904

Table 3: Consumption Compensation ($\lambda\%$) for Different (β, ϕ) Values

(a) Lucas (1987) Benchmark Values				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950, 0.971, 0.985$	0.008	0.042	0.08	0.17
(b) Beveridge-Nelson Decomposition 1901-2000				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950$	0.99	1.93	2.26	2.74
$\beta = 0.971$	1.75	2.35	2.54	2.97
$\beta = 0.985$	3.45	2.74	2.77	3.14
(c) Beveridge-Nelson Decomposition 1901-1941				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950$	2.21	5.63	10.42	∞
$\beta = 0.971$	3.92	7.35	13.28	∞
$\beta = 0.985$	7.83	9.15	16.29	∞
(d) Beveridge-Nelson Decomposition 1947-2000				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950$	0.14	0.25	0.27	0.30
$\beta = 0.971$	0.24	0.29	0.30	0.31
$\beta = 0.985$	0.47	0.33	0.32	0.32
(e) Hodrick-Prescott Filter for Trend 1901-2000				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950, 0.971, 0.985$	0.03	0.16	0.32	0.64
(f) Hodrick-Prescott Filter for Trend 1901-1941				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950$	0.06	0.32	0.64	1.29
(g) Hodrick-Prescott Filter for Trend 1947-2000				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950, 0.971, 0.985$	0.01	0.04	0.08	0.17
(h) Linear Time Trend 1901-2000				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950, 0.971, 0.985$	0.26	1.29	2.60	5.26
(i) Linear Time Trend 1901-1941				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950, 0.971, 0.985$	0.41	2.08	4.20	8.58
(j) Linear Time Trend 1947-2000				
β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950, 0.971, 0.985$	0.14	0.70	1.41	2.84

(k) Linear Time Trend with break 1901-2000

β Equivalent in a Yearly Basis	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 20$
$\beta = 0.950, 0.971, 0.985$	0.10	0.48	0.96	1.92

