Testing Models of Low-Frequency Variability*

Ulrich K. Müller
Princeton University
Department of Economics
Princeton, NJ, 08544

Mark W. Watson
Princeton University
Department of Economics
and Woodrow Wilson School
Princeton, NJ, 08544

This version: October 2006
First draft: March 2006

Abstract

We develop a framework to assess how successfully standard time series models explain low-frequency variability of a data series. The low-frequency information is extracted by computing a finite number of weighted averages of the original data, where the weights are low-frequency trigonometric series. The properties of these weighted averages are then compared to the asymptotic implications of a number of common time series models. We apply the framework to twenty U.S. macroeconomic and financial time series using frequencies lower than the business cycle.

JEL classification: C22, E32

Keywords: Long Memory, Local-to-Unity, Unit Root Test, Stationarity Test, Business Cycle Frequency, Heteroskedasticity

*The first draft of this paper was written for the Federal Reserve Bank of Atlanta conference in honor of the twenty-fifth anniversary of the publication of Beveridge and Nelson (1981), and we thank the conference participants for their comments. We also thank Tim Bollerslev, David Dickey, John Geweke and Barbara Rossi for useful comments and discussions, and Rafael Dix Carneiro for excellent research assistance. Support was provided by the National Science Foundation through grants SES-0518036 and SES-0617811.

1 Introduction

Persistence and low-frequency variability has been an important and ongoing empirical issue in macroeconomics and finance. Nelson and Plosser (1982) sparked the debate in macroeconomics by arguing that many macroeconomic aggregates follow unit-root autoregressions. Beveridge and Nelson (1981) used the logic of the unit-root model to extract stochastic trends from macro series, and showed that variations in these stochastic trends were a large, sometimes dominant, source of variability in the series. Meese and Rogoff's (1983) finding, that random walk forecasts of exchange rates dominated other forecasts, focused attention on the unit root model in international finance. And in finance, interest in the random walk model arose naturally because of its relation to the efficient markets hypothesis (Fama (1970)).

This empirical interest led to the development of econometric methods for testing the unit root hypothesis, and for estimation and inference in systems that contain integrated series. More recently, the focus has shifted towards more general models of persistence, such as the fractional (or long memory) model and the local-to-unity autoregression, which nest the unit root model as a special case, or in the local level model which allows an alternative nesting of the $I(0)$ and $I(1)$ models. While these models are designed to explain low-frequency behavior of time series, fully parametric versions of the models have implications for higher frequency variation, and efficient statistical procedures thus exploit both low and high frequency variations for inference. This raises the natural concern about the robustness of such inference to alternative formulations of higher frequency variability. These concerns have been addressed by, for example, constructing unit-root tests using autoregressive models that are augmented with additional lags as in Said and Dickey (1984), or by using various nonparametric estimators for long-run covariance matrices and (as in Geweke and Porter-Hudak (1983) (GPH)) for the fractional parameter. As useful as these approaches are, there still remains a question of how successful these various methods are in controlling for unknown or misspecified high frequency variability.

This paper takes a different approach. It *begins* by specifying the low-frequency band of interest. For example, the empirical analysis presented in Section 4 focuses on frequencies lower than the business cycle, that is periods greater than eight years. Using this frequency cut-off, the analysis then extracts the low-frequency component of the series of interest by computing weighted averages of the data, where the weights are low-frequency trigonometric series. Inference about the low-frequency variability of the series is exclusively based on the properties of these weighted averages, disregarding other aspects of the original data. The number of weighted averages, say q, that capture the low-frequency variability is small in typical applications. For example, only $q = 13$ weighted averages almost completely capture the lower than business cycle variability in postwar macroeconomic time series (for any sampling frequency). This suggests basing inference on asymptotic approximations in which q is fixed as the sample size tends to infinity. Such asymptotics yield a q-dimensional multivariate Gaussian limiting distribution for the weighted averages, with a covariance matrix that depends on the specific model of low-frequency variability. Inference about alternative models or model parameters can thus draw on the well-developed statistical theory concerning multivariate normal distributions.

An alternative to the methods proposed here is to use time domain filters, such as band-pass or other moving average filters, to isolate the low-frequency variability of the data. The advantage of the transformations that we employ is that they conveniently discretize the low-frequency information of the original data into q data points, and they are applicable beyond the $I(0)$ models typically analyzed with moving average linear filters.

There are several advantages to focusing exclusively on the low-frequency variability components of the data. The foremost advantage is that many empirical questions are naturally formulated in terms of low-frequency variability. For example, the classic Nelson and Plosser (1982) paper asks whether macroeconomic series such as real GNP tend to revert to a deterministic trend over periods longer than the business cycle, and macroeconomic questions involving balanced growth involve the covariability of series over frequencies lower than the business cycle. Questions of potential mean-reversion in asset prices or real exchange rates are often phrased in terms of long "horizons" or low frequencies. Because the statistical models studied here were developed to answer these kinds of low-frequency questions, it is natural to evaluate the models on these terms.

In addition, large literatures have developed econometric methods in the local-to-unity framework, and also in the fractional framework. These methods are presumably valid only if, at a minimum, their assumed framework accurately describes the low-frequency behavior of the time series under study. The tests developed here may thus also be used as specification tests for the appropriateness of these methods. Other advantages, including robustness to high frequency misspecification and ease of implementation (because the methods rely on simple properties of the multivariate normal distribution), have already been mentioned.

An important caveat is that reliance on low-frequency methods will result in a loss of information and efficiency for empirical questions involving all frequencies. Thus, for example, questions about balanced growth are arguably properly answered by the approach developed here, while questions about martingale difference behavior involve a constant spectrum over all frequencies, and focusing only on low frequencies entails a loss of information.

Several papers have addressed other empirical and theoretical questions in similar frameworks. Bierens (1997) derives estimation and inference procedures for cointegration relationships based on a finite number of weighted averages of the original data, with a joint Gaussian limiting distribution. Phillips (2006) pursues a similar approach with an infinite number of weighted averages. Phillips (1998) provides a theoretical analysis of 'spurious regressions' of various persistent time series on a finite (and also infinite) number of deterministic regressors. Müller (2004) finds that long-run variance estimators based on a finite number of trigonometrically weighted averages is optimal in a certain sense. All these approaches exploit the known asymptotic properties of weighted averages for a given model of low-frequency variability. In contrast, the focus of this paper is to test alternative models of low-frequency variability and their parameters.

The plan of the paper is as follows. The next section introduces the three classes of models that we will consider: fractional models, local-to-unity autoregressions, and the local level model, parameterized as an unobserved components model with a large $I(0)$ component and a small unit root component. This section discusses the choice of weights for extracting the low-frequency components and the model-specific asymptotic distributions of the resulting weighted averages. Section 3 develops tests of the models based on these asymptotic distributions. Section 4 uses the methods of Section 3 to study the low-frequency properties of twenty macroeconomic and financial time series.[1] Section 5 offers some concluding remarks.

[1] As it turns out, slight modifications of the weighting function and test statistics of Sections 2 and 3 result

2 Models and Low-Frequency Transformations

Let y_t, $t = 1, \cdots, T$ denote the observed time series, and consider the decomposition of y_t into unobserved deterministic and stochastic components

$$y_t = d_t + u_t.$$

This paper focuses on the low-frequency variability of the stochastic component u_t; the deterministic component is modelled as a constant $d_t = \mu$, or as a constant plus linear trend $d_t = \mu + \beta t$, with unknown parameters μ and β.

We consider three leading models used in finance and macroeconomics to model low-frequency variability. The first is a fractional ("long-memory") model; stationary versions of the model have a spectral density $S(\lambda) \propto |\lambda|^{-2d}$ as $\lambda \to 0$, where d is the fractional parameter. We consider stationary and integrated versions of the model. The second model is the AR model with largest root close to unity; using standard notation we write the dominant AR coefficient as $\rho_T = (1 - c/T)$, so that the process is characterized by the local-to-unity parameter c. For this model, normalized versions of u_t converge in distribution to an Ornstein-Uhlenbeck process with diffusion parameter $-c$, and for this reason we will often refer to this model as the OU model. We consider level and integrated versions of the OU model. The third model that we consider decomposes u_t into an $I(0)$ and $I(1)$ component, $u_t = w_t + (g/T) \sum_{s=1}^{t} \eta_t$, where $(w_t, \eta_t)'$ are $I(0)$ with long-run covariance matrix $\sigma^2 I_2$, and g is a parameter that governs the relative importance of the $I(1)$ component. In this "Local Level" (LL) model (c.f. Harvey (1989)) both components are important for the low-frequency variability of u_t. Again, we consider level and integrated versions of the LL model.

2.1 Asymptotic Representation of the Models

As shown in Theorem 1 below, the low-frequency variability implied by each of these models can be characterized by the stochastic properties of the partial sum process for u_t, so for our purposes it suffices to define each model in terms of the behavior of these partial sums.

in considerable compuational simplifications. These simplifications, while not used in the empirical analysis reported in Section 4, may be useful to applied researchers and are described in the appendix. A description of the computational details together with replication files may be found under www.princeton.edu/∼mwatson.

Letting W denote a Wiener process and σ a generic non-zero scalar constant, the specific assumptions for each model, and their integrated counterparts are:[2]

1(a) **Stationary fractional model** (*FR*): u_t follows a stationary fractional model with parameter $-1/2 < d < 1/2$, where $T^{-1/2-d}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \sigma W^d(\cdot)$, where W^d is a "type I" fractional Brownian motion defined as $W^d(s) = A(d)\int_{-\infty}^{0}\left[(s-\lambda)^d - (-\lambda)^d\right] dW(\lambda) + A(d)\int_0^s (s-\lambda)^d dW(\lambda)$ and $A(d) = \left(\frac{1}{2d+1} + \int_0^\infty \left[(1+\lambda)^d - \lambda^d\right]^2 d\lambda\right)^{-1/2}$.

1(b) **Integrated fractional model**: u_t follows a fractional model with parameter $1/2 < d < 3/2$, when the first differences $u_t - u_{t-1}$ (with $u_0 = 0$) follow a stationary fractional model with parameter $d-1$.

2 **Local-to-unity model** (OU): $u_t = \rho_T u_{t-1} + \eta_t$, $\rho_T = 1 - c/T$ and $T^{-1/2}\sum_{t=1}^{[\cdot T]}\eta_t \Rightarrow \sigma W(\cdot)$. Assumptions about the initial condition, u_0, depend on whether the model is stable ($c > 0$) or not ($c \leq 0$). In the stable model, u_0 is drawn from the stationary limiting distribution and $T^{-1/2}u_{[\cdot T]} \Rightarrow \sigma J_c(\cdot)$, where J_c is the stationary Ornstein-Uhlenbeck (OU) process $J_c(s) = Ze^{-sc}/\sqrt{2c} + \int_0^s e^{-c(s-\lambda)}dW(\lambda)$, with $Z \sim \mathcal{N}(0,1)$ independent of W. In the unstable model ($c \leq 0$), $u_0 = 0$, and $T^{-1/2}u_{[sT]} \Rightarrow \sigma\int_0^s e^{-c(s-\lambda)}dW(\lambda)$.

3 **Integrated local-to-unity model** (*I-OU*): u_t follows an integrated local-to-unity model with parameter c if the first differences $u_t - u_{t-1}$ (with $u_0 = 0$) follow a local-to-unity model, where for simplicity we restrict the analysis to the stable model with $c > 0$.

4 **Local level model** (LL): u_t follows a local level model with parameter $g \geq 0$, when $u_t = w_t + \frac{g}{T}\sum_{s=1}^{t}\eta_s$, and $(T^{-1/2}\sum_{t=1}^{[\cdot T]} w_t, T^{-1/2}\sum_{t=1}^{[\cdot T]}\eta_t)' \Rightarrow \sigma(W_1(\cdot), W_2(\cdot))'$, where W_1 and W_2 are independent standard Wiener processes.

5 **Integrated local level model** (*I-LL*): u_t follows an integrated local level model with parameter $g \geq 0$ if the first differences $u_t - u_{t-1}$ (with $u_0 = 0$) follow a local level model with parameter g.

[2] Formally, the specifications (2)-(5) require u_t and y_t to be modelled as double arrays, but we omit any dependence on T to ease notation.

Table 1 summarizes the assumptions about convergence of the partial sum process for each model and provides a description of the covariance kernel of the limiting process. A large number of primitive conditions have been used to justify these limits. Specifically, for the stationary fractional model (1a), weak convergence to the fractional Wiener process W^d has been established under various primitive conditions for u_t by Taqqu (1975) and Chan and Terrin (1995)—see Marinucci and Robinson (1999) for additional references and discussion. Mandelbrot and Ness (1968) showed that W^d so defined has almost surely continuous sample paths. Model (1b) uses Velasco's (1999) definition of a fractional process for $1/2 < d < 3/2$. The local-to-unity model (2) and local level model (4) rely on a functional central limit theorem applied to $(w_t, \eta_t)'$; various primitive conditions are given, for example, in McLeish (1974), Wooldridge and White (1988), Phillips and Solo (1992), and Davidson (2002); see Stock (1994) for general discussion.

The unit root and $I(0)$ models are nested in several of the models in Table 1. The unit root model corresponds to the integrated fractional model (1b) with $d = 1$, the OU model (2) with $c = 0$, and the integrated local level model (5) with $g = 0$. Similarly, the $I(0)$ model corresponds to the stationary fractional model (1a) with $d = 0$ and the local level model (4) with $g = 0$.

The objective of this paper is to assess how well these specifications explain the low-frequency variability of u_t. Since the deterministic component d_t is unknown, we restrict attention to statistics that are functions of the least-square residuals of a regression of y_t on a constant (denoted u_t^μ) or on a constant and time trend (denoted u_t^τ). Because $\{u_t^i\}_{t=1}^T$, $i = \mu, \tau$ are maximal invariants to the groups of transformations $\{y_t\}_{t=1}^T \to \{y_t + m\}_{t=1}^T$ and $\{y_t\}_{t=1}^T \to \{y_t + m + bt\}_{t=1}^T$, respectively, there is no loss of generality in basing inference on functions of $\{u_t^i\}_{t=1}^T$ for tests that are invariant to these transformations.

We extract the information about the low-frequency variability of u_t by considering a fixed number (q) of weighted averages of u_t^i, $i = \mu, \tau$, where the weights are known and deterministic low-frequency trigonometric series. We discuss specific choices for these functions below, but first provide a general result about the joint asymptotic distribution of these q weighted averages. Here and below, the limits of integrals are understood to be zero and one, if not indicated otherwise.

Theorem 1 *Suppose there exists α and $\sigma > 0$ such that $T^{-\alpha}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \sigma G(\cdot)$, where G is a mean-zero Gaussian process with almost surely continuous sample paths and $k(r,s) = E[G(r)G(s)]$. Define*

$$\begin{aligned} k^{\mu}(r,s) &= k(r,s) - rk(1,s) - sk(r,1) + rsk(1,1) \\ k^{\tau}(r,s) &= k^{\mu}(r,s) - 6s(1-s)\int k^{\mu}(r,\lambda)d\lambda \\ &\quad -6r(1-r)\int k^{\mu}(\lambda,s)d\lambda + 36rs(1-s)(1-r)\int\int k^{\mu}(\nu,\lambda)d\nu d\lambda, \end{aligned}$$

and let $\Psi(\cdot) = (\Psi_1(\cdot), \cdots, \Psi_q(\cdot))'$, where $\Psi_l : [0,1] \mapsto \mathbb{R}$, $l = 1, \cdots, q$, are functions with continuous derivatives ψ_l. Then

$$X_T \equiv T^{-\alpha}\sum_{t=1}^{T} \Psi(t/T)u_t^i \Rightarrow X \sim \mathcal{N}(0, \sigma^2\Sigma)$$

where the j,lth element of Σ is given by $\int\int \psi_j(r)\psi_l(s)k^i(r,s)drds$.

The joint distribution of the q weighted averages of u_t^i, $i = \mu, \tau$ is thus asymptotically normal with a covariance matrix that is, up to scale, determined by the covariance kernel of G.

If X_T captures the information in y_t about the low-frequency variability of u_t, then the question of model fit for a specific stochastic model becomes the question whether X_T is approximately distributed $\mathcal{N}(0, \sigma^2\Sigma)$. For the models introduced above, $\Sigma = \Sigma(\theta)$ is a known function of the model parameter $\theta \in \{d, c, g\}$ for the models in Table 1, and σ^2 is an unknown constant governing the low-frequency scale of the process. (For example, σ^2 is the long-run variance of η_t in the local-to-unity model.) Because q is finite, that is our asymptotics keep q fixed as $T \to \infty$, it is not possible to estimate σ^2 consistently using the q elements in X_T. This suggests restricting attention to scale invariant tests of X_T. Imposing scale invariance has the additional advantage that the value of α in $X_T = T^{-\alpha}\sum_{t=1}^{T} \Psi(t/T)u_t^i$ does not need to be known.

