Econometrica, Vol. 86, No. 3 (May, 2018), 775–804

LONG-RUN COVARIABILITY

ULRICH K. MÜLLER
Department of Economics, Princeton University

MARK W. WATSON
Department of Economics, Princeton University

We develop inference methods about long-run comovement of two time series. The parameters of interest are defined in terms of population second moments of low-frequency transformations ("low-pass" filtered versions) of the data. We numerically determine confidence sets that control coverage over a wide range of potential bivariate persistence patterns, which include arbitrary linear combinations of $I(0)$, $I(1)$, near unit roots, and fractionally integrated processes. In an application to U.S. economic data, we quantify the long-run covariability of a variety of series, such as those giving rise to balanced growth, nominal exchange rates and relative nominal prices, the unemployment rate and inflation, money growth and inflation, earnings and stock prices, etc.

KEYWORDS: Bandpass regression, fractional integration, great ratios.

1. INTRODUCTION

ECONOMIC THEORIES OFTEN HAVE stark predictions about the covariability of variables over long-horizons: consumption and income move proportionally (permanent income/life cycle model of consumption) as do nominal exchange rates and relative nominal prices (long-run PPP), the unemployment rate is unaffected by the rate of price inflation (vertical long-run Phillips curve), and so forth. But there is a limited set of statistical tools to investigate the validity of these long-run propositions. This paper expands this set of tools.

Two fundamental problems plague statistical inference about long-run phenomena. The first is the paucity of sample information: there are few "long-run" observations in the samples typically used in empirical analyses of long-run relations. The second is that inference critically depends on the data's long-run persistence. Random walks yield statistics with different probability distributions than i.i.d. data, for example, and observations from persistent autoregressions or fractionally integrated processes yield statistics with their own unique probability distributions. Taken together, these two problems conspire to make long-run inference particularly difficult: proper inference depends critically on the exact form of long-run persistence, but there is limited sample information available to empirically determine this form.

This paper develops methods designed to provide reliable inference about long-run covariability for a wide range of persistence patterns. The methods rely on a relatively small number of low-frequency averages of the data to measure the data's long-run variability and covariability. Our focus is on parameters that characterize the population second moments of these low-frequency averages.

Ulrich K. Müller: umueller@princeton.edu
Mark W. Watson: mwatson@princeton.edu

This work was presented as the Fisher–Schultz lecture at the 2016 European meetings of the Econometric Society. We thank David Papell, Ken West, three referees, and participants at several seminars for their comments. Support was provided by the National Science Foundation through Grant SES-1627660.

 https://doi.org/10.3982/ECTA15047

Our main contribution is to provide empirical researchers with a relatively easy-to-use method for constructing confidence intervals for long-run correlation coefficients, linear regression coefficients, and standard deviations of regression errors. These confidence intervals are valid for $I(0)$, $I(1)$, near unit roots, fractionally integrated models, and linear combinations of variables with these forms of persistence. Using a set of pre-computed "approximate least favorable distributions," the confidence intervals readily follow from the formulae discussed in Section 4.[1]

The outline of the paper is as follows. The next section introduces two empirical examples, the long-run relationship between consumption and GDP and between short- and long-term nominal interest rates, and defines the notion of long-run variability and covariability used throughout the paper. Our definition involves the population second moments of long-run projections, where these projections are similar to low-pass filtered versions of the data (e.g., Baxter and King (1999) and Hodrick and Prescott (1997)). In the long-run projections we employ, long-run covariability is equivalently captured by the covariability of a small number q of trigonometrically weighted averages of the data. The population second moments of the projections therefore correspond to an average of the spectrum (or pseudo-spectrum when the spectrum does not exist) over a narrow low-frequency band. Thus, this paper's long-run covariance parameters are those from low-frequency band spectrum regression (as in Engle (1974)), extended to allow for processes with more than $I(0)$ persistence. A key distinction between this paper and previous semiparametric approaches to the joint low-frequency behavior of persistent time series (see, for instance, Phillips (1991), Marinucci and Robinson (2001), Chen and Hurvich (2003), Robinson and Hualde (2003), Robinson (2008), or Shimotsu (2012)) is that in our asymptotic analysis, we keep q fixed as a function of the sample size. This ensures that the small-sample paucity of low-frequency information is reflected in our asymptotic approximations, as in Müller and Watson (2008), which yields more reliable inference in samples typically used in empirical macroeconomics.