GNP and Consumption of Non-durables and Services per-capita

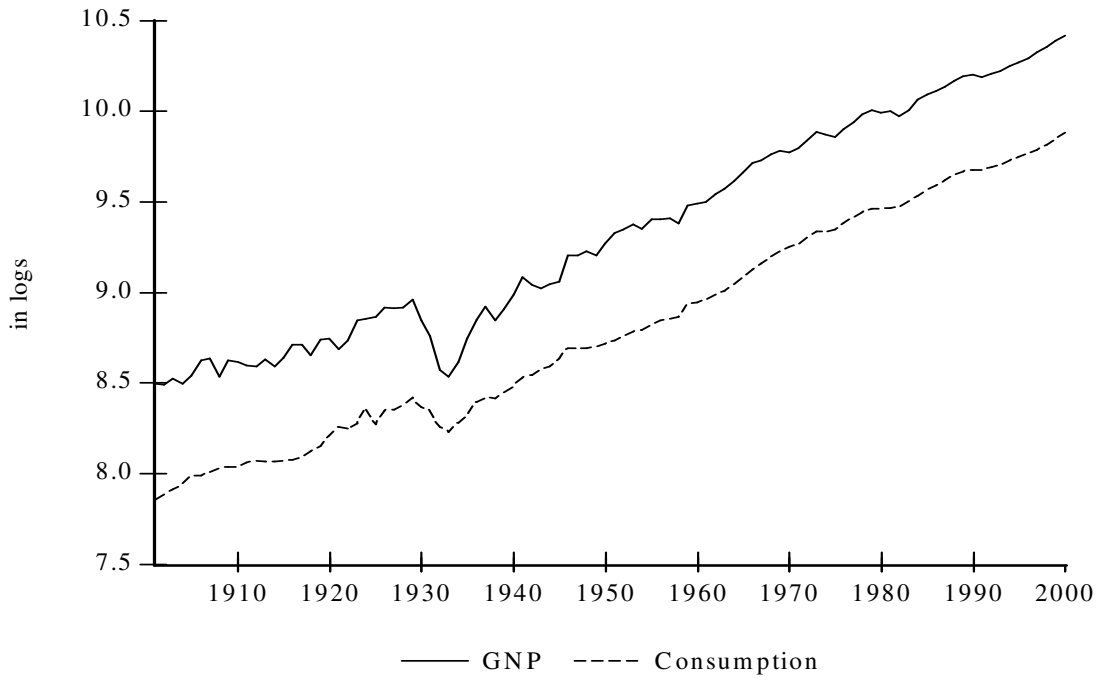