Thus, consider the following maximal invariant to the group of transformation $X_T \to aX_T$, $a \neq 0$,

$$v_T = X_T/\sqrt{X_T'X_T}.$$

Under the conditions of Theorem 1, by the Continuous Mapping Theorem, $v_T \Rightarrow X/\sqrt{X'X}$. The density of $v = (v_1, \cdots, v_q)' = X/\sqrt{X'X}$ with respect to the uniform measure on the surface of a q dimensional unit sphere is given by (see, for instance, Kariya (1980) or King (1980))

$$f_v(\Sigma) = C|\Sigma|^{-1/2}\left(v'\Sigma^{-1}v\right)^{-q/2} \tag{1}$$

where the positive constant $C = \frac{1}{2}\Gamma(q/2)\pi^{-q/2}$, and $\Gamma(\cdot)$ is the Gamma function. For a given model for u_t, the asymptotic distribution of v_T depends only on q and the model parameter θ. Our strategy therefore is to base inference about the models in Table 1 using tests based on v_T, $\Sigma(\theta)$, and the probability density function (1).

2.2 Continuity of the Fractional and Local-to-Unity Models

Before discussing the choice of functions Ψ and test statistics, it is useful to take a short detour to discuss the continuity of $\Sigma(\theta)$ for two of the models. In the local-to-unity model, there is a discontinuity at $c = 0$ in our treatment of the initial condition and this leads to different covariance kernels in Table 1; similarly, in the fractional model there is a discontinuity at $d = 1/2$ as we move from the stationary to the integrated version of the model. As it turns out, these discontinuities do not lead to discontinuities of the density of v in (1) as a function of c and d.

This is easily seen in the local-to-unity model (2). Location invariance implies that it suffices to consider the asymptotic distribution of $T^{-1/2}(u_{[\cdot T]} - u_1)$. As noted by Elliott (1999), in the stable model $T^{-1/2}(u_{[\cdot T]} - u_1) \Rightarrow J^c(\cdot) - J^c(0) = Z(e^{-sc} - 1)/\sqrt{2c} + \int_0^s e^{-c(s-\lambda)} dW(\lambda)$, and $\lim_{c\to 0}(e^{-sc} - 1)/\sqrt{2c} = 0$, so that the asymptotic distribution of $T^{-1/2}(u_{[\cdot T]} - u_1)$ is continuous at $c = 0$.

The calculation for the fractional model is somewhat more involved. Note that the density (1) of v remains unchanged under reparametrizations $\Sigma \to a\Sigma$ for any $a > 0$. Because $\Sigma(d)$ is a linear function of $k^\mu(r,s)$, it therefore suffices to show that

$$\lim_{\epsilon \downarrow 0} \frac{k^\mu_{FR(d-\epsilon)}(r,s)}{k^\mu_{I\text{-}FR(d+\epsilon)}(r,s)} = a \tag{2}$$

for some constant $a > 0$ that does not depend on (r,s), where k_{FR} and k_{I-FR} are the covariance kernels for the stationary and integrated fractional models. As shown in the appendix, (2) holds with $a = 2$, so that the density of v is continuous at $d = 1/2$.[3]

[3]This result suggests a definition of a *demeaned* fractional process with $d = 1/2$ as any process whose

2.3 Choice of Ψ and the Resulting $\Sigma(\theta)$

Our choice of $\Psi = (\Psi_1, \cdots, \Psi_q)'$ is guided by two goals. The first goal is that Ψ should extract low-frequency variations of u_t and, to the extent possible, be uncontaminated by higher frequency variations. The second goal is that Ψ should produce a diagonal (or nearly diagonal) covariance matrix Σ, as this facilitates the interpretation of X_T and simplifies the testing problem. In particular, when Σ is diagonal, the models' implications for persistence in u_t become implications for specific forms of heteroskedasticity in X_T.

Many choices of Ψ (Fourier expansions, cosine transforms, etc.) do a good job extracting low-frequency variability, but do not produce diagonal Σ. (For example, see Akdi and Dickey (1998) for an analysis of the unit root model using Fourier expansions.). Eigenfunctions of the covariance kernel of the demeaned/detrended Wiener process do a good job on both fronts. By construction, these functions yield a diagonal Σ for the $I(1)$ model, and because they are orthogonal, the also yield a diagonal Σ in the $I(0)$ model. Thus, these functions yield a diagonal Σ for all values of g in the local level model. For the fractional model, the eigenfunctions produce a diagonal Σ when $d = 0$ and $d = 1$, and as we show below, a nearly diagonal Σ for other values of d. A similar result holds for the local-to-unity model. Moreover, because of the trigonometric form of the eigenfunctions, they can be used to isolate the low-frequency variation in the data.

Let $k_W^\mu(r,s)$ and $k_W^\tau(r,s)$ denote the covariance kernels of the demeaned and detrended Wiener process. The following theorem characterizes their eigenfunctions:

Theorem 2 *Let*

$$\begin{aligned}
\varphi_l^\mu(s) &= \sqrt{2}\cos(\pi l s), \ \textit{for } l \geq 1 \\
\varphi_l^\tau(s) &= \begin{cases} \sqrt{2}\cos(\pi s(l+1)) & \textit{for odd } l \geq 1 \\ \sqrt{\frac{2\omega_{l/2}}{\omega_{l/2}-\sin(\omega_{l/2})}}(-1)^{(l+2)/2}\sin(\omega_{l/2}(s-1/2)) & \textit{for even } l \geq 2 \end{cases}
\end{aligned}$$

partial sums converge to a Gaussian process with covariance kernel that is given by an appropriately scaled limit of k_{FR}^μ as $d \uparrow 1/2$; see equations (11) and (12) in the appendix. The possibility of a continuous extension across all values of d renders Velasco's (1999) definition of fractional processes with $d \in (1/2, 3/2)$ as the partial sums of a stationary fractional process with parameter $d-1$ considerably more attractive, as it does not lead to a discontinuity at the boundary $d = 1/2$, at least for demeaned data with appropriately chosen scale.

$\varphi_0^\mu(s) = \varphi_{-1}^\tau(s) = 1$ *and* $\varphi_0^\tau(s) = \sqrt{3}(1-2s)$, *where* ω_j *is the* j*th positive root of* $\cos(\omega/2) = 2\sin(\omega/2)/\omega$. *The set of orthonormal functions* $\{\varphi_l^\mu\}_{l=0}^\infty$ *and* $\{\varphi_l^\tau\}_{l=-1}^\infty$ *are the eigenfunctions of* $k_W^\mu(r,s)$ *and* $k_W^\tau(r,s)$ *with associated eigenvalues* $\{\lambda_l^\mu\}_{l=0}^\infty$ *and* $\{\lambda_l^\tau\}_{l=-1}^\infty$, *respectively, where* $\lambda_0^\mu = 0$ *and* $\lambda_l^\mu = (l\pi)^{-2}$ *for* $l \geq 1$ *and* $\lambda_{-1}^\tau = \lambda_0^\tau = 0$, $\lambda_l^\tau = (l\pi+\pi)^{-2}$ *for odd* $l \geq 1$ *and* $\lambda_l^\tau = (\omega_{l/2})^{-2}$ *for even* $l \geq 2$.

In the empirical analysis we use $\Psi_l(s) = \varphi_l^\mu(s)$ for demeaned series and $\Psi_l(s) = \varphi_l^\tau(s)$ for detrended series.

To see how well these functions extract low-frequency variability, consider the R^2 of a continuous time regression of a generic periodic series $\sin(\pi\vartheta s + \phi)$ on $\Psi_1(s), \cdots, \Psi_q(s)$, where $\vartheta \geq 0$ and $\phi \in [0, \pi)$. Ideally, this R^2 should equal unity for $\vartheta \leq \vartheta_0$ and zero for $\vartheta > \vartheta_0$, for all phase shifts $\phi \in [0,\pi)$, where ϑ_0 corresponds to the pre-specified cut-off frequency. With $\Psi_l = \varphi_l^\mu$ and $\vartheta_0 = q$ in the mean case and with $\Psi_l = \varphi_l^\tau$ and $\vartheta_0 = q+1$ in the trend case, regressing $\sin(\pi\vartheta s+\phi)$ on Ψ yields the following values of R^2

$$
\begin{aligned}
R_\mu^2 &= \frac{8\vartheta^3}{\pi(2\pi\vartheta + \sin(2\phi) - \sin(2(\phi+\vartheta\pi)))} \sum_{l=1}^{q} \frac{(\cos(\phi) - (-1)^l \cos(\phi+\vartheta\pi))^2}{(\vartheta^2 - l^2)^2} \\
R_\tau^2 &= \frac{8\vartheta}{2\pi\vartheta + \sin(2\phi) - \sin(2(\phi+\vartheta\pi))} \left[\frac{\vartheta^2}{\pi} \sum_{l=1}^{[q/2+1/2]} \frac{(\cos(\phi) - \cos(\phi+\vartheta\pi))^2}{(\vartheta^2 - 4l^2)^2} \right. \\
&\left. + 4\pi(\cos(\pi\vartheta/2+\phi))^2 \sum_{l=1}^{[q/2]} \omega_l \frac{(\omega_l \cos(\omega_l/2)\sin(\pi\vartheta/2) - \pi\vartheta\cos(\pi\vartheta/2)\sin(\omega_l/2))^2}{(\omega_l^2 - \pi^2\vartheta^2)^2(\omega_l - \sin\omega_l)} \right]
\end{aligned}
$$

(with these expressions extended by continuous limits at discontinuity points of ϑ). R_μ^2 and R_τ^2 converge to zero as $\vartheta \to \infty$ for all fixed values of q and ϕ, so that these choices for Ψ do not extract any high frequency information. Figure 1 depicts R_μ^2 as a function of ϑ for $\vartheta_0 = 14$. In the top panel, for each value of ϑ, R_μ^2 is averaged over all values for the phase shift $\phi \in [0,\pi)$, in the middle panel, R_μ^2 is maximized over ϕ and in the bottom panel, R_μ^2 is minimized over ϕ. The eigenfunctions come reasonably close to the ideal of extracting all information about cycles of frequency $\vartheta \leq \vartheta_0$ ($R^2 = 1$) and no information about cycles of frequency $\vartheta > \vartheta_0$ ($R^2 = 0$). Also shown in the figures are the R^2 using the Fourier expansions $\Psi_l(s) = \sqrt{2}\sin(\pi(l+1)s)$ for l odd and $\Psi_l(s) = \sqrt{2}\cos(\pi l s)$ for l even, and these have comparable performance. Results for R_τ^2 are similar, although they show lower R^2 values for small values of ϑ.

Table 2 summarizes the size of the off-diagonal elements of Σ for various values of $\theta \in \{d, c, g\}$ in the FR, OU and LL models. It presents the average absolute correlation when $\vartheta_0 = 15$, a typical value in the empirical analysis. The average absolute correlation is zero or close to zero for all considered parameter values.

Because Σ is (essentially) diagonal, the models can be compared by considering the diagonal elements of Σ. Figure 1 plots the square roots of these diagonal elements for the various models considered in Table 2. Evidently, more persistent models produce larger variances for low-frequency components, a generalization of the familiar 'periodogram' intuition that for stationary u_t, the variance of $\sqrt{2/T}\sum_{t=1}^T \cos(\pi l t/T)u_t$ is an approximately unbiased estimator of the spectral density at frequency $l/2T$. For example, for the unit root model ($d = 1$ in the fractional model or $c = 0$ in the local-to-unity model), the standard deviation of X_1 is 15 times larger than the standard deviation of X_{15}. In contrast, when $d = 0.25$ in the fractional model the relative standard deviation of X_1 falls to 2, and when $c = 5$ in the local-to-unity model, the relative standard deviation of X_1 is 7. In the $I(0)$ model ($d = 0$ in the fractional model or $g = 0$ in the local level model), $\Sigma = I_q$, and all of the standard deviations are unity.

3 Test Statistics

This section discusses several test statistics for the models. As discussed above, under the conditions of Theorem 1, the transformed data satisfies $v_T \Rightarrow v = X/\sqrt{X'X}$, where $X \sim \mathcal{N}(0, \Sigma)$. The low-frequency characteristics of the models are summarized by the covariance matrix Σ, so we derive optimal tests based on v against specific alternatives for Σ. Since $v_T \Rightarrow v$, these tests are asymptotically valid for a definition of the models as described in Table 1 above.

Three specific tests are discussed. First, relative persistence in the models is associated with the relative size of diagonal elements of Σ, so the first test focuses on the form of heteroskedasticity in X. Second, while the limits in Table 1 obtain also for certain forms of heteroskedasticity in u_t, the limits change when u_t has slowly varying second moments, and this leads to changes in Σ. Thus, the second test asks whether there is pronounced enough low-frequency heteroskedasticity in u_t as to invalidate the partial sum limits shown in Table 1 and Theorem 1. Third, because each of the models implies that the distribution of v is characterized by a single parameter, it is straightforward to derive point-optimal tests, and we provide such tests for the $I(1)$ and $I(0)$ models. The section concludes with a discussion of the potential for discriminating between the various models as such, that is without specifying specific parameter values.

3.1 Testing for Alternative Forms of Heteroskedasticity in X_T

Let Σ_0 denote the value of Σ under a particular null model and parameter θ_0. We consider tests against alternatives of the form $\Sigma = \Lambda\Sigma_0\Lambda$ where Λ is a diagonal matrix. The relative values of the diagonal elements of Λ amplify or attenuate the relative variance of the elements of X, and represent u_t-processes with a different persistence structure than what is implied by the null model. Power can be achieved for a range of alternatives by considering a range of values for Λ. A convenient way to specify the range of alternatives is to represent Λ as $\Lambda = \text{diag}(\exp(\delta_1), \cdots, \exp(\delta_q))$, where $\delta = (\delta_1, \cdots, \delta_q)'$ is a mean zero Gaussian vector with $E[\delta\delta'] = \gamma^2\Omega$. Under the null hypothesis $\gamma = 0$ and $\Sigma = \Sigma_0$, while under the alternative $\gamma \neq 0$, and the deviation from the null depends on the realization of δ. The covariance matrix Ω determines which kind of deviations are more likely. Modelling δ as a random vector allows the alternative to flexibly capture a wide range of specific alternatives. Conditional on δ, the alternative covariance matrix $\Lambda\Sigma_0\Lambda$ has the lth diagonal elements multiplied by $\exp(2\delta_l)$ compared to Σ_0, while the correlation structure remains unchanged.

Formally, consider the null and alternative hypotheses

$$\begin{aligned} &H_0 : v \text{ has density } f_v(\Sigma_0) \\ &H_1 : v \text{ has density } E_\delta f_v(\Lambda\Sigma_0\Lambda) \end{aligned} \tag{3}$$

where E_δ denotes integration over the measure of δ and f_v is defined in (1). Let e be a $q \times 1$ vector of ones, and ι_j the $q \times 1$ vector with a one in the jth row and zeros elsewhere. After calculations that closely mirror those of Nyblom (1989), one obtains that the locally best test at $\gamma = 0$ rejects for large values of

$$\text{LB} = (q/2+1)\frac{b'\Omega b}{(v'\Sigma_0^{-1}v)^2} + 2\frac{e'\Omega b}{v'\Sigma_0^{-1}v} - \frac{\text{tr}\, B\Omega}{v'\Sigma_0^{-1}v} \tag{4}$$

where b and B are a $q \times 1$ vector and a $q \times q$ matrix, respectively, with elements

$$\begin{aligned} b_j &= v_j \iota_j' \Sigma_0^{-1} v \\ B_{jl} &= \begin{cases} v_l v_j \iota_l' \Sigma_0^{-1} \iota_j \text{ for } l \neq j \\ v_j \iota_j' \Sigma_0^{-1} v + v_j^2 \iota_j' \Sigma_0^{-1} \iota_j \text{ for } l = j \end{cases} \end{aligned}$$

In our empirical analysis we have used several choices for Ω associated with stochastic processes for δ characterized by "breaks" of a random size, "trends" with random slopes, and random walk variation. All provided similar empirical conclusions, and to save space we will only present results for the test in which δ follows the demeaned random walk (denoted LBIM): $\delta_l = \tilde{\delta}_l - q^{-1} \sum_{j=1}^{q} \tilde{\delta}_j$, where $\tilde{\delta}_l = \tilde{\delta}_{l-1} + \varepsilon_l$ with $\tilde{\delta}_0 = 0$ and $\varepsilon_l \sim iid\mathcal{N}(0,1)$. The demeaning centers the alternative model for Σ at the null model, and also results in a simplification of the statistic because $e'\Omega = 0$ for a demeaned δ.

3.2 Testing for Low-Frequency Heteroskedasticity in u_t

Limiting results for partial sums like those shown in Table 1 are robust to time varying variances of the driving disturbances, as long as the time variation is a stationary short memory process; this implies that the values of Σ in Theorem 1 are similarly robust to such forms of heteroskedasticity. However, instability in the second moment of financial and macroeconomic data is often of quite persistent (e.g., Bollerslev, Engle, and Nelson (1994) and Andersen, Bollerslev, Christoffersen, and Diebold (2006), Balke and Gordon (1989), Kim and Nelson (1999) and McConnell and Perez-Quiros (2000)), and it is interesting to ask whether second moments of u_t exhibit enough low-frequency variability as to invalidate limits like those shown in Table 1. To investigate this, we nest each of the models considered thus far in a more general model that allows for such low-frequency heteroskedasticity, derive the resulting value of Σ for the more general model, and construct an optimal test using this as the alternative.