Section 3 derives the large-sample normality of the q pairs of trigonometrically weighted averages and introduces a flexible parameterization of the joint long-run persistence properties of the underlying stochastic process. The large-sample framework developed in Section 3 reduces the problem of inference about long-run covariability parameters into the problem of inference about the covariance matrix of a $2q$-dimensional multivariate normal random vector. Section 4 reviews relevant methods for solving this parametric small-sample problem. Section 5 uses the resulting inference methods and post-WWII data from the United States to empirically study several familiar long-run relations involving balanced growth (GDP, consumption, investment, labor income, and productivity), the term structure of interest rates, the Fisher correlation (inflation and interest rates), the Phillips correlation (inflation and unemployment), PPP (exchange rates and price ratios), money growth and inflation, consumption growth and real returns, and the long-run relationship between stock prices, dividends, and earnings.

2. LONG-RUN PROJECTIONS AND COVARIABILITY

2.1. *Two Empirical Examples of Long-Run Covariability*

We begin by examining the long-run covariability of GDP and consumption and of short- and long-term nominal interest rates. These data motivate and illustrate the methods developed in this paper.

[1]The replication file contains a matlab function for computing these confidence intervals and is available at www.princeton.edu/~mwatson.

Consumption and income: One of the most celebrated and studied long-run relationships in economics concerns income and consumption. The long-run stability of the consumption/income ratio is one of economics' "Great Ratios" (Klein and Kosobud (1961)) and even a casual glance at the U.S. data suggests the two variables move together closely in the long run. Consider, for example, the evolution of U.S. real per-capita GDP and consumption over the post-WWII period. In the 17 years from 1948 through 1964, GDP increased by 62% and consumption increased by 52%. Over the next 17 years (1965–1981), both GDP and consumption grew more slowly, by only 30%. Growth rebounded during 1982 to 1998, when GDP grew by 43% and consumption increased 55%, but slowed again over 1999–2015 when GDP grew by only 17% and consumption increased by only 23%. Over these 17-year periods, there was substantial variability in the average annual rate of growth of GDP (2.9%, 1.4%, 2.1%, and 0.9% per year, respectively over the subsamples), and these changes were roughly matched by consumption (annual average growth rates of 2.5%, 1.5%, 2.6%, and 1.2%). Thus, over periods of 17 years, GDP and consumption exhibited substantial long-run variability and covariability in the post-WWII sample period.[2]

With this in mind, the first two panels of Figure 1 plot the average growth rates of GDP and consumption over six non-overlapping subsamples in 1948–2015. Figure 1(a) plots the average growth rates against time, and Figure 1(b) is a scatterplot of the six average growth rates for consumption against the corresponding values for GDP. Each of the six subsamples contains roughly 11 years, spans of history longer than the typical business cycle, and in this sense capture "long-run" variability in GDP and consumption. Average GDP and consumption growth over these subsamples exhibited substantial variability and (from the scatterplot) roughly one-for-one covariability.

Figure 1(c) sharpens the analysis by plotting "low-pass" transformations of the series designed to isolate variation in the series with periods longer than 11 years, computed as projections onto low-frequency periodic functions. These low-frequency projections produce series that are essentially the same as low-pass moving averages,[3] but are easier to analyze. These low-frequency projections are computed as follows. Let x_t, $t = 1, \dots, T$ denote a time series (e.g., growth rates of GDP or consumption). We use cosine functions for the periodic functions; let $\Psi_j(s) = \sqrt{2}\cos(js\pi)$ denote the function with period $2/j$ (where the factor $\sqrt{2}$ simplifies a calculation below), $\Psi(s) = [\Psi_1(s), \Psi_2(s), \dots, \Psi_q(s)]'$ denote a vector of these functions with periods 2 through $2/q$, and Ψ_T denote the $T \times q$ matrix with tth row given by $\Psi((t-1/2)/T)'$, so the jth column of Ψ_T has period $2T/j$. The GDP and consumption data in Figure 1 span $T = 272$ quarters, so setting $q = 12$ captures periodicities longer than $272/6 \approx 45$ quarters, or 11.3 years. The projection of x_t onto $\Psi((t-1/2)/T)$ for $t = 1, \dots, T$ yields the fitted values