Figure 1

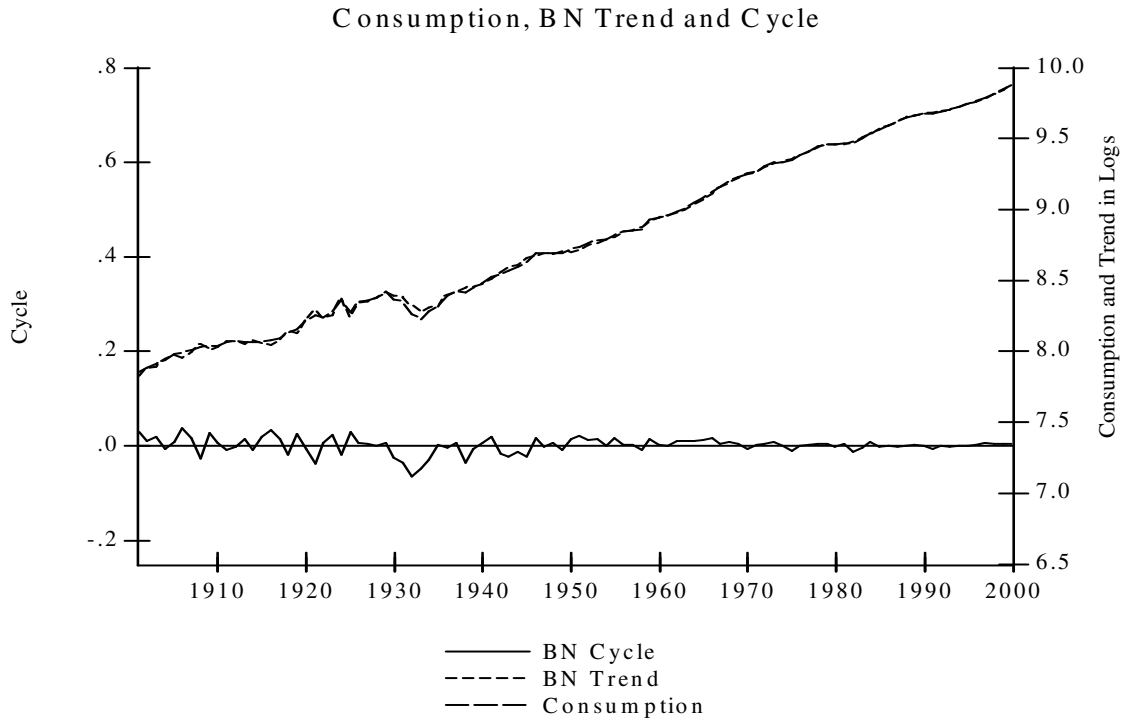


Figure 2

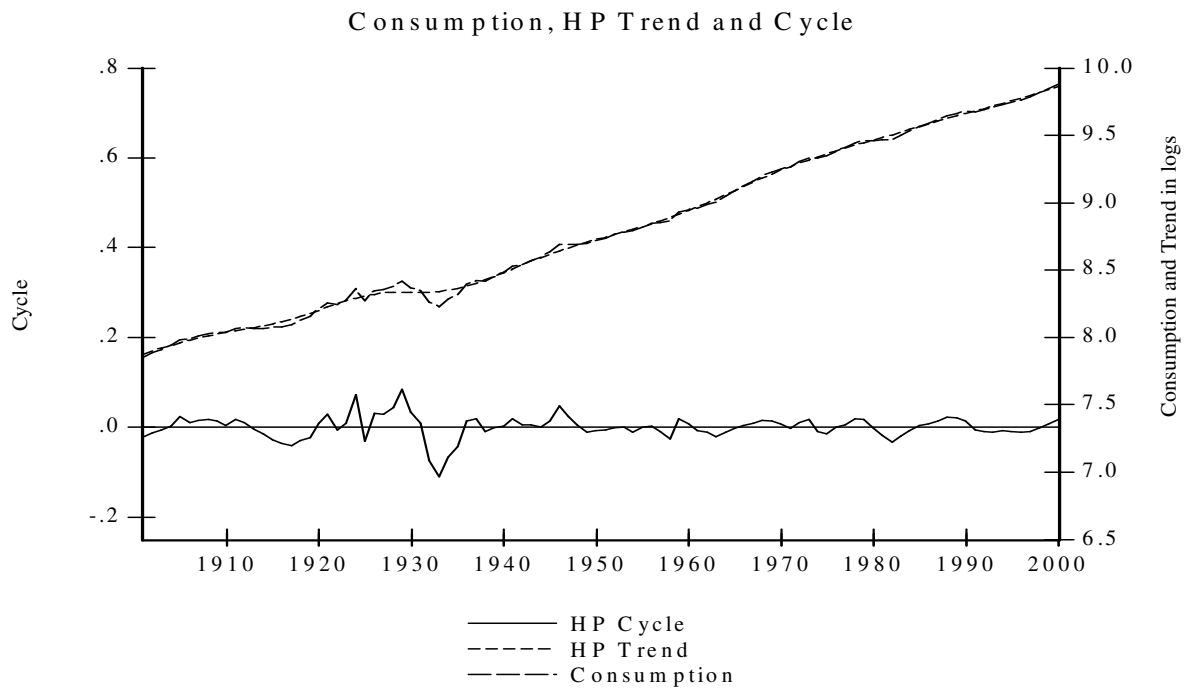


Figure 3