Thus, for each of the low-frequency models, consider a version of the model with low-frequency heteroskedastic driving disturbances in their natural moving average representations: let $h(\cdot)$ be a continuous function on the unit interval, and consider models for $\{u_t\}$ that satisfy $T^{-\alpha}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \sigma\tilde{G}(\cdot)$, where

$$\begin{aligned}
FR &: \begin{cases} \tilde{G}(s) = A(d)\int_{-\infty}^{0}((s-\lambda)^d-(-\lambda)^d)dW(\lambda) \\ \qquad\qquad +A(d)\int_0^s(s-\lambda)^d h(\lambda)dW(\lambda),\ d\in(-1/2,1/2) \\ \tilde{G}(s) = \frac{A(d-1)}{d}\int_{-\infty}^{0}((s-\lambda)^d-(-\lambda)^{d-1}(sd-\lambda))dW(\lambda) \\ \qquad\qquad +\frac{A(d-1)}{d}\int_0^s(s-\lambda)^d h(\lambda)dW(\lambda),\ d\in(1/2,3/2) \end{cases} \\
OU &: \begin{cases} \tilde{G}(s) = \frac{1}{c}\int_{-\infty}^{0}(e^{c\lambda}-e^{-c(s-\lambda)})dW(\lambda)+\frac{1}{c}\int_0^s(1-e^{-c(s-\lambda)})h(\lambda)dW(\lambda),\ c>0 \\ \tilde{G}(s) = \frac{1}{c}\int_0^s(1-e^{-c(s-\lambda)})h(\lambda)dW(\lambda),\ c\le 0 \end{cases} \\
LL &: \tilde{G}(s) = \int_0^s h(\lambda)dW_1(\lambda)+g\int_0^s(s-\lambda)h(\lambda)dW_2(\lambda),\ g\ge 0 \\
I\text{-}OU &: \begin{array}{l} \tilde{G}(s) = c^{-2}\int_{-\infty}^{0}(e^{-c(s-\lambda)}-(1-cs)e^{c\lambda})dW(\lambda) \\ \qquad\qquad +c^{-2}\int_0^s(e^{-(s-\lambda)}-c(s-\lambda)-1)h(\lambda)dW(\lambda),\ c>0 \end{array} \\
I\text{-}LL &: \tilde{G}(s) = \int_0^s(s-\lambda)h(\lambda)dW_1(\lambda)+\frac{1}{2}g\int_0^s(s-\lambda)^2h(\lambda)dW_2(\lambda),\ g\ge 0
\end{aligned} \tag{5}$$

With $h(s)=1$, these definitions for $\tilde{G}(s)$ yield $G(s)$ as defined in Table 1. Note that the function h only affects the stochastic component of $\tilde{G}(s)$ that stems from the in-sample innovations, but leaves unaffected terms associated with initial conditions, such as $\frac{1}{c}\int_{-\infty}^{0}(e^{-c(s-\lambda)}-e^{c\lambda})dW(\lambda)$ in the stable local-to-unity model. The idea is that $h(t/T)$ describes the square root of the time varying long-run variance of the in-sample driving disturbances at date $t\ge 1$, while maintaining the assumption that stable models were stationary prior to the beginning of the sample. This restriction allows to write the covariance kernel of $\tilde{G}(s)$ as the sum of two pieces, and the one that captures the pre-sample innovations remains unaffected by h. Especially in the fractional model, such a decomposition is computationally convenient, as noted by Davidson and Hashimadze (2006).

For any of the models and any continuous function h, it is possible to compute the covariance kernel for $\tilde{G}$, and via Theorem 1, the covariance matrix of X. For example, suppose that u_t is $I(0)$ and $h(s)=\sqrt{1+2a\cos(\pi s)}$ and $|a|<1/2$. For this process, the j,lth element of Σ is given by $\int\Psi_j(s)\Psi_l(s)h^2(s)ds$, and in the mean case with $\Psi_l(s)=\sqrt{2}\cos(\pi ls)$

for $l = 1, \cdots, q$, we obtain

$$\Sigma = \begin{pmatrix} 1 & a & 0 & \cdots & 0 & 0 \\ a & 1 & a & \cdots & 0 & 0 \\ 0 & a & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & 1 \end{pmatrix}.$$

Evidently low-frequency heteroskedasticity in u_t leads to autocorrelations in X, and for this example the autocorrelation has the form of an MA(1) model.

Let $\Sigma(\theta_0, h)$ denote the value of Σ associated with a model with parameter θ_0 and heteroskedasticity function h. The homoskedastic versions of the models from Table 1 then lead to $\Sigma = \Sigma(\theta_0, 1)$ while their heteroskedastic counterparts lead to $\Sigma = \Sigma(\theta_0, h)$. The power of a test of the null hypothesis $\Sigma = \Sigma(\theta_0, 1)$ against the alternative $\Sigma = \Sigma(\theta_0, h)$ depends on the assumed form of the function h. To produce a test with good power for a wide range of h, we consider a flexible model for h in which $h = e^{\kappa W^*}$, where W^* is a standard Wiener process on the unit interval independent of G, and κ is a parameter. Thus, we consider the hypotheses

$$\begin{aligned} &H_0 : v \text{ has density } f_v(\Sigma(\theta_0, 1)) \\ &H_1 : v \text{ has density } E_{W^*} f_v(\Sigma(\theta_0, e^{\kappa W^*})). \end{aligned} \tag{6}$$

The constant κ governs whether tests maximize power against models with very pronounced low-frequency heteroskedasticity (κ large) or model with barely noticeable low-frequency heteroskedasticity (κ small). By the Neyman-Pearson Lemma and the form of f_v (1), an optimal test of (6) rejects for large values of

$$H = \frac{E_{W^*}[|\Sigma(\theta_0, e^{\kappa W^*})|^{-1/2}[v'\Sigma(\theta_0, e^{\kappa W^*})^{-1}v]^{-q/2}]}{[v'\Sigma(\theta_0, 1)^{-1}v]^{-q/2}}. \tag{7}$$

In the empirical section below, we implement this test with $\kappa = 1.3$. This value is motivated by the observation that for the mean case and $q = 15$, the 10% level optimal test with $\kappa = 1.3$ achieves power of approximately 50% against the alternative for which it is optimal.

3.3 Low-Frequency POI Tests for the $I(0)$ and $I(1)$ Models

Finally, we test the $I(0)$ and $I(1)$ null hypotheses using low-frequency point-optimal tests. Specifically, in the context of the local-to-unity model we test the unit root model $c = c_0 = 0$ against the alternative model with $c = c_1$ using the likelihood ratio statistic

$$\text{LFUR} = v'\Sigma(c_0)^{-1}v/v'\Sigma(c_1)^{-1}v$$

where the values of c_1 are those suggested by Elliott, Rothenberg, and Stock (1996) ($c_1 = 7$ for demeaned series and $c_1 = 13.5$ for detrended series). We label the statistic LFUR as a reminder that it is a low-frequency unit root test statistic.

We similarly test the $I(0)$ null hypothesis against the point alternative of a local level model with parameter $g = g_1 > 0$ (which is the same nesting of the $I(0)$ model as employed in Nyblom (1989) and Kwiatkowski, Phillips, Schmidt, and Shin (1992)). A calculation shows that the likelihood ratio statistic rejects for large values of

$$\text{LFST} = \left(\sum_{l=1}^{q} v_l^2\right) / \left(\sum_{l=1}^{q} \frac{v_l^2}{1 + g_1^2\lambda_l}\right)$$

where λ_l are the eigenvalues defined in Theorem 2. We follow Stock (1994) and set $g_1 = 8$ in the mean case and $g_1 = 13$ in the trend case.

3.4 Discrimination Between Models

So far, we have discussed tests that seek to establish whether a low-frequency model with a specific parameter value is a plausible data generating mechanism for the transformed data v_T. Alternatively, one might ask whether a model as such, with unspecified parameter value, is rejected in favor of another model. A large number of inference procedures have been developed for specific low-frequency models, such as the local-to-unity model and the fractional model. Yet, typically there is considerable uncertainty about the appropriate low-frequency model for a given series. A high-power discrimination procedure would therefore have obvious practical appeal.

In the following, we focus on the problem of discriminating between the three continuous bridges between the $I(0)$ and the $I(1)$ model: the fractional model with $0 \leq d \leq 1$, the local-to-unity model with $c \geq 0$ and the local level model with $g \geq 0$. These models are obviously similar in the sense that they all nest (or arbitrarily well approximate) the $I(0)$ and $I(1)$ model. More interestingly, a recent literature has pointed out that (non-degenerate) regime switching models and fractional models are similar along many dimensions—see, for example, Parke (1999), Diebold and Inoue (2001), and Davidson and Sibbertsen (2005). Since the local level model can be viewed as a short memory model with time varying mean, this question is closely related to the similarity of the fractional model with $0 < d < 1$ and the local level model with $g > 0$.

This suggests that it will be challenging to discriminate between low-frequency models using information contained in v_T. A convenient way to quantify the difficulty is to compute the total variation distance between the models. Recall that the total variation distance between two probability measures is defined as the largest absolute difference the two probability measures assign to the same event, maximized over all possible events. Let Σ_0 and Σ_1 be the covariance matrices of X induced by two models and specific parameter values. Using a standard equality (see, for instance, Pollard (2002), page 60), the total variation distance between the two probability measures described by the densities $f_v(\Sigma_0)$ and $f_v(\Sigma_1)$ is given by

$$\mathrm{TVD}(\Sigma_0, \Sigma_1) = \tfrac{1}{2} \int |f_v(\Sigma_0) - f_v(\Sigma_1)| d\eta$$

where η is the uniform measure on the surface of a q dimensional unit sphere. There is no obvious way to analytically solve this integral, but it can be evaluated using Monte Carlo integration. To see how, write

$$\begin{aligned} \mathrm{TVD}(\Sigma_0, \Sigma_1) &= \int \mathbf{1}[f_v(\Sigma_1) < f_v(\Sigma_0)](f_v(\Sigma_0) - f_v(\Sigma_1)) d\eta \\ &= \int \mathbf{1}[\mathrm{LR} < 1](1 - \mathrm{LR}) f_v(\Sigma_0) d\eta \end{aligned} \quad (8)$$

where $\mathrm{LR} = f_v(\Sigma_1)/f_v(\Sigma_0)$. Thus, $\mathrm{TVD}(\Sigma_0, \Sigma_1)$ can be approximated by drawing v's under $f_v(\Sigma_0)$ and averaging the resulting values of $\mathbf{1}[\mathrm{LR} < 1](1 - \mathrm{LR})$.[4]

[4] It is numerically advantageous to rely on (8) rather than on the more straightforward expression $\mathrm{TVD}(\Sigma_0, \Sigma_1) = \frac{1}{2} \int |1 - \mathrm{LR}| f_v(\Sigma_0) d\eta$ for the numerical integration, since $\mathbf{1}[\mathrm{LR} < 1](1 - \mathrm{LR})$ is bounded and thus possesses all moments, which is not necessarily true for $|1 - \mathrm{LR}|$.

Let $\Sigma_i(\theta)$ denote the covariance matrix of X for model $i \in \{FR, OU, LL\}$ with parameter value θ, and consider the quantity

$$D_{i,j}(\theta) = \min_{\gamma \in \Gamma} \mathrm{TVD}(\Sigma_i(\theta), \Sigma_j(\gamma))$$

where $\Gamma = [0, 1]$ for $i = FR$ and $\Gamma = [0, \infty)$ for $i \in \{OU, LL\}$. If $D_{i,j}(\theta)$ is small, then there is a parameter value $\gamma_0 \in \Gamma$ for which the distribution of v with $\Sigma = \Sigma_j(\gamma_0)$ is close to the distribution of v with $\Sigma = \Sigma_i(\theta)$, so it will be difficult to discriminate model i from model j if indeed $\Sigma = \Sigma_i(\theta)$. More formally, consider any model discrimination procedure between models i and j based on v, which correctly chooses model i when $\Sigma = \Sigma_i(\theta)$ with probability p. By definition of the total variation distance, the probability of the event "procedure selects model i" under $\Sigma = \Sigma_j(\gamma)$ is at least $p - \mathrm{TVD}(\Sigma_i(\theta), \Sigma_j(\gamma))$. If $D_{i,j}(\theta)$ is small, then either the probability of mistakenly selecting model i is large for some $\Sigma = \Sigma_j(\gamma_0)$, $\gamma_0 \in \Gamma$, or the probability p of correctly selecting model i is small. In the language of hypothesis tests, for any test of the the null hypothesis that $\Sigma = \Sigma_j(\gamma)$, $\gamma \in \Gamma$ against the alternative that $\Sigma = \Sigma_i(\theta)$, $\theta \in \Theta$, the sum of the probabilities of Type I and Type II error are bounded below by $1 - \max_{\theta \in \Theta} D_{i,j}(\theta)$.

The value of $D_{i,j}(\theta)$ is an (increasing) function of q. Figure 2 plots $D_{i,j}(\theta)$ for each of the model pairs for $q = 15$, which corresponds to 60 years of data with interest focused on frequencies lower than 8-year cycles. Panel (a) plots $D_{FR,OU}(d)$ and $D_{FR,LL}(d)$ and panels (b) and (c) contain similar plots for the OU and LL models. Evidently $D_{i,j}(\theta)$ is small throughout. For example, for all values of d, the largest distance of the fractional model to the local-to-unity and local level model is less than 30%, and the largest distance between the OU and LL models is less than 50%. For comparison, the total variation distance between the $I(0)$ and $I(1)$ model for $q = 15$ is about 92%. Total variation distance using detrended data is somewhat smaller than the values shown in Figure 2. Evidently then, it is impossible to discriminate between these standard models with any reasonable level of confidence using sample sizes typical in macroeconomic applications, at least based on the below business cycle variability in the series summarized by v_T. Indeed, to obtain, say, $\max_{0 \leq d \leq 1} D_{FR,OU}(d) \approx 0.9$, one would need a sample size of 480 years (corresponding to $q = 120$).

4 Empirical Results

4.1 Data

In this section we study twenty macroeconomic and financial time series using the low-frequency methods discussed in the last section. We analyze post-war quarterly versions of important macroeconomic aggregates (real GDP, aggregate inflation, nominal and real interest rates, productivity, and employment) and longer annual versions of related series (real GNP from 1869-2004, nominal and real bond yields from 1900-2004, and so forth). We also study several cointegrating relations by analyzing differences between series (such as long-short interest rate spreads) or logarithms of ratios (such as consumption-income or dividend-price ratios). A detailed description of the data is given in the Appendix. As usual, several of the data series are transformed by taking logarithms, and as discussed above, the deterministic component of each series is modeled as a constant or a linear trend. Table A.1 summarizes these transformations for each series.

Figure 4 shows a three-panel plot for each series. The first two panels show times series plots of the demeaned/detrended values of the series (u_t^μ or u_t^τ) as appropriate, and the first differences of the series. The third panel shows plots v_T, the low-frequency transformations of the series, where q, the number of elements in v_T, was chosen to isolate frequencies lower than the business cycle. Using the standard 6-32 quarter definition of business cycle periodicity, this means that attention is restricted to frequencies lower than $2\pi/32$ for quarterly series and $2\pi/8$ for annual series. The post-war quarterly series span the period 1952:1-2005:3, so that $T = 215$, and $q = [2T/32] = 13$ for the demeaned series and $q = 12$ for the detrended series. Each annual time series is available for a different sample period (real GNP is available from 1869-2004, while bond rates are available from 1900-2004, real exchange rates from 1791-2004, for example), so the value of q is series-specific. One series (returns on the SP500) contains daily observations from 1928-2005, and for this series $q = 17$.

4.2 Results

The relatively short sample (less than 60 years of data for many of the series), makes it impossible to carry out sharp statistical inference about model parameters. That is, because of the nature of the data, confidence sets will often contain a wide ranges of values for d, c, and g. With this in mind, the empirical analysis is guided by four key questions:

1. (a) Is the unit root model ($d = 1$ in the fractional model, $c = 0$ in the local level model, $g = 0$ in the integrated local level model) consistent with data?

   (b) Is the $I(0)$ model ($d = 0$ in the fractional model, $g = 0$ in the local level model) consistent with data?

2. Are some models rejected for all parameter values? If so, does this arise because they inadequately describe the persistence in the data (that is, provide poor fits for the diagonal elements of Σ), or rather because they ignore low-frequency heteroskedasticity (that is, provide poor fits for the off-diagonal elements of Σ)?

3. Which models fit the data better?

4. Are inferences about the models based on the low-frequency components of the data similar to inferences from standard methods (unit root tests, estimators of the fractional parameter d, and so forth)?