$$\widehat{x}_t = X_T' \Psi\big((t-1/2)/T\big), \tag{1}$$

where X_T are the projection (linear regression) coefficients, $X_T = (\Psi_T'\Psi_T)^{-1}\Psi_T' x_{1:T}$, and $x_{1:T}$ is the $T \times 1$ vector with tth element given by x_t. The fitted values from these projections, say $(\widehat{x}_t, \widehat{y}_t)$, are plotted in Figure 1(c) and evolve much like the averages in

[2]Consumption is personal consumption expenditures (including durables) from the NIPA; Section 5 shows results for non-durables, services, and durables separately. Both GDP and consumption are deflated by the PCE deflator, so that output is measured in terms of consumption goods, and expressed in per-capita terms using the civilian non-institutionalized population over the age of 16. The Supplemental Material (Müller and Watson (2018)) contains data sources and descriptions for all data used in this paper.

[3]The Supplemental Material provides a comparison.

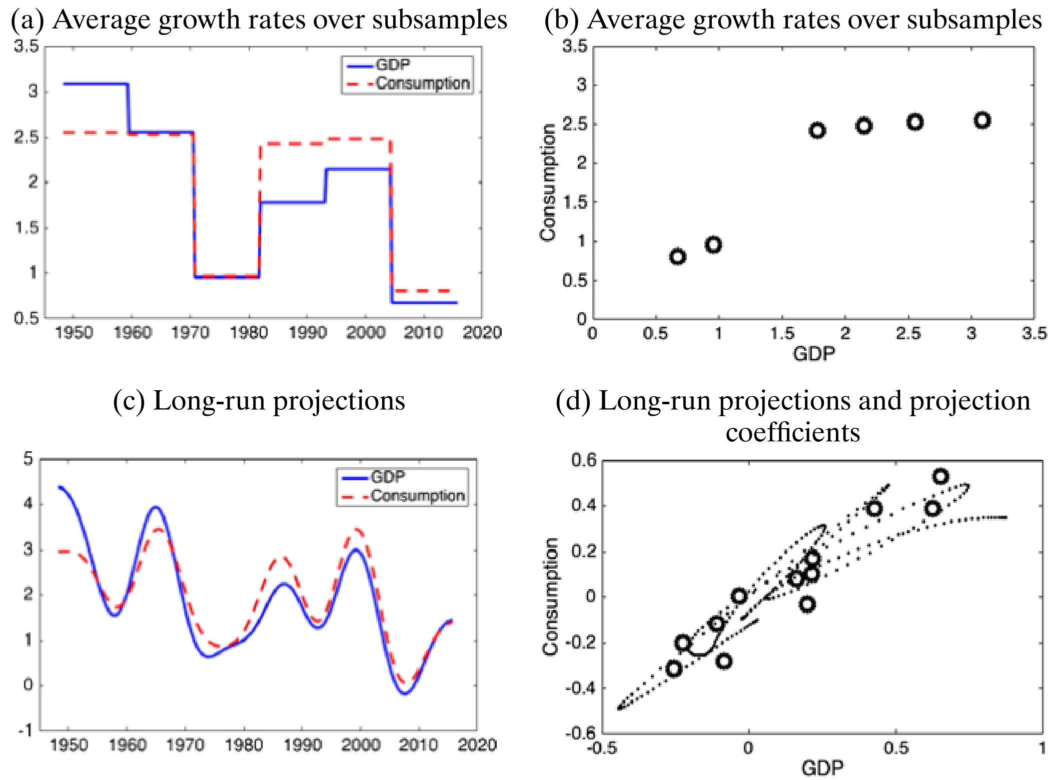

FIGURE 1.—Long-run average growth rates of consumption and GDP. *Notes*: Panel (a) shows sample averages of the variables over the period shown. Panel (b) is a scatterplot of the variables in (a). Panel (c) plots the projections of the data onto the low-frequency cosine terms discussed in the text, where sample means have been added to projections so they are consistent with the averages plotted in Figure 1(a). The small dots in panel (d) are a scatterplot of the variables in (c) (after scaling) and the large circles are a scatterplot of the projection coefficients (X_{jT}, Y_{jT}) from (c).

Figure 1(a), but better capture the smooth transition from high-growth to low-growth periods.