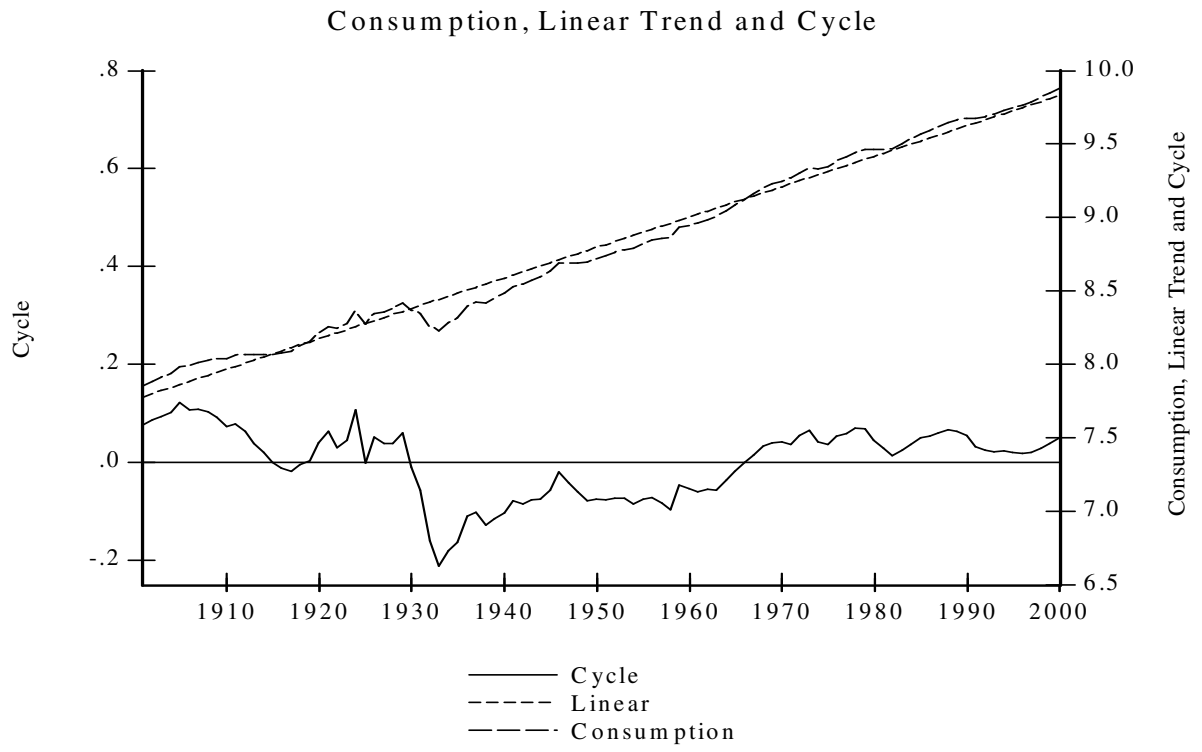


Figure 4

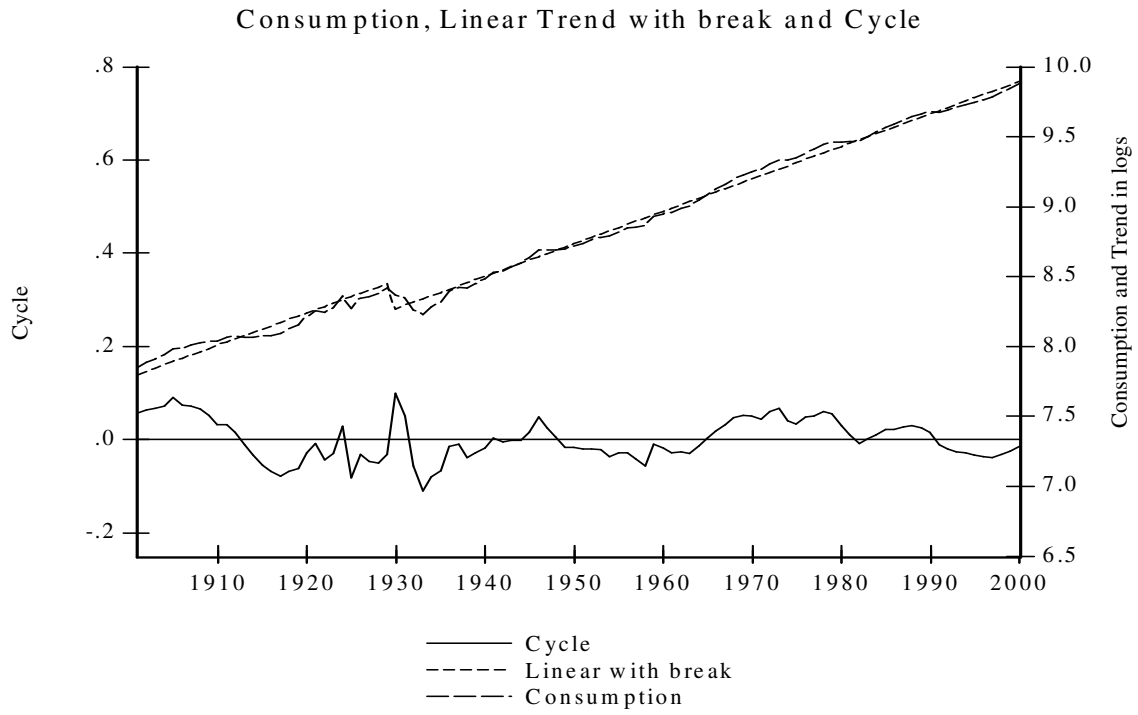


Figure 5