The empirical results for all twenty series are summarized in Tables 3-4 and Figures 5-6. These tables and figures are organized to provide answers to the four key questions. Table 3 shows p-values for tests of the $I(0)$ and $I(1)$ models. Results are shown for each series and for the LBIM, H and LFUR and LFST tests. Figure 5 plots the p-values for the LBIM and H tests for each series, for each of the five models and for a fine grid of parameter values: $-0.49 \leq d \leq 1.49$ for the fractional model, $0 \leq c \leq 30$ for the local-to-unity model, and $0 \leq g \leq 30$ for the local level model. Figure 6 plots the log-likelihood values for each series and for each of the five models. Finally, results for standard statistical tests and estimators are summarized in Table 4. This table shows p-values for the DFGLS unit-root test of Elliott, Rothenberg and Stock (1996) and the stationarity test of Nyblom (1989) (using a HAC covariance matrix as suggested in Kwiatkowski, Phillips, Schmidt, and Shin (1992)). It also shows estimated values of d and standard errors from the sort of regressions suggested in Geweke and Porter-Hudak (1983) (GPH). The GPH-regression estimators and standard errors are implemented as described in Robinson (2003).

The remainder of this section discusses the empirical results for each of the series.

Real GDP/GNP. The post-war quarterly real GDP data are consistent with a unit-root model, but not the $I(0)$ model. From Table 3, the p-values is 0.02 for the LBIM statistic for the $I(0)$ and the p-value for the LFST statistic is similarly small. In contrast, p-values for the test statistics for the $I(1)$ null are large. Heteroskedasticity is evident in the plot for $(1-L)y_t$ shown in Figure 4 (associated with the decrease in volatility in the post-1983 period), but evidently this heteroskedasticity is not so severe that the $I(1)$ model is rejected using the H statistic. From Figure 5, confidence intervals for the persistent parameters in the models are wide: The LBIM 90% confidence intervals include all values of d greater than 0.16 in the fractional model, all values of c considered in the the local-to-unity (OU) model, and values of g greater than 5.5 in the local level model. As discussed above, these wide confidence intervals are associated with the limited amount of low-frequency information in the 54-year sample period. Figure 6 indicates that the fractional model with $d = 0.75$, the OU model with $c = 7$, the local level model with $g = 30$, and the integrated local level model with $g = 0$, provide roughly equivalent fits to the models. (The log-likelihood values for these models differ by less than 0.2.) Finally, Table 4 suggests that similar conclusions would be reached used a battery of standard procedures: the unit root null is not rejected by the DFGLS test, the $I(0)$ null is rejected by the Nyblom/KPSS test, and GPH regressions yield point estimates of d similar to the low-frequency MLE, although the GPH confidence intervals are narrower than those obtained using the LBIM test. However, the GPH regressions use only $[n^{0.5}] = 14$, and $[n^{0.65}] = 32$ observations, so that these confidence intervals might have less coverage than suggested by asymptotic theory.

The results are different using the annual observations on real GNP from 1869-2004. From Table 3, both the $I(1)$ and $I(0)$ models are rejected for this series. Indeed, from Figure 5, the H-statistic rejects *all* parameter values for *all* of the models. Apparently, the low-frequency heteroskedasticity in the series is so severe, that the limits in Table 1 and Theorem 1 are not relevant for the long-annual GNP series. This heteroskedasticity is evident in Figure 4 and coincides with the post-World War II decline in volatility, and the resulting serial correlation in v_T is also evident in the figure. This conclusion—that low-frequency heteroskedasticity in the time series is so severe that it leads to rejection of the models—will be repeated for several of the series studied here. Finally, Table 4 shows that standard statistics are inconclusive about the appropriate process for this series. Neither the unit root or stationarity tests reject at the 5% level, and the GPH statistics are rather nonsensical.

Inflation. The unit-root model for inflation is not rejected using the post-war quarterly data, while the $I(0)$ model is rejected. Results are shown for inflation based on the GDP deflator, but similar conclusions follow from the PCE deflator and CPI. Stock and Watson (2005) document instability in the "size" of the unit root component (corresponding to the value of g in the local level model) over the post-war period, but apparently this instability is not so severe that it leads to rejections based on the tests considered here. Figure 5 shows that the fractional model with $d \geq 0.4$, the OU model with $c \leq 15$, and local level model with $g \geq 5$ are not rejected, and Figure 6 shows that models with $d = 0.8$, $c = 5$, and $g = 30$, provide comparably good fits. Table 4 indicates that the same qualitative results follow from standard methods.

Quite different results are obtained from the long-annual (1869-2004) series. Notably, the long-annual series shows less persistence than the postwar quarterly series: Figure 6 shows that the best fitting models are the fractional model with $d = 0.30$ and the local level model with $g = 30$; Table 3 shows that both the $I(0)$ and $I(1)$ models are rejected; and Figure 4 shows that the LBIM statistics reject the OU model for all values of $c < 30$, and the LBIM confidence set for d is $0.04 \leq d \leq 0.49$. Figure 4 shows pronounced heteroskedasticity in the series, again associated with postwar decrease in volatility. This volatility leads a rejection of many of the models using the H-statistic, and Figure 4 shows essentially no overlap in the LBIM and H confidence sets. Again, the severe low-frequency heteroskedasticity in the series yields statistics that are not consistent with the standard partial sum limits shown in Table 1.

Labor productivity and employee hours. Labor productivity is very persistent. The $I(0)$ model is rejected but the $I(1)$ model is not. Figure 5 shows that values of d less than 0.95 and values of c greater than 3 are rejected by the LBIM test. From Figure 6, the best fitting model is the *I-OU* model with $c = 12$ which fits the low-frequency data slightly better than the fractional model with $d = 1.49$ and much better than the OU and local level models. This persistence is evident in the plot of the first differences in Figure 4: there are long-swings in trend productivity growth in the post-war period associated with the productivity slowdown of the 1970s and 1980s and the productivity rebound of the 1990s. The standard statistics reported in Table 4 understate this persistence; for example, the GPH confidence intervals for d are more concentrated around the unit root model.

The behavior of employee hours per capita has received considerable attention in the recent VAR literature (see Gali (1999), Christiano, Eichenbaum, and Vigfusson (2003), Pesavento and Rossi (2005), and Francis and Ramey (2006a)). The results shown here are consistent with unit-root but not $I(0)$ low-frequency behavior. This result is evident from the low-frequency results summarized in Table 3 and Figure 5, and from the standard procedures in Table 4. Francis and Ramey (2006b) discuss demographic trends that are potentially responsible for the high degree of persistence in this series.

Interest rates. Postwar nominal interest rates are consistent with a unit-root but not an $I(0)$ process. Figure 4 shows heteroskedasticity in the series, most notably an increase in the volatility of long-rates in the second half of the sample (see Watson (1999) for discussion) and this leads to low p-values for the H-statistic for many models. The long annual bond rates (1900-2004) show even more heteroskedasticity, and this, together with the longer sample period for the long-annual data, yields H-statistics with p-values that are essentially zero for all the models considered. Again, low-frequency heteroskedasticity is so pronounced that the overall low-frequency behavior of this series is not well described by any of the models.

The results for real interest rates are similar. The statistics that focus on persistence, such as the LBIM tests or the statistics reported in Table 4, suggest relatively little persistence in u_t. (For example, the $I(0)$ model not rejected.) However, there is evidence for low-frequency heteroskedasticity in these series, and there is little overlap in the H-statistic and LBIM-statistic confidence intervals.

Real exchange rates. A large empirical literature has examined the unit root or near unit root behavior of real exchange rates. The data used here—annual observations on the real dollar/pound real exchange rate from 1791-2004—come in large part from one important empirical study in this literature, Lothian and Taylor (1996). Table 4 shows that standard tests reject both the $I(0)$ and unit root models, and GPH regressions suggest values for d around 0.5. Table 3 and Figure 5 shows that the low-frequency LBIM tests yield similar conclusions. Figure 5 shows that local-to-unity models with large values of c are not rejected by the LBIM statistic, but Figure 6 shows that these model fits the data poorly relative to the fractional model with d close to 0.5 or a local level model with a reasonably large random walk component ($g = 30$). The H-statistics reject for few models, suggesting that low-frequency heteroskedasticity in these real exchange rates is not very pronounced.

Cointegrating relations. Several of the data series, such as the spread between 10-year and 1-year Treasury bond rates, represent error correction terms from putative cointegrating relationships. Under the hypothesis of cointegration, these series should be $I(0)$. The $I(0)$ model is not rejected for long-short interest rate spread, and models with little persistence (the fractional model with values of d close to zero or the local level model with small values of g) provide the best fits to the low-frequency components of the series. This is not the conclusion that would be reached using standard tests: from Table 4 the Nyblom/KPSS statistic has a p-value of only 0.01, and the confidence interval for d depends critically on whether $[n^{0.5}]$ or $[n^{0.65}]$ observations are used in the regression.

Real unit labor costs (the logarithm of the ratio of labor productivity to real wages, $y - n - w$ in familiar notation) exhibit limited persistence: the $I(1)$ model is rejected by the LBIM test, but the $I(0)$ model is not rejected, and models with a low degree of persistence provide the best fits. However, the LFST statistic has a p-value of only 0.01 providing some evidence against the $I(0)$ null. That said, looking across all of the results, models with low persistence are not rejected, and a cointegration model for $y - n - w$ appears generally consistent with the data.

The "balanced growth" cointegrating relation between consumption and income (e.g., King, Plosser, Stock, and Watson (1991)) fares less well, where the $I(1)$ model is not rejected, but the $I(0)$ model is rejected. This $I(1)$ characterization of the series is consistent with the low-frequency variation in the series summarized in Table 3 and Figures 5-6, and with results from the standard statistics reported in Table 4. The apparent source of this rejection is the large increase in the consumption-income ratio over the 1985-2004 period, a subject that has attracted much recent attention (for example, see Lettau and Ludvigson (2004) for an explanation based on increases in asset values.) The investment-income relationship also appears to be at odds with the null of cointegration, although this rejection depends in part on the particular series used for investment and its deflator.

Finally, the stability of the logarithm of the earnings-stock price ratio or dividend-price ratio, and the implication of this stability for the predictability of stock prices, has been an ongoing subject of controversy (see Campbell and Yogo (2006) for a recent discussion). Using Campbell and Yogo's (2006) annual data for the SP500 from 1880-2002, both the $I(0)$ and $I(1)$ models are rejected. Confidence intervals constructed using the LBIM statistic suggest less persistence than a unit root (for example the LBIM confidence interval for the fractional model includes $0.38 \leq d \leq 0.89$). However the low-frequency heteroskedasticity in the series leads to a rejection of essentially all of the models using the H statistic. The shorter (1928-2004) CRSP dividend-yield (also from Campbell and Yogo (2006)), displays more low-frequency persistence, less heteroskedasticity, and is consistent with the $I(1)$ model but not the $I(0)$ model.

Volatility of stock returns. Ding, Granger, and Engle (1993) analyzed the absolute value of daily returns from the SP500 and showed that the autocorrelations decayed in a way that was remarkably consistent with a fractional process. Low frequency characteristics of the data summarized in Figures 5 and 6 are consistent with this finding. Both the unit-root and $I(0)$ models are rejected by the LBIM statistic, but models with somewhat less persistence than the unit root, such as the fraction model with $0.13 < d < 0.73$, are not rejected. The GPH statistics (which are now based on a large number of observations, $n = 20643$ so that $[n^{0.5}] = 143$, and $[n^{0.65}] = 637$) suggest some instability across frequencies: $\widehat{d} = 0.38$ using $[n^{0.65}]$ and $\widehat{d} = 0.46$ using $[n^{0.5}]$, and Figure 6 shows the best fitting model based on the low-frequency data has $d = 0.48$. This suggests an important role for the role of the frequency cut-off for the analysis, a point made by Andersen and Bollerslev (1997) in the context of volatility modeling and by Bollerslev and Mikkelsen (1999) in their study of long-term equity anticipation securities (LEAPS) on the SP500. Low-frequency changes in the volatility in the series are evident in Figure 4, and Figure 5 shows that the H-test rejects all of the models considered.

5 Conclusions

Standard specification tests for time series examine a model's appropriateness over the whole spectrum. In contrast, the methodology developed here isolates a model's low-frequency implications by focusing exclusively on the properties of a finite number of weighted averages of the original data. For example, by choosing the weights as trigonometric series with periods larger than eight years, our empirical analysis considers whether any of five standard models of persistence successfully explain the variability of twenty macro and financial time series at frequencies lower than the business cycle.

Three main findings stand out. First, despite the narrow focus, very few of the series are compatible with the $I(0)$ model. This hold true even for some putative cointegration error correction terms. Most macroeconomic series and relationships thus exhibit pronounced non-trivial dynamics below business cycle frequencies. In contrast, the unit root model is often consistent with the observed low-frequency variability.

Second, our theoretical results on the similarity of the low-frequency implications of alternative models imply that it is essentially impossible to discriminate between these models based on low-frequency information using sample sizes typically encountered in emprical work. When using any one of these one parameter low-frequency models for empirical work, one thus must either rely on extraneous information to argue for the correct model choice, or one must take these models seriously over a much wider frequency band. Neither of these two options is particularly attractive for many applications, which raises the question whether econometric techniques can be developed that remain valid for a wide range of low-frequency models.

Third, maybe the most important empirical conclusion is that for many series there seems to be too much low-frequency variability in the second moment to provide good fits for any of the models. From an economic perspective, this underlines the importance of understanding the sources and implications of such low-frequency volatility changes. From a statistical perspective, this finding motivates further research into methods that allow for substantial time variation in second moments.

A Appendix

A.1 Proofs of Theorems 1 and 2

Proof of Theorem 1:

Define $S_t = \sum_{s=1}^{t} u_s$ and $S_t^i = \sum_{s=1}^{t} u_s^i$, $i = \mu, \tau$. With $T^{-\alpha} S_{[\cdot T]} \Rightarrow \sigma G(\cdot)$, we find by least squares algebra and the CMT

$$\begin{aligned} T^{-\alpha} S_{[sT]}^{\mu} &= T^{-\alpha} S_{[sT]} - sT^{-\alpha} S_T + R_T^{\mu}(s) \\ T^{-\alpha} S_{[sT]}^{\tau} &= T^{-\alpha} S_{[sT]}^{\mu} - 6s(1-s) \int T^{-\alpha} S_{[\lambda T]}^{\mu} d\lambda + R_T^{\tau}(s) \end{aligned}$$

where $\sup_{s \in [0,1]} |R_T^i(s)| \xrightarrow{p} 0$ for $i = \mu, \tau$. Thus, by the CMT

$$T^{-\alpha} \sigma^{-1} S_{[sT]}^{\mu} \Rightarrow G(s) - sG(1) \equiv G^{\mu}(s) \tag{9}$$

$$T^{-\alpha} \sigma^{-1} S_{[sT]}^{\tau} \Rightarrow G^{\mu}(s) - 6s(1-s) \int G^{\mu}(\lambda) d\lambda \equiv G^{\tau}(s) \tag{10}$$

and

$$E[G^{\mu}(r) G^{\mu}(s)] = E[(G(r) - rG(1))(G(s) - sG(1))] = k^{\mu}(r,s)$$

and similarly, $k^{\tau}(r,s) = E[G^{\tau}(r) G^{\tau}(s)]$. By summation by parts

$$\begin{aligned} \sum_{t=1}^{T} \Psi_l(t/T) u_t^i &= S_T^i \Psi_l(1) - \sum_{t=1}^{T} S_{t-1}^i (\Psi_l(t/T) - \Psi_l((t-1)/T)) \\ &= -\int S_{[sT]}^i \psi_l(s) ds + \int S_{[sT]}^i (\psi_l(s) - \frac{\Psi_l([sT]/T + T^{-1}) - \Psi_l([sT]/T)}{T^{-1}}) ds, \end{aligned}$$

since $S_T^i = 0$ for $i = \mu, \tau$. Application of the mean-value theorem yields

$$\sup_{s \in [0,1]} |\psi_l(s) - \frac{\Psi_l([sT]/T + T^{-1}) - \Psi_l([sT]/T)}{T^{-1}}| \leq \sup_{s \in [0,1]} \sup_{s' \in [0,1], |s'-s| \leq 2T^{-1}} |\psi_l(s) - \psi_l(s')| \to 0$$

and the uniform convergence follows from continuity (and hence uniform continuity) of $\psi_l(\cdot)$ on $[0,1]$. Thus

$$\begin{aligned} & T^{-\alpha} |\int S_{[sT]}^i (\psi_l(s) - \frac{\Psi_l([sT]/T + T^{-1}) - \Psi_l([sT]/T)}{T^{-1}}) ds| \\ \leq & \sup_{s \in [0,1]} |\psi_l(s) - \frac{\Psi_l([sT]/T + T^{-1}) - \Psi_l([sT]/T)}{T^{-1}}| \int |T^{-\alpha} S_{[sT]}^i| ds \xrightarrow{p} 0 \end{aligned}$$

since $\int |T^{-\alpha} S^i_{[sT]}| ds \Rightarrow \sigma \int |G^i(s)| ds$ by the CMT. Using this result row by row and the convergences (9) and (10), we obtain by the CMT

$$\begin{aligned} X_T &= T^{-\alpha} \sum_{t=1}^{T} \Psi(t/T) u_t^i \\ &= -\int T^{-\alpha} S^i_{[sT]} \psi(s) ds + o_p(1) \Rightarrow -\sigma \int G^i(s) \psi(s) ds \end{aligned}$$

where $\psi(\cdot) = (\psi_1(\cdot), \cdots, \psi_q(\cdot))'$, and the result follows.