The matrix Ψ_T has two properties that simplify calculations and interpretation of the long-run projections. First, $\Psi_T' l_T = 0$, where l_T is a vector of ones, so that $\widehat{x}_t$ also corresponds to the projection of $x_t - \overline{x}_{1:T}$ onto $\Psi((t-1/2)/T)$, where $\overline{x}_{1:T}$ is the sample mean. More generally, X_T and $\widehat{x}_t$ are invariant to location shifts in the x_t-process, so with $x_t = \mu + u_t$, the properties of X_T and $\widehat{x}_t$ do not depend on the typically unknown value of μ.[4] Second, $T^{-1}\Psi_T'\Psi_T = I_q$, so X_T corresponds to simple cosine-weighted averages of the data (i.e., are the "cosine transforms" of $\{x_t\}$)

$$X_T = T^{-1}\Psi_T' x_{1:T}. \tag{2}$$

The orthogonality of the cosine regressors Ψ_T leads to a tight connection between the variability and covariability in the long-run projections $(\widehat{x}_t, \widehat{y}_t)$ plotted in Figure 1(c) and

[4]If the x_t-process contains a linear trend, say $x_t = \mu_0 + \mu_2 t + u_t$, then alternative periodic functions that are orthogonal to a time trend can be used so that X_T and $\widehat{x}_t$ do not depend on (μ_0, μ_1). See Müller and Watson (2008) for one set of such functions.

the cosine transforms (X_{jT}, Y_{jT}):

$$T^{-1}\sum_{t=1}^{T}\begin{pmatrix}\widehat{x}_t\\ \widehat{y}_t\end{pmatrix}\begin{pmatrix}\widehat{x}_t & \widehat{y}_t\end{pmatrix}=T^{-1}\begin{pmatrix}X_T'\\ Y_T'\end{pmatrix}\Psi_T'\Psi_T\begin{pmatrix}X_T & Y_T\end{pmatrix}=\begin{pmatrix}X_T'X_T & X_T'Y_T\\ Y_T'X_T & Y_T'Y_T\end{pmatrix}. \tag{3}$$

Thus, the sample covariability of the T time series projections $(\widehat{x}_t, \widehat{y}_t)$ coincides with the sample covariability of the q projection coefficients/cosine transforms (X_T, Y_T).[5] This is shown in Figure 1(d), which shows a scatterplot of (scaled versions) of the projections $(\widehat{x}_t, \widehat{y}_t)$, shown as small dots, and the projection coefficients (X_{jT}, Y_{jT}), shown as large circles. While the scatterplots capture the same variability and covariability in long-run movements in GDP and consumption growth, the projection coefficients eliminate much of the serial correlation evident in the $(\widehat{x}_t, \widehat{y}_t)$ scatterplot.