The proof of Theorem 2 relies in part on the following Lemma.

Lemma 1 *Suppose $\{\phi_l\}_{l=0}^{\infty}$ is an orthonormal basis of $L^2[0,1]$, and $\varsigma_l \in L^2[0,1]$, $l = 0, 1, 2, \ldots$ are orthonormal. If $\int_0^1 \phi_l(s) \varsigma_l(s) ds > 1/\sqrt{2}$ for all $l \geq 0$, then $\{\varsigma_l\}_{l=0}^{\infty}$ is an orthonormal basis of $L^2[0,1]$, too.*

Proof. For $f_1, f_2 \in L^2[0,1]$, write $\langle f_1, f_2 \rangle$ for $\int_0^1 f_1(s) f_2(s) ds$.

Suppose otherwise. Then there exists a function $f \in L^2[0,1]$ with $\langle f, f \rangle = 1$ such that for all $l \geq 0$, $\langle f, \varsigma_l \rangle = 0$. Since $\{\phi_l\}_{l=0}^{\infty}$ is a basis, there exists a real sequences c_l with $\sum_{l=0}^{\infty} c_l^2 = 1$ so that $\lim_{n \to \infty} \int (f(s) - \sum_{l=0}^{n} c_l \phi_l(s))^2 ds = 0$. Let $\tilde{f} = \sum_{l=0}^{\infty} c_l \varsigma_l$. Since $\{\varsigma_l\}_{l=0}^{\infty}$ is orthonormal, $\langle \tilde{f}, \tilde{f} \rangle = 1$. Denote $\varsigma_l^* = \varsigma_l - \langle \phi_l, \varsigma_l \rangle \phi_l$ and note that $\langle \varsigma_l^*, \varsigma_l^* \rangle = 1 - \langle \phi_l, \varsigma_l \rangle^2 < 1/2$. We have

$$\langle f, \tilde{f} \rangle = \langle f, \sum_{l=0}^{\infty} c_l \phi_l \rangle + \langle f, \sum_{l=0}^{\infty} c_l \varsigma_l^* \rangle.$$

By the Cauchy-Schwarz inequality, $\langle f, \sum_{l=0}^{\infty} c_l \varsigma_l^* \rangle^2 \leq \sum_{l=0}^{\infty} c_l^2 \langle \varsigma_l^*, \varsigma_l^* \rangle \leq 1/2$. But $\langle f, \sum_{l=0}^{\infty} c_l \phi_l \rangle = \sum_{l=0}^{\infty} c_l^2 \langle \phi_l, \varsigma_l \rangle > 1/\sqrt{2}$, so that $\langle f, \tilde{f} \rangle > 0$, which contradicts $\langle f, \varsigma_l \rangle = 0$ for all $l \geq 0$. ■

Proof of Theorem 2:

Standard calculations show that

$$\begin{aligned} k_W^{\mu}(r,s) &= \min(s,r) + \tfrac{1}{3} - (r+s) + \tfrac{1}{2}(r^2 + s^2) \\ k_W^{\tau}(r,s) &= \min(s,r) + \tfrac{2}{15} + \tfrac{6}{5} rs - \tfrac{11}{10}(r+s) + 2(r^2+s^2) - (r^3+s^3) \\ &\quad -3(r^2 s + r s^2) + 2(r^3 s + r s^3). \end{aligned}$$

Noting that for any real $\lambda \neq 0$, $s > 0$ and ϕ

$$\begin{aligned}
\int_0^s \sin(\lambda u + \phi) u du &= (\sin(\lambda s + \phi) - \lambda s \cos(\lambda s + \phi) - \sin(\phi))/\lambda^2 \\
\int_0^s \sin(\lambda u + \phi) u^2 du &= (2s\lambda \sin(\lambda s + \phi) + (2 - \lambda^2 s^2)\cos(\lambda s + \phi) - 2\cos(\phi))/\lambda^3 \\
\int_0^s \sin(\lambda u + \phi) u^3 du &= (3(\lambda^2 s^2 - 2)\sin(\lambda s + \phi) + \lambda s(6 - \lambda^2 s^2)\cos(\lambda s + \phi) + 6\sin(\phi))/\lambda^4
\end{aligned}$$

it is straightforward, but highly tedious, to confirm that $\int_0^1 E[W^i(s)W^i(r)]\varphi_l^i(s)ds = \lambda_l^i \varphi_l^i(r)$ for $l = 0, 1, \ldots$ when $i = \mu$ and for $l = -1, 0, 1, 2, \ldots$ when $i = \tau$.

To show that $\{\varphi_l^\mu\}_{l=0}^\infty$ and $\{\varphi_l^\tau\}_{l=-1}^\infty$ are the complete set of eigenfunctions, it suffices to show that they form a basis of $L^2[0,1]$. This is a well known result for $\{\varphi_l^\mu\}_{l=0}^\infty$, because the cosine expansion is the real part of the usual Fourier expansion. For $\{\varphi_l^\tau\}_{l=-1}^\infty$, note that $\varphi_{-1}^\tau = \varphi_0^\mu$ and $\varphi_l^\tau = \varphi_{l+1}^\mu$ for odd $l \geq 1$. Furthermore, $\int \varphi_0^\tau(s)\varphi_1^\mu(s)ds = 4\sqrt{6}/\pi^2 > 1/\sqrt{2}$. It is not hard to see that the jth positive root ω_j of $\cos(\omega/2) = 2\sin(\omega/2)/\omega$ satisfies $(2j+1)\pi - \pi/6 < \omega_j < \pi(2j+1)$. Therefore, $1 < \sqrt{\omega_j/(\omega_j - \sin(\omega_j))} < \sqrt{17\pi/(17\pi - 3)} < 1.03$ for all $j \geq 1$, and by an exact second order Taylor expansion of $\sin(\omega(s - 1/2))$ around $\omega = \pi(2j+1)$

$$\begin{aligned}
&\sup_{s\in[0,1]} \Bigg| \varphi_l^\tau(s) - \sqrt{\frac{\omega_{l/2}}{\omega_{l/2} - \sin(\omega_{l/2})}}\varphi_{l+1}^\mu(s) \\
&\qquad - \sqrt{\frac{\omega_{l/2}}{\omega_{l/2} - \sin(\omega_{l/2})}}(-1)^{(l+2)/2}\sqrt{2}\sin(\pi(l+1)s)(s - \tfrac{1}{2})(\omega_{l/2} - l\pi - \pi)\Bigg| \\
&\leq 1.03\frac{\sqrt{2}\pi^2}{144} < 0.1
\end{aligned}$$

for all even $l \geq 2$. Since $\int \sin(\pi l s)\cos(\pi l s)(1 - 2s)ds = (2l\pi)^{-1}$ for $l \geq 1$, by the Cauchy-Schwarz inequality, we thus find $\int \varphi_l^\tau(s)\varphi_{l+1}^\mu(s)ds > 0.9 - 1.03/24 > 0.85$ for all even $l \geq 2$. Completeness of $\{\varphi_l^\tau\}_{l=-1}^\infty$ now follows from Lemma 1.

A.2 Continuity of fractional process at $d = 1/2$:

By the definition of $k_{FR(d)}^\mu(r,s)$ and $k_{I\text{-}FR(d)}^\mu(r,s)$, we find for $s \leq r$

$$\begin{aligned}
k_{FR(d)}^\mu(r,s) = \tfrac{1}{2}[&s^{1+2d} + r^{1+2d} - (r-s)^{1+2d} + 2rs \\
&- s(1 - (1-r)^{1+2d} + r^{1+2d}) - r(1 - (1-s)^{1+2d} + s^{1+2d})]
\end{aligned}$$

and

$$k^{\mu}_{I\text{-}FR(d)}(r,s) = \frac{1}{4d(1+2d)}[-r^{1+2d}(1-s) - s(s^{2d} + (r-s)^{2d} + (1-r)^{2d} - 1)$$
$$+ r(s^{1+2d} + 1 - (1-s)^{2d} + (r-s)^{2d}) + sr((1-s)^{2d} + (1-r)^{2d} - 2))]$$

so that

$$\lim_{d\uparrow 1/2} k^{\mu}_{FR(d)}(r,s) = \lim_{d\downarrow 1/2} k^{\mu}_{I\text{-}FR(d)}(r,s) = 0.$$

Now for $0 < s < r$, using that for any real $a > 0$, $\lim_{x\downarrow 0}(a^x - 1)/x = \ln a$, we find

$$\lim_{d\uparrow 1/2} \frac{k^{\mu}_{FR(d)}(r,s)}{1/2-d} = -(1-r)^2 s \ln(1-r) - r^2(1-s)\ln r - r(1-s)^2 \ln(1-s)$$
$$+(r-s)^2 \ln(r-s) + (r-1)s^2 \ln(s) \tag{11}$$

and

$$\lim_{d\uparrow 1/2} \frac{k^{\mu}_{FR(d)}(r,r)}{1/2-d} = 2(1-r)r(-(1-r)\ln(1-r) - r\ln r). \tag{12}$$

Performing the same computation for $k^{\mu}_{I\text{-}FR(d)}(r,s)$ yields the result.

A.3 Some Simplifications of the Methodology

The empirical analysis in Section 4 is based on the weighting functions $\Psi_l(s) = \varphi^{\mu}_l(s)$ and $\Psi_l(s) = \varphi^{\tau}_l(s)$ for the mean and trend case, and the tests described in Section 3. While conceptually straightforward, some of the tests are tedious to perform, and here we suggest three simplifications that yield similar results.

First, note that φ^{τ}_l in differ from φ^{μ}_{l+1} only for even l, and not by very much: if ω_j was defined by the roots of $\cos(\omega/2) = 0$, one would obtain $\varphi^{\tau}_l = \varphi^{\mu}_{l+1}$ also for even l. Especially for l large, the additional term $2\sin(\omega/2)/\omega$ in the definition of ω_j only leads to a minor distortion. One might hence avoid the computation of ω_j and set $\Psi^{\tau}_l = \varphi^{\mu}_{l+1}$ for $l = 1, \cdots, q$, without generating large off-diagonal elements in Σ in the I(1) model. Unreported results show that with Ψ^{τ}_l so defined, Σ remains very close to diagonal in most models.

Second, the critical value of the LB statistic (4) depends on Σ_0, and hence must be computed by simulation for each null model. An alternative is to base the test instead on $v^* = Qv/\sqrt{v'Q'Qv}$ for some matrix Q satisfying $Q\Sigma_0 Q' = I_q$; in this case the null hypothesis about the parameter Σ^* of the density of v^* becomes $H_0 : \Sigma^* = I_q$, and $v^* \sim Z/\sqrt{Z'Z}$ under the null hypothesis, where $Z \sim \mathcal{N}(0, I_q)$. We suggest choosing Q lower triangular, so that Q^{-1} is the Choleski decomposition of Σ_0. Because $\Sigma(\theta)$ is approximately diagonal for most empirically relevant models, there are only small differences between tests based on v and tests based on v^*.

For $\Sigma^* = I_q$, $b_l = v_l^{*2}$ and B becomes a diagonal matrix with elements $B_{ll} = 2v_l^{*2}$, resulting in a test statistic for martingale variation in δ which simplifies to

$$\text{LBIM}^* = (q/2+1)\sum_{l=1}^{q}(\sum_{j=1}^{l} v_j^{*2} - \overline{v^{*2}})^2 - \frac{1}{3q}\sum_{l=1}^{q}[6l^2 - 6l(1+q) + (1+q)(1+2q)]v_l^{*2}.$$

where $\overline{v^{*2}} = q^{-1}\sum_{l=1}^{q} v_l^{*2}$. The first term, which dominates the statistic for large q, is the usual Nyblom (1989) locally best test statistic for a martingale variation in the mean of v_l^{*2}.

Third, recall from the discussion in Section 3.2 that severe heteroskedasticity in u_t generates autocorrelated X. This motivates the simple test statistic

$$H^* = \frac{1}{q}\sum_{l=1}^{q}\frac{|\hat{\rho}_l^*|}{l}$$

where $\hat{\rho}_l^* = q^{-1}\sum_{l=1}^{q} v_l^*$. Results not reported here show that for $q = 15$ in mean case, a 10% level test based on H^* has in most models about 6 percentage points less power than the 10% level optimal test (7) against the alternative the test (7) is optimal against. The critical value of H^* again only depends on q.

A.4 Data Appendix

Table A1 lists the series used in section 4, the sample period, data frequency transformation, and data source and notes.

References

Akdi, Y., and D. Dickey (1998): "Periodograms of Unit Root Time Series: Distributions and Tests," *Communications in Statistics: Theory and Methods*, 27, 69–87.

Andersen, T., and T. Bollerslev (1997): "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of Finance*, 52, 975–1005.

Andersen, T., T. Bollerslev, P. Christoffersen, and F. Diebold (2006): *Volatility: Practical Methods for Financial Applications*. Princeton University Press, Princeton.

Balke, N., and R. Gordon (1989): "The Estimation of Prewar Gross National Product: Methodology and New Evidence," *Journal of Political Economy*, 94, 38–92.

Beveridge, S., and C. Nelson (1981): "A New Approach to Decomposition of Economics Time Series Into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics*, 7, 151–174.

Bierens, H. (1997): "Nonparametric Cointegration Analysis," *Journal of Econometrics*, 77, 379–404.

Bollerslev, T., R. Engle, and D. Nelson (1994): "ARCH Models," in *Handbook of Econometrics Vol. IV*, ed. by R. Engle, and D. McFadden. Elsevier Science, Amsterdam.

Bollerslev, T., and H. Mikkelsen (1999): "Long-Term Equity Anticipation Securities and Stock Market Volatility Dynamics," *Journal of Econometrics*, 92, 75–99.

Campbell, J., and M. Yogo (2006): "Efficient Tests of Stock Return Predictability," *forthcoming in Journal of Financial Economics.*

Chan, N., and N. Terrin (1995): "Inference for Unstable Long-Memory Processes with Applications to Fractional Unit Root Autoregressions," *Annals of Statistics*, 23, 1662–1683.

Christiano, L., M. Eichenbaum, and R. Vigfusson (2003): "What Happens After a Technology Shock," *NBER Working Paper 9819.*

Davidson, J. (2002): "Establishing Conditions for the Functional Central Limit Theorem in Nonlinear and Semiparametric Time Series Processes," *Journal of Econometrics*, 106, 243–269.

Davidson, J., and N. Hashimadze (2006): "Type I and Type II Fractional Brownian Motions: A Reconsideration," *mimeo, University of Exeter.*

Davidson, J., and P. Sibbertsen (2005): "Generating Schemes for Long Memory Processes: Regimes, Aggregation and Linearity," *Journal of Econometrics*, 128, 253–282.

Diebold, F., and A. Inoue (2001): "Long Memory and Regime Switching," *Journal of Econometrics*, 105, 131–159.

Ding, Z., C. Granger, and R. Engle (1993): "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1, 83–116.

Elliott, G. (1999): "Efficient Tests for a Unit Root When the Initial Observation is Drawn From its Unconditional Distribution," *International Economic Review*, 40, 767–783.

Elliott, G., T. Rothenberg, and J. Stock (1996): "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

Fama, E. (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383–417.

Francis, N., and V. Ramey (2006a): "Is the Technology-Driven Business Cycle Hypothesis Dead?," *forthcoming in Journal of Monetary Economics.*

——— (2006b): "Measures of Per Capita Hours and their Implications for the Technology-Hours Debate," *mimeo, U.C. San Diego.*

Gali, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 89, 249–271.

Geweke, J., and S. Porter-Hudak (1983): "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

Harvey, A. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Kariya, T. (1980): "Locally Robust Test for Serial Correlation in Least Squares Regression," *Annals of Statistics*, 8, 1065–1070.

Kim, C.-J., and C. Nelson (1999): "Has the Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *The Review of Economics and Statistics*, 81, 608–616.

King, M. (1980): "Robust Tests for Spherical Symmetry and their Application to Least Squares Regression," *The Annals of Statistics*, 8, 1265–1271.

King, R., C. Plosser, J. Stock, and M. Watson (1991): "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81, 819–840.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin (1992): "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics*, 54, 159–178.

Lettau, M., and S. Ludvigson (2004): "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption," *American Economic Review*, 94, 276–299.

Lothian, J., and M. Taylor (1996): "Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries," *Journal of Political Economy*, 104, 488–509.