Short-term and long-term interest rates. The second empirical example involves short- and long-term nominal interest rates, as measured by the rate on 3-month U.S. Treasury bills, x_t, and the rate on 10-year U.S. Treasury bonds, y_t, from 1953 through 2015. Figure 2 plots the levels of short- and long-term interest rates, (x_t, y_t), along with their long-run projections, $(\widehat{x}_t, \widehat{y}_t)$, and cosine transforms, (X_T, Y_T). This sample includes only 63 years, so periodicities longer than 11 years are extracted using long-run projections with $q = 11$. Figure 2(a) shows that these long-run projections capture the rise in interest rates from the beginning of the sample through the early 1980s and then their subsequent decline. The projections for long-term interest rates closely track the projections for short-term rates and, given the connection between the projections and cosine transforms, X_{jT} and Y_{jT} are highly correlated (Figure 2(b)).

These two data sets differ markedly in their persistence: GDP and consumption growth rates are often modeled as low-order MA models, while nominal interest rates are highly serially correlated. Yet, the variables in both data sets exhibit substantial long-run variation and covariation which is readily evident in the long-run projections $(\widehat{x}_t, \widehat{y}_t)$ or equivalently (from (3)) the projection coefficients (X_T, Y_T). This suggests that the covariance/variance properties of (X_T, Y_T) are a useful starting point for defining the long-run covariability properties of stochastic processes exhibiting a wide range of persistent patterns.

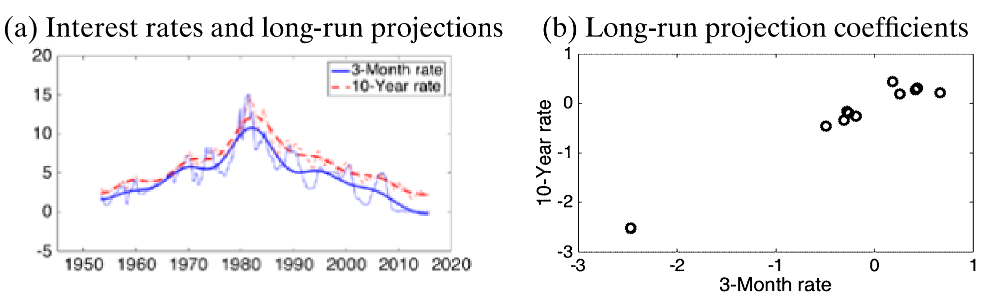

FIGURE 2.—Long-run projections of short- and long-term interest rates. *Notes*: Panel (a) plots the projections of the data onto the low-frequency cosine terms discussed in the text, where sample means have been added to projections so they are consistent with the levels of the data also shown in the figure. Panel (b) is a scatterplot of the projection coefficients (X_{jT}, Y_{jT}) from (a).

[5] Alternative low-frequency weights, such as Fourier transforms, have the same orthogonality properties and could be used in place of the cosine transforms. While the general analysis accommodates these alternative weights, our numerical analysis uses the cosine weights presented in the text.

2.2. *A Measure of Long-Run Covariability Using Long-Run Projections*

A straightforward definition of long-run covariability is based on the population analogue of the sample second moment matrices in (3). Let Σ_T denote the covariance matrix of $(X_T', Y_T')'$, partitioned as $\Sigma_{XX,T}$, $\Sigma_{XY,T}$, etc., and define

$$\begin{aligned}\Omega_T &= T^{-1}\sum_{t=1}^{T} E\left[\begin{pmatrix}\widehat{x}_t\\ \widehat{y}_t\end{pmatrix}\begin{pmatrix}\widehat{x}_t & \widehat{y}_t\end{pmatrix}\right]\\ &= \sum_{j=1}^{q} E\left[\begin{pmatrix}X_{jT}\\ Y_{jT}\end{pmatrix}\begin{pmatrix}X_{jT}\\ Y_{jT}\end{pmatrix}'\right] = \begin{pmatrix}\mathrm{tr}(\Sigma_{XX,T}) & \mathrm{tr}(\Sigma_{XY,T})\\ \mathrm{tr}(\Sigma_{YX,T}) & \mathrm{tr}(\Sigma_{YY,T})\end{pmatrix},\end{aligned} \tag{4}$$

where the equalities directly follow from (3).