Mandelbrot, B., and J. V. Ness (1968): "Fractional Brownian Motions, Fractional Noise and Applications," *SIAM Review*, 10, 422–437.

Marinucci, D., and P. Robinson (1999): "Alternative Forms of Fractional Brownian Motion," *Journal of Statistical Planning and Inference*, 80, 111–122.

McConnell, M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's," *American Economic Review*, 90, 1464–1476.

McLeish, D. (1974): "Dependent Central Limit Theorems and Invariance Principles," *The Annals of Probability*, 2, 620–628.

Meese, R., and K. Rogoff (1983): "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics*, 14, 3–24.

Müller, U. (2004): "A Theory of Robust Long-Run Variance Estimation," *mimeo, Princeton University.*

Nelson, C., and C. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series — Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139–162.

Nyblom, J. (1989): "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223–230.

Parke, W. (1999): "What is Fractional Integration?," *Review of Economics and Statistics*, 81, 632–638.

Pesavento, E., and B. Rossi (2005): "Do Technology Shocks Drive Hours Up or Down? A Little Evidence from an Agnostic Procedure," *Macroeconomic Dynamics*, 9, 478–488.

Phillips, P. (1998): "New Tools for Understanding Spurious Regression," *Econometrica*, 66, 1299–1325.

——— (2006): "Optimal Estimation of Cointegrated Systems with Irrelevant Instruments," *Cowles Foundation Discussion Paper 1547.*

Phillips, P., and V. Solo (1992): "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971–1001.

Pollard, D. (2002): *A User's Guide to Measure Theoretic Probability*. Cambridge University Press, Cambridge, UK.

Robinson, P. (2003): "Long-Memory Time Series," in *Time Series with Long Memory*, ed. by P. Robinson, pp. 4–32. Oxford University Press, Oxford.

Said, S., and D. Dickey (1984): "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika*, 71, 2599–607.

Stock, J. (1994): "Unit Roots, Structural Breaks and Trends," in *Handbook of Econometrics*, ed. by R. Engle, and D. McFadden, vol. 4, pp. 2740–2841. North Holland, New York.

Stock, J., and M. Watson (2005): "Has Inflation Become Harder to Forecast?," *mimeo, Princeton University*.

Taqqu, M. (1975): "Convergence of Integrated Processes of Arbitrary Hermite Rank," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 50, 53–83.

Velasco, C. (1999): "Non-Stationary Log-Periodogram Regression," *Journal of Econometrics*, 91, 325–371.

Watson, M. (1999): "Explaining the Increased Variability in Long-Term Interest Rates," *Federal Reserve Bank of Richmond–Economic Quarterly*, 85, 71–96.

Wooldridge, J., and H. White (1988): "Some Invariance Principles and Central Limit Theorems for Dependent Heterogeneous Processes," *Econometric Theory*, 4, 210–230.

Table 1: Asymptotic Properties of Partial Sums of Popular Time Series Models

	Process	Parameter	Partial sum convergence	covariance kernel $k(r,s)$, $s \leq r$
1a	FR	$-\frac{1}{2} < d < \frac{1}{2}$	$T^{-1/2-d}\sigma^{-1}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow W^d(\cdot)$	$\frac{1}{2}(r^{2d+1} + s^{2d+1} - (r-s)^{2d+1})$
1b	FR	$\frac{1}{2} < d < \frac{3}{2}$	$T^{-1/2-d}\sigma^{-1}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \int_0^{\cdot} W^{d-1}(\lambda)d\lambda$	$\frac{(r-s)^{2d+1}+(1+2d)(rs^{2d}+r^{2d}s)-r^{2d+1}-s^{2d+1}}{4d(1+2d)}$
2a	OU	$c > 0$	$T^{-3/2}\sigma^{-1}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \int_0^{\cdot} J^c(\lambda)d\lambda$	$\frac{2cs-1+e^{-cs}+e^{-cr}-e^{-c(r-s)}}{2c^3}$
2b	OU	$c \leq 0$	$T^{-3/2}\sigma^{-1}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \int_0^{\cdot}\int_0^{\nu} e^{-c(\nu-\lambda)}dW(\lambda)d\nu$	$\begin{cases} \frac{2cs-2+2e^{-cr}+2e^{-cs}-e^{-cr}(e^{-cs}+e^{cs})}{2c^3} & \text{for } c<0 \\ \frac{1}{6}(3rs^2 - s^3) \text{ for } c=0 \end{cases}$
3	I-OU	$c > 0$	$T^{-5/2}\sigma^{-1}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \int_0^{\cdot}\int_0^{\nu} J^c(\lambda)d\lambda d\nu$	$\frac{3-sc(3+c^2s^2)+3rc(1-cs+c^2s^2)-3e^{-cs}(1+cr)-3e^{-cr}(1+cs-e^{cs})}{6c^5}$
4	LL	$g \geq 0$	$T^{-1/2}\sigma^{-1}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow W_1(\cdot) + g\int_0^{\cdot} W_2(\lambda)d\lambda$	$s + \frac{1}{6}g^2(3rs^2 - s^3)$
5	I-LL	$g \geq 0$	$T^{-3/2}\sigma^{-1}\sum_{t=1}^{[\cdot T]} u_t \Rightarrow \int_0^{\cdot} W_1(\lambda)d\lambda + g\int_0^{\cdot}\int_0^{\nu} W_2(\lambda)d\lambda d\nu$	$\frac{1}{6}(3rs^2 - s^3) + \frac{1}{120}g^2(10r^2s^3 - 5rs^4 + s^5)$

Notes: W, W_1 and W_2 are independent standard Wiener processes, W^d is "type I" fractional Brownian Motion defined as $W^d(s) = A(d)\int_{-\infty}^0 \left[(s-\lambda)^d - (-\lambda)^d\right] dW(\lambda) + A(d)\int_0^s (s-\lambda)^d dW(\lambda)$ where $A(d) = \left(\frac{1}{2d+1} + \int_0^\infty \left[(1+\lambda)^d - \lambda^d\right]^2 d\lambda\right)^{-1/2}$ and J^c is the stationary Ornstein-Uhlenbeck process $J^c(s) = Ze^{-sc}/\sqrt{2c} + \int_0^s e^{-c(s-\lambda)}dW(\lambda)$ with $Z \sim \mathcal{N}(0,1)$ independent of W.

Table 2
Average Absolute Correlations for $\Sigma(\theta)$

Fractional Model	***d* = −0.25**	***d* = 0.00**	***d* = 0.25**	***d* = 0.75**	***d* = 1.00**	***d* = 1.25**
Demeaned	0.03	0.00	0.01	0.01	0.00	0.03
Detrended	0.03	0.00	0.01	0.01	0.00	0.02
Local-to-Unity Model	***c* = 30**	***c* = 20**	***c* = 15**	***c* = 10**	***c* = 5**	***c* = 0**
Demeaned	0.02	0.02	0.02	0.02	0.02	0.00
Detrended	0.02	0.02	0.02	0.02	0.01	0.00
Local Level Model	***g* = 0**	***g* = 2**	***g* = 5**	***g* = 10**	***g* = 20**	***g* = 30**
Demeaned	0.00	0.00	0.00	0.00	0.00	0.00
Detrended	0.00	0.00	0.00	0.00	0.00	0.00

Notes: Entries in the table are the average values of the absolute values of the correlations associated with $\Sigma(\theta)$ with $q = 15$ for the demeaned model and $q = 14$ for the detrended model.

Table 3: *P*-values for $I(0)$ and $I(1)$ Models

Series	I(0)			I(1)		
	LBIM	H	LFST	LBIM	H	LFUR
Real GDP (PWQ)	0.02	0.12	0.01	0.50	0.89	0.34
Real GNP (Long Annual)	0.00	0.00	0.01	0.30	0.00	0.01
Inflation (PWQ)	0.00	0.03	0.02	0.92	0.28	0.22
Inflation (Long Annual)	0.05	0.09	0.01	0.00	0.00	0.00
Productivity	0.00	0.19	0.00	0.15	0.71	0.94
Hours	0.01	0.61	0.00	0.50	0.53	0.45
10YrTBond	0.00	0.22	0.00	0.92	0.10	0.49
1YrTBond	0.00	0.17	0.01	0.43	0.14	0.24
3mthTbill	0.00	0.21	0.01	0.43	0.07	0.25
Bond Rate	0.00	0.00	0.00	0.97	0.00	0.32
Real Tbill Rate	0.19	0.11	0.26	0.19	0.04	0.06
Real Bond Rate	0.24	0.01	0.16	0.00	0.00	0.00
Dollar/Pound Real Ex. Rate	0.00	0.09	0.00	0.00	0.22	0.00
Unit Labor Cost	0.30	0.45	0.01	0.00	0.50	0.03
TBond Spread	0.99	0.10	0.19	0.00	0.06	0.00
real C-GDP	0.00	0.15	0.00	0.39	0.55	0.85
real I-GDP	0.00	0.03	0.00	0.65	0.03	0.65
Earnings/Price (SP500)	0.00	0.00	0.00	0.02	0.01	0.07
Div/Price (CRSP)	0.00	0.01	0.00	0.73	0.31	0.51
Abs.Returns (SP500)	0.01	0.01	0.01	0.00	0.05	0.01

Notes: This table shows *p*-values for the $I(0)$ and $I(1)$ model.

Table 4: DFGLS, Nyblom/KPSS and GPH Results

Series	DFGLS *p*-value	Nyblom/ KPSS *p*-value	GPH Regressions: $\hat{d}$ (SE)			
			Levels		Differences	
			$n^{0.5}$	$n^{0.65}$	$n^{0.5}$	$n^{0.65}$
Real GDP (PWQ)	0.16	<0.01	1.00 (0.17)	0.98 (0.11)	-0.19 (0.17)	-0.09 (0.11)
Real GNP (Long Annual)	0.07	0.10	0.98 (0.19)	0.98 (0.13)	-0.84 (0.19)	-0.46 (0.13)
Inflation (PWQ)	0.14	0.02	0.84 (0.17)	0.91 (0.11)	-0.10 (0.17)	-0.02 (0.11)
Inflation (Long Annual)	0.09	0.01	0.52 (0.19)	0.30 (0.13)	-0.85 (0.19)	-0.93 (0.13)
Productivity	0.84	<0.01	0.95 (0.17)	0.97 (0.11)	0.07 (0.17)	-0.03 (0.11)
Hours	0.50	<0.01	0.75 (0.17)	0.99 (0.11)	-0.11 (0.17)	0.09 (0.11)
10YrTBond	0.21	<0.01	1.05 (0.17)	1.08 (0.11)	0.13 (0.17)	0.10 (0.11)
1YrTBond	0.09	0.01	0.85 (0.17)	0.95 (0.11)	-0.05 (0.17)	-0.03 (0.11)
3mthTbill	0.13	0.01	0.74 (0.17)	1.00 (0.11)	-0.21 (0.17)	0.04 (0.11)
Bond Rate	0.17	<0.01	1.07 (0.20)	1.22 (0.14)	0.08 (0.20)	0.09 (0.14)
Real Tbill Rate	0.01	0.04	0.72 (0.17)	0.79 (0.11)	-0.19 (0.17)	-0.12 (0.11)
Real Bond Rate	<0.01	0.21	0.46 (0.20)	0.35 (0.14)	-0.55 (0.20)	-0.65 (0.14)
Dollar/Pound Real Ex. Rate	0.03	<0.01	0.54 (0.17)	0.43 (0.11)	-0.44 (0.17)	-0.55 (0.11)
Unit Labor Cost	0.00	<0.01	0.57 (0.17)	0.75 (0.11)	-0.55 (0.17)	-0.31 (0.11)
TBond Spread	<0.01	0.01	0.18 (0.17)	0.61 (0.11)	-0.80 (0.17)	-0.41 (0.11)
real C-GDP	0.91	<0.01	0.96 (0.17)	0.96 (0.11)	0.19 (0.17)	-0.10 (0.11)
real I-GDP	0.39	<0.01	0.62 (0.17)	0.87 (0.11)	-0.74 (0.17)	-0.25 (0.11)
Earnings/Price (SP500)	0.02	0.01	0.70 (0.19)	0.62 (0.14)	-0.36 (0.19)	-0.35 (0.14)
Div/Price (CRSP)	0.59	0.01	0.72 (0.23)	0.72 (0.16)	-0.28 (0.23)	-0.43 (0.16)
Abs.Returns (SP500)	<0.01	<0.01	0.46 (0.05)	0.38 (0.03)	-0.52 (0.05)	-0.61 (0.03)

Notes: The entries in the column labeled *DFGLS* are *p*-values for the DFGLS test of Elliott, Rothenberg and Stock (1996). The entries in the column labeled *Nyblom* are *p*-values for the Nyblom (1989) $I(0)$ test (using a HAC covariance matrix as suggested in Kwiatkowski, Phillips, Schmidt and Shin (1992)). Results are computed using a Newey-West HAC estimator with $0.75\times T^{1/3}$ lags. The results in the columns labeled *GPH Regressions* are the estimated values of d and standard errors computed from regressions using the lowest $n^{0.5}$ or $n^{0.65}$ periodogram ordinates. The GPH regressions and standard errors were computed as described in the Robinson (2003): specifically, the *GPH* regressions are of the form $ln(p_i) = \beta_0 + \beta_1 ln(\omega_i)$ + error, where p_i is the i'th periodogram ordinate and ω_i is the corresponding frequency, the estimated value of $\hat{d} = -\hat{\beta}_1/2$, where $\hat{\beta}_1$ is the OLS estimator, and the standard error of $\hat{d}$ is $SE(\hat{d}) = \pi/\sqrt{24m}$, where m is the number of periodogram ordinates used in the regression.

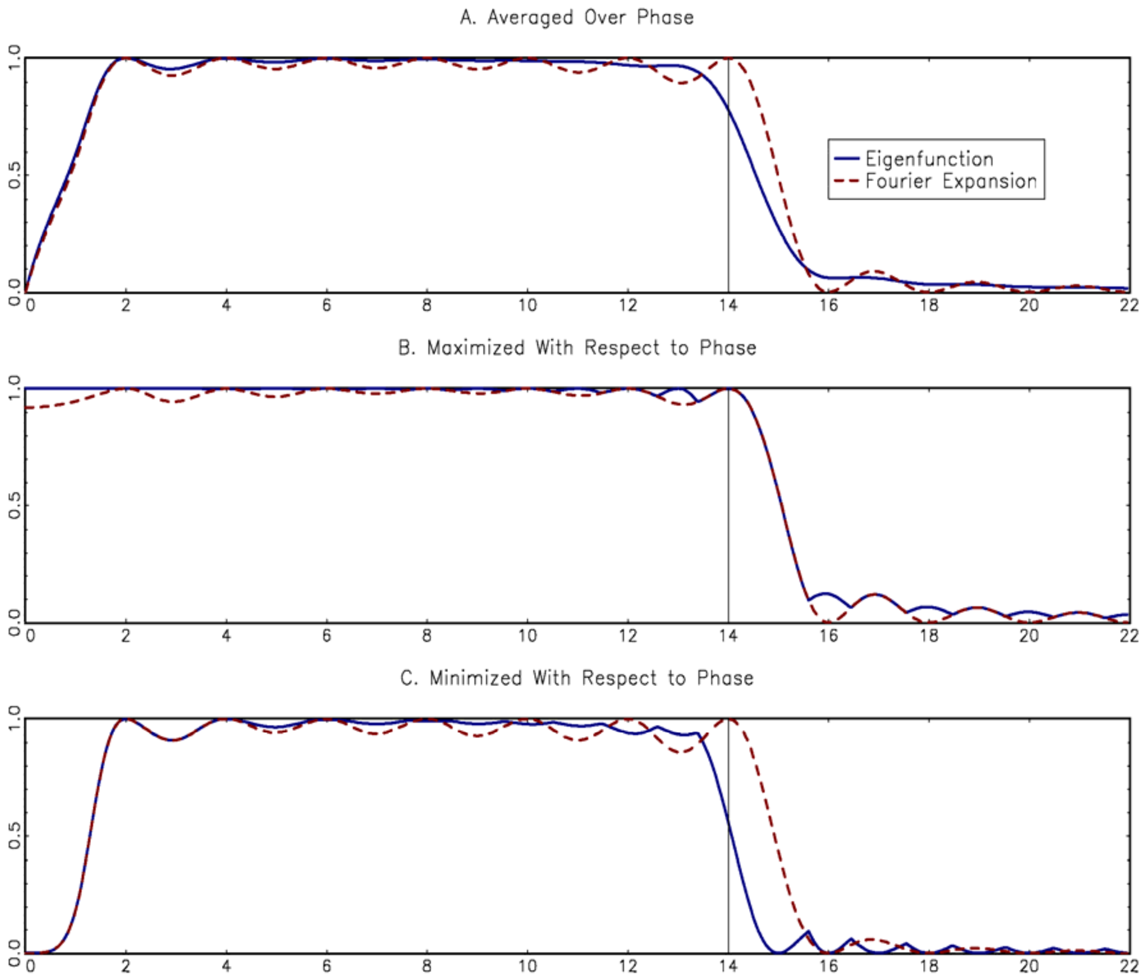
Table A1
Data Description and Sources

Series	Sample Period	F	Tr	Source and Notes
Real GDP	1952:1-2005:3	Q	ln τ	DRI: GDP157
Real GNP (Long Annual)	1869-2004	A	ln τ	1869-1928: Balke and Gordon (1989) 1929-2004: BEA (Series are linked in 1929)
Inflation	1952:1-2005:3	Q	lev μ	DRI: 400×ln(GDP272(*t*)/GDP272(*t*−1))
Inflation (Long Annual)	1870-2004	A	lev μ	GNP Deflator (PGNP): 1869-1928: Balke and Gordon (1989) 1929-2004: BEA (Series are linked in 1929) Inflation Series is 100×ln(PGNP(*t*)/PGNP(*t*−1))
Productivity	1952:1-2005:2	Q	ln τ	DRI: LBOUT (Output per hour, business sector)
Hours	1952:1-2005:2	Q	ln τ	DRI: LBMN(*t*)/P16(*t*) (Employee hours/population)
10YrTBond	1952:1-2005:3	Q	lev μ	DRI: FYGT10
1YrTBond	1952:1-2005:3	Q	lev μ	DRI: FYGT1
3mthTbill	1952:1-2005:2	Q	lev μ	DRI:FYGM3
Bond Rate	1900-2004	A	lev μ	NBER: M13108 (1900-1946) DRI: FYAAAI (1947-2004)
Real Tbill Rate	1952:1-2005:2	Q	lev μ	DRI: FYGM3(*t*)-400×ln(GDP273(*t*+1)/GDP273(*t*))
Real Bond Rate	1900-2004	A	lev μ	R(t) – 100×ln(PGNP(*t*)/PGNP(*t*−1)) R(t) = Bond Rate (described above) PGNP = GNP deflator (described above)
Dollar/Pound Real Ex. Rate	1791-2004	A	ln μ	1791-1990: Lothian and Taylor (1996) 1991-2004: FRB (Nominal Exchange Rate) BLS (US PPI Finished Goods) IFS (UK PPI Manufactured Goods)
Unit Labor Cost	1952:1-2005:2	Q	ln μ	DRI: LBLCP(*t*)/LBGDP(*t*)
TBond Spread	1952:1-2005:3	Q	lev μ	DRI: FYGT10-FYGT1
real C-GDP	1952:1-2005:3	Q	lnrμ	DRI: GDP 158/GDP157
real I-GDP	1952:1-005:3	Q	lnrμ	DRI: GDP 177/ GDP 157
Earnings/Price (SP500)	1880-2002	A	lnrμ	Campbell and Yogo (2006)
Div/Price (CRSP)	1926-2004	A	lnrμ	Campbell and Yogo (2006)
Abs.Returns (SP500)	1/3/1928-1/22/2005	D	lnrμ	SP: SP500(t) is the closing price at date *t*. Absolute returns are |ln[SP500(*t*)/SP500(*t*−1)]|

Notes: The column labeled *F* shows the data frequency (*A*: annual, *Q*: quarterly, and *D*: daily). The column labeled *Tr* (transformation) show the transformation: demeaned levels (lev μ), detrended levels (lev τ), demeaned logarithms (ln μ), detrended logarithms (ln τ), and lnr denotes the logarithm of the indicated ratio. In the column labeled *Source and Notes*, DRI denotes the DRI Economics Database (formerly Citibase) and NBER denotes the NBER historical data base.

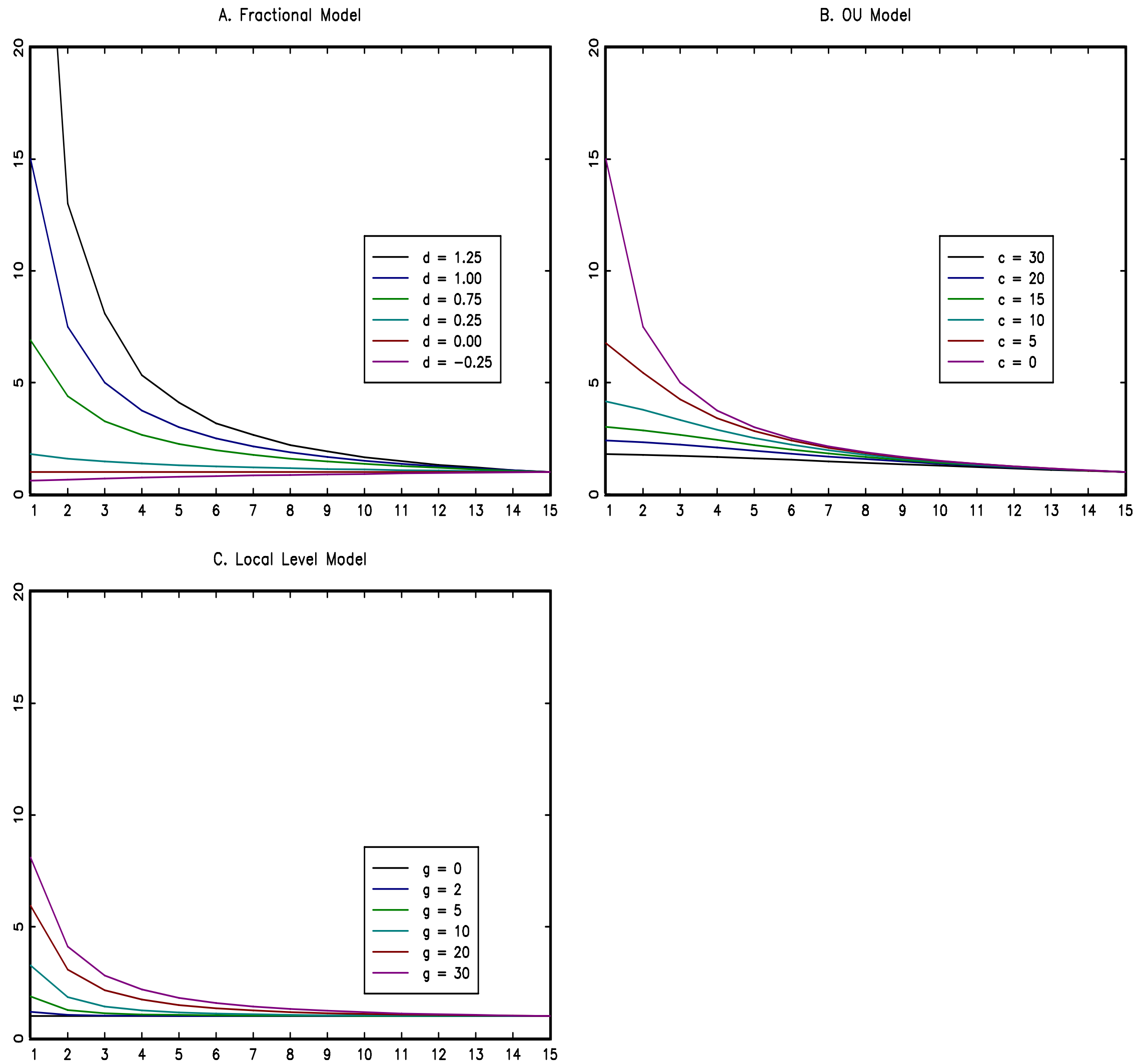
Figure 1

R^2 regression of $\sin(\pi\vartheta s+\phi)$ onto $\Psi_1(s)$ … $\Psi_{14}(s)$



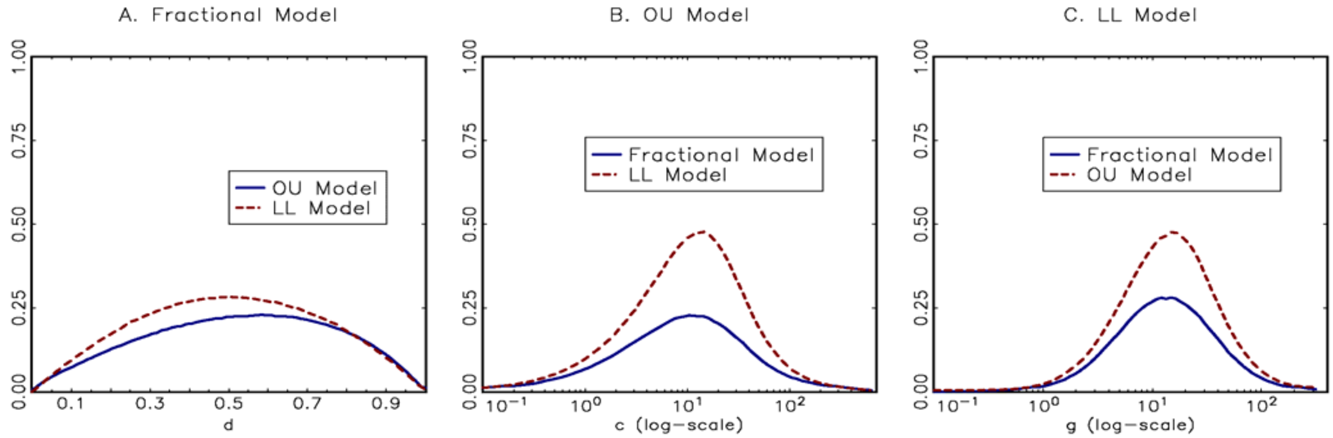
Notes**:** These figures show the R^2 of a continuous time regression of a generic periodic series $\sin(\pi\vartheta s+\phi)$ onto $\Psi_1(s)$, … , $\Psi_{14}(s)$. Panel (a) shows the R^2 value averaged over values of $\phi \in [0, \pi)$, panel (b) shows the R^2 maximized over these values of ϕ for each ϑ, and panel (c) shows the R^2 minimized over these values of ϕ for each ϑ. The solid curve shows results using the eigenfunctions $\varphi_l^{\mu}(s)$ from Theorem 2, and the dashed curve shows results using Fourier expansions.

Figure 2
Standard Deviation of X_l Implied by Different Models



Notes: These figures show the square roots of the diagonal elements of $\Sigma(\theta)$ for different values of the parameter $\theta = (d, c, g)$, where $\Sigma(\theta)$ is computed for the demeaned data. Larger values of d and g, and smaller values of c, yield relatively larger standard deviations of X_1.

Figure 3
Total Variation Distance



Notes: Results are shown for the demeaned case with q = 15.

Figure 4
Detrended/Demeaned Levels, First Differences, and Ψ Transformed Data

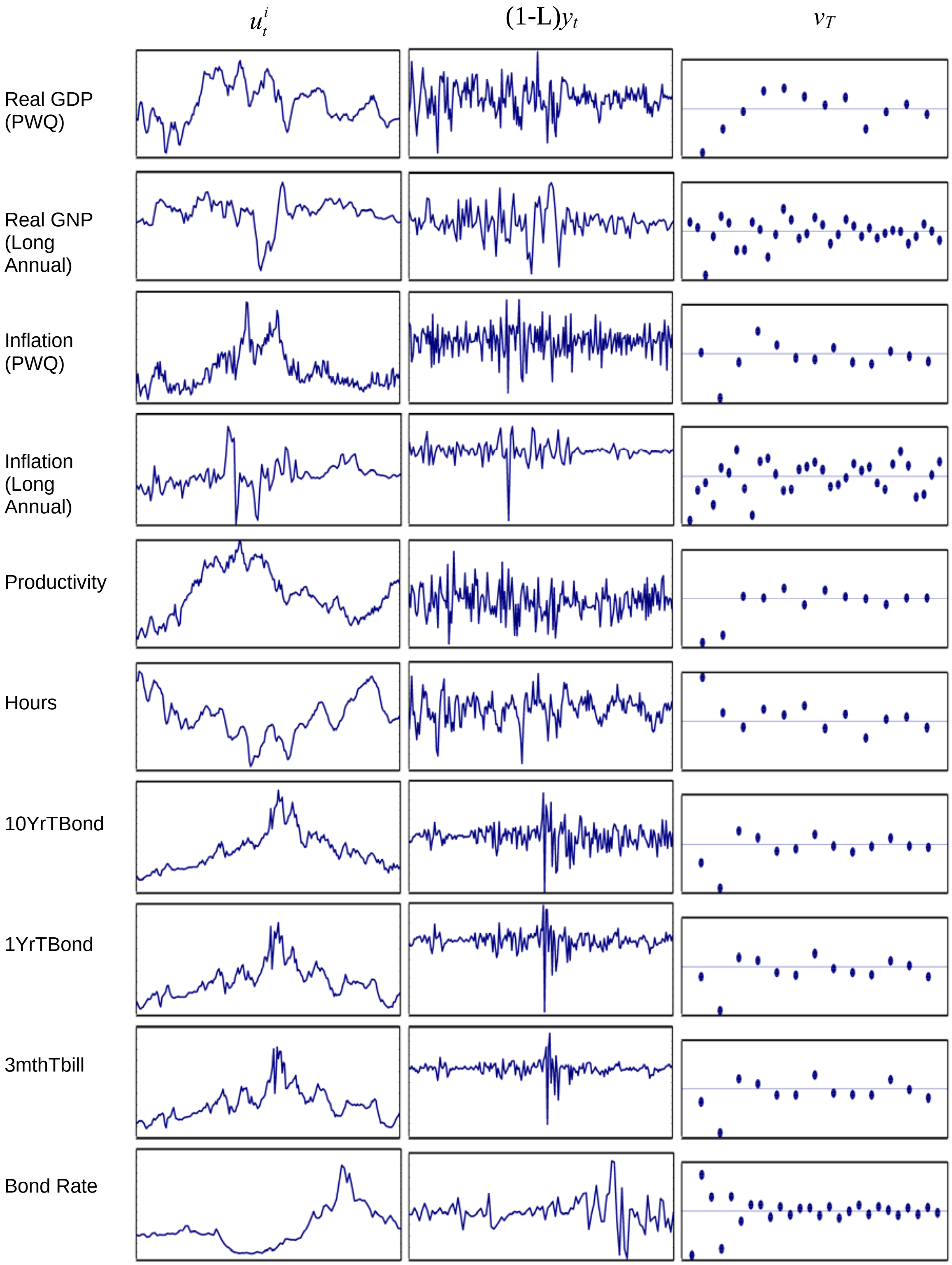


Figure 4 (Continued)

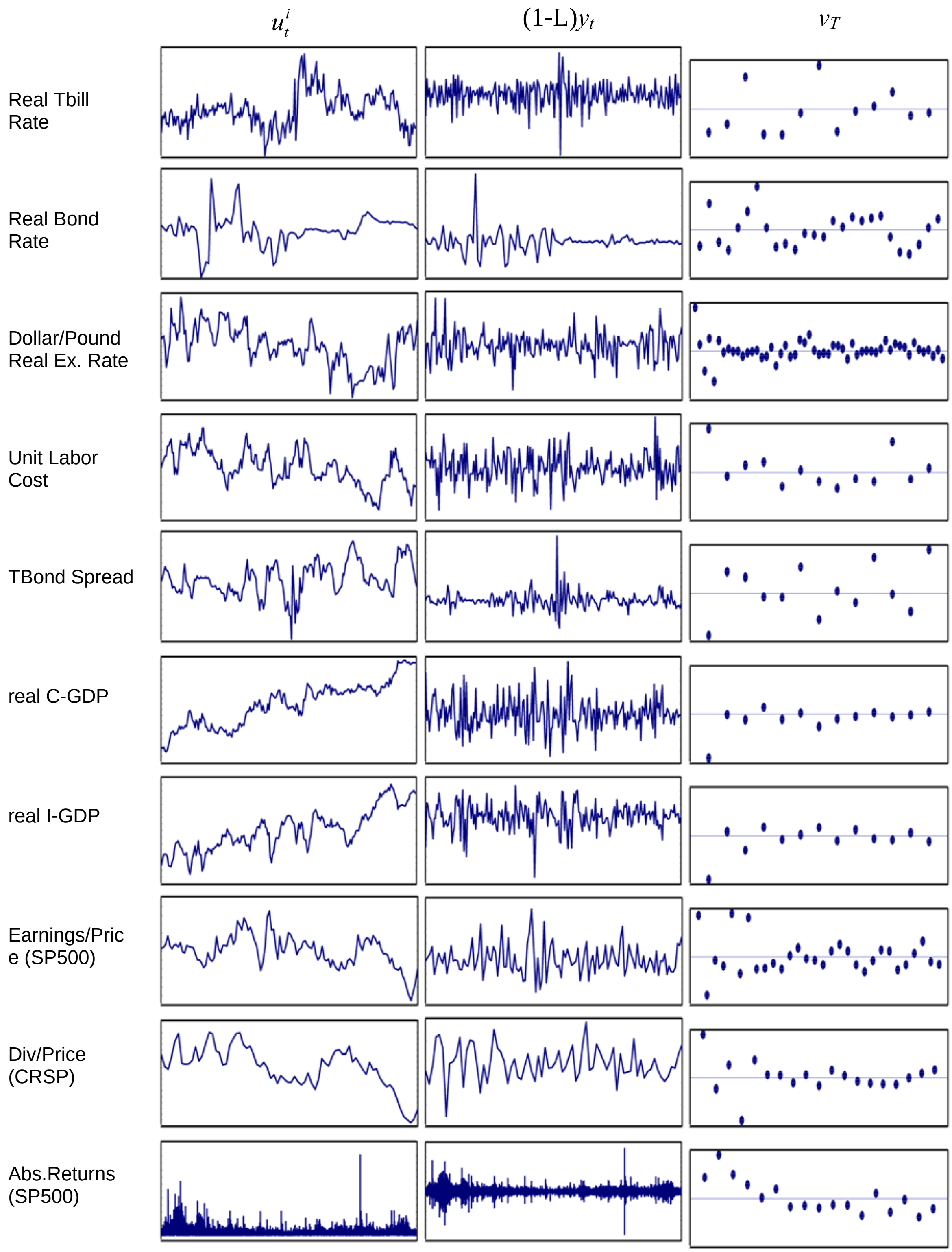


Figure 5
p-values for LBIM (solid line) and H (dotted line) tests

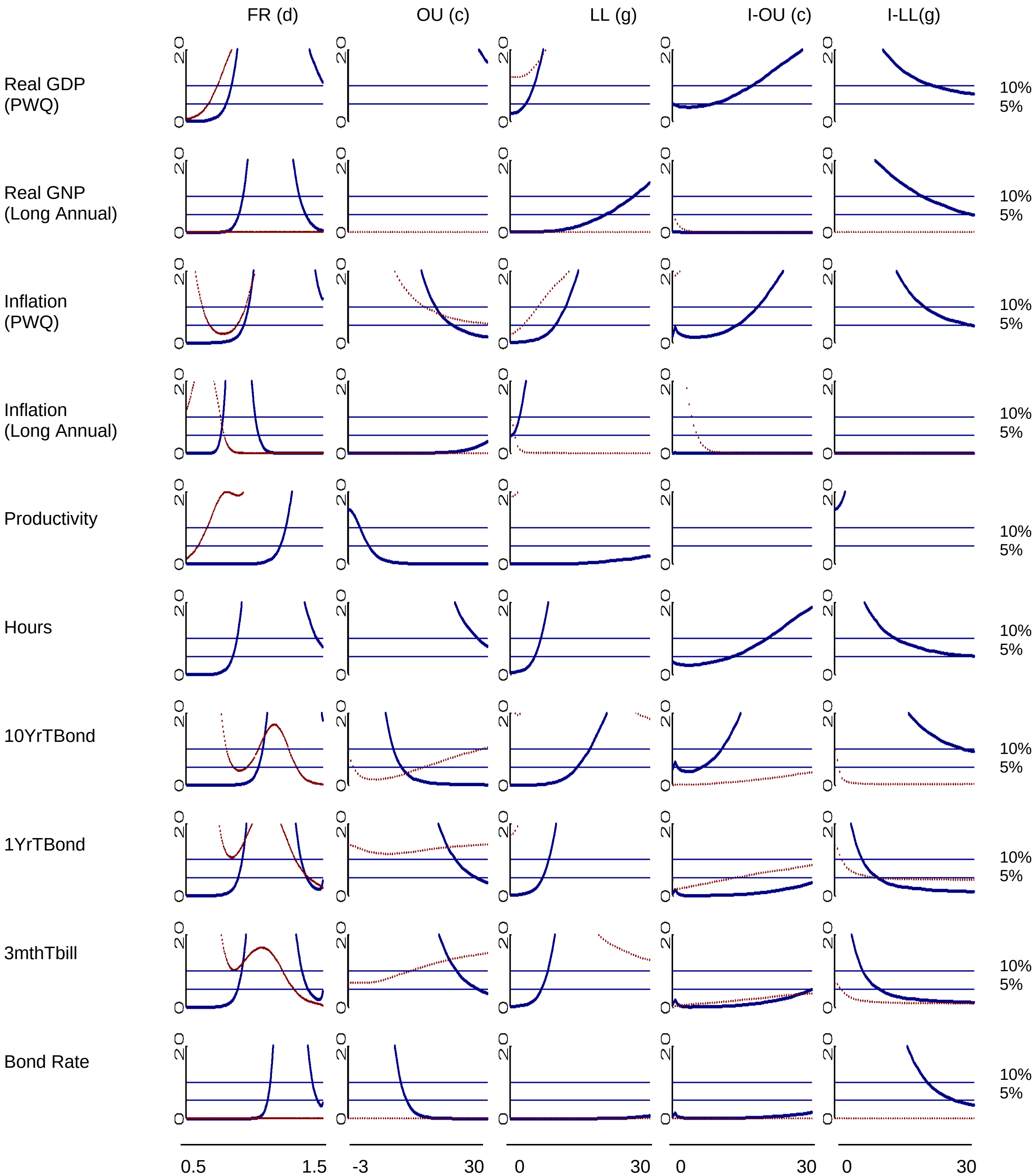


Figure 5 (continued)
***p*-values for LBIM (solid line) and H (dotted line) tests**

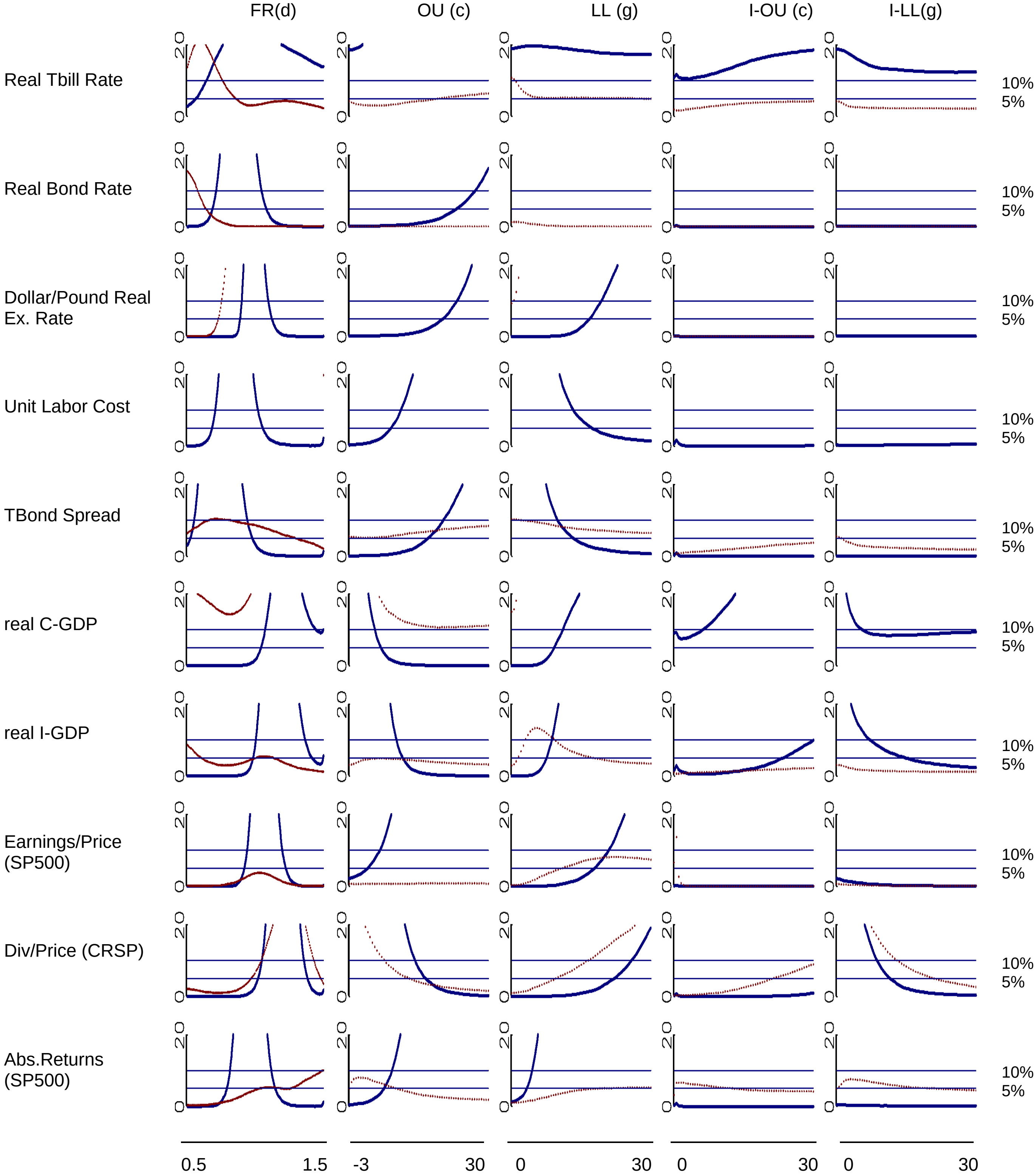


Figure 6
Low-Frequency Log-Likelihood Values

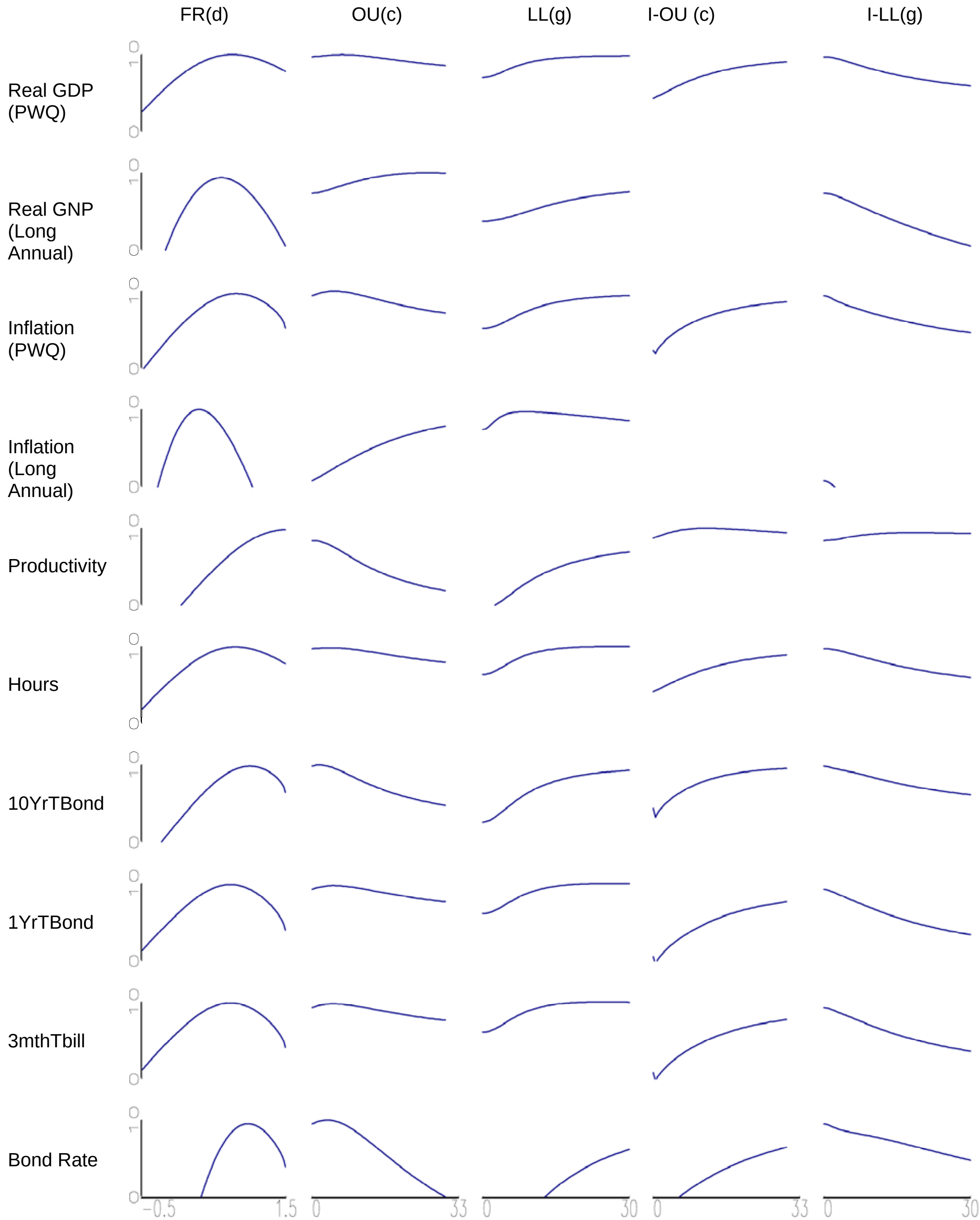
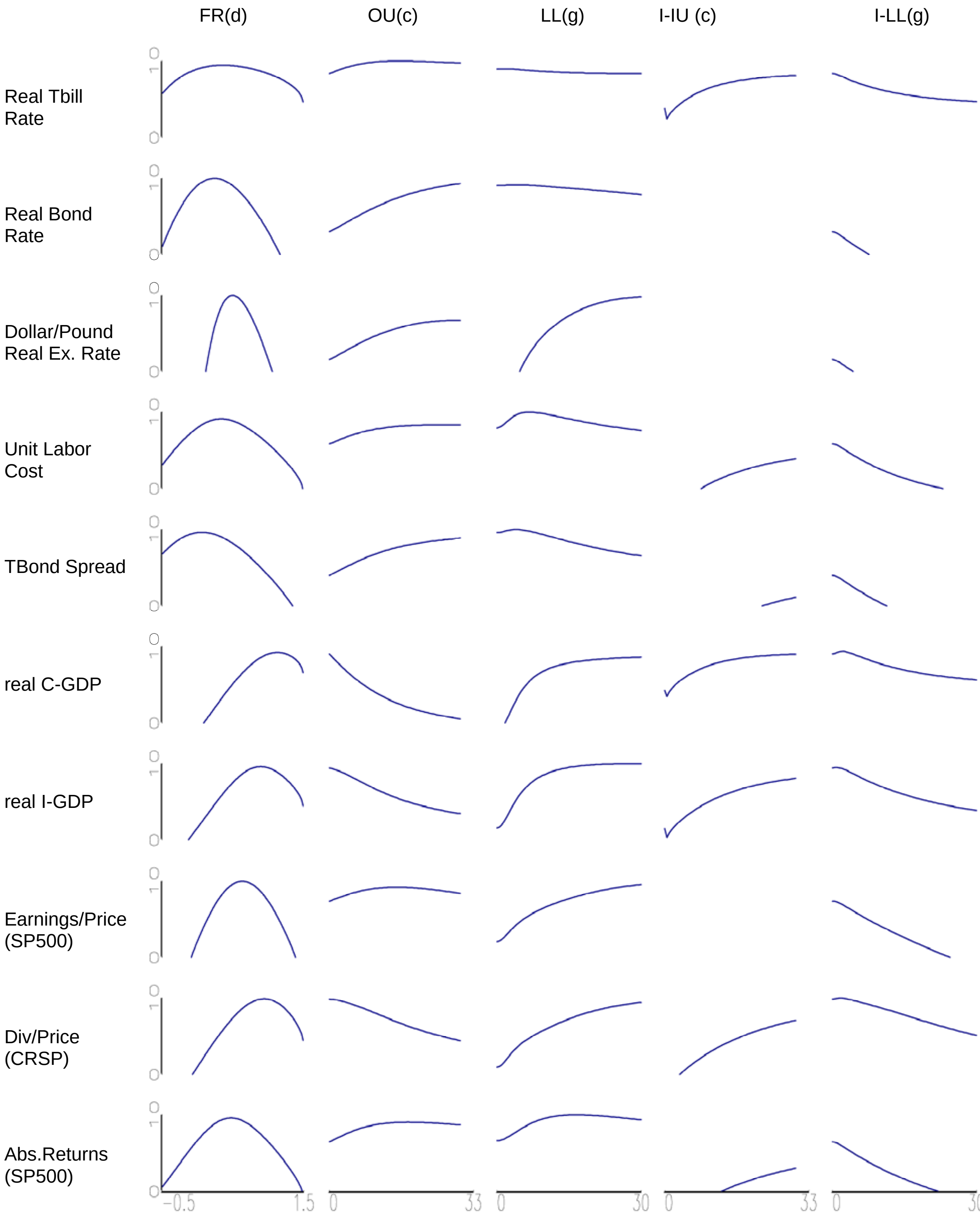


Figure 6 (Continued)
Low-Frequency Log-Likelihood Values



Notes to Figures

Figure 4: For each series the first panel plots the demeaned/detrended value of the series, the second panel plots the first difference of the series, and the final panel plots the low-frequency transformation v_T.

Figure 5: Each plot shows the *p*-value for the LBIM test (solid blue curve) and the H test (dotted red curve) computed using a fine grid of parameter values.

Figure 6: Each panel shows the log-likelihood value computed from the low-frequency maximal invariant v_T The likelihood values are normalized so that the value of the maximized log-likelihood across models is equal to 10.