The 2×2 matrix Ω_T is the average covariance matrix of the long-run projections $(\widehat{x}_t, \widehat{y}_t)$ in a sample of length T, and provides a summary of the variability and covariability of the long-run projections over repeated samples. Equivalently, by the second equality, Ω_T also measures the covariability of the cosine transforms (X_T, Y_T). Corresponding long-run correlation and linear regression parameters follow from the usual formulae

$$\begin{aligned}\rho_T &= \Omega_{xy,T}/\sqrt{\Omega_{xx,T}\Omega_{yy,T}},\\ \beta_T &= \Omega_{xy,T}/\Omega_{xx,T},\\ \sigma^2_{y|x,T} &= \Omega_{yy,T} - (\Omega_{xy,T})^2/\Omega_{xx,T},\end{aligned} \tag{5}$$

where $(\Omega_{xy,T}, \Omega_{xx,T}, \Omega_{yy,T})$ are the elements of Ω_T. The linear regression coefficient β_T solves the population least-squares problem

$$\beta_T = \operatorname*{argmin}_b E\left[T^{-1}\sum_{t=1}^{T}(\widehat{y}_t - b\widehat{x}_t)^2\right], \tag{6}$$

so that β_T is the coefficient in the population best linear prediction of the long-run projection $\hat{y}_t$ by the long-run projection $\hat{x}_t$, $\sigma^2_{y|x,T}$ is the average variance of the prediction error, and ρ_T^2 is the corresponding population R^2. These parameters thus measure the population linear dependence of the long-run variation of (x_t, y_t). Equivalently, by the second equality in (4), β_T also solves

$$\beta_T = \operatorname*{argmin}_b E\left[\sum_{j=1}^{q}(Y_{jT} - bX_{jT})^2\right] \tag{7}$$

with a corresponding interpretation for $\sigma^2_{y|x,T}$ and ρ_T^2. Thus, these parameters equivalently measure the (population) linear dependence in the scatterplots in Figures 1(d) and 2(b).

The covariance matrix, Ω_T, or equivalently, $(\rho_T, \beta_T, \sigma^2_{y|x,T})$, are the long-run population parameters that are the focus of our analysis. These parameters depend on the periods used to define the "long run," that is, the value of q used to construct the long-run projections. In the empirical examples discussed above, we chose periods longer than 11 years, that is, periods longer than the U.S. business cycle. This led us to use $q = 12$ for GDP and consumption (with sample size $T = 272$ quarters) and $q = 11$ for interest rates (with

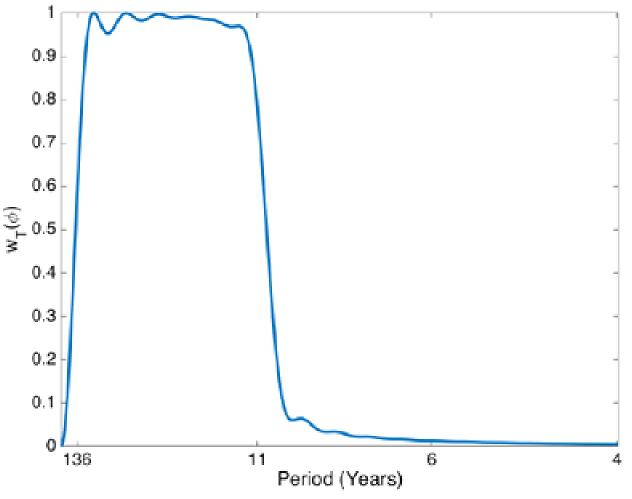

FIGURE 3.—Spectral weight for low frequencies with $T = 272$ quarters and $q = 12$. *Notes*: The figure plots $|\sum_{j=1}^{q}\sum_{t=1}^{T}\Psi_j((t-0.5)/T)e^{-\mathrm{i}t\phi}|^2$ for frequencies $0 \le \phi \le 2\pi/16$. The horizontal axis is labeled with periods measured in years ($= 0.5\pi/\phi$).

sample size $T = 252$ quarters). If we had instead been interested in periods longer than 20 years, we would have chosen $q = 7$ for GDP/consumption ($2T/q \approx 78$ quarters) and $q = 6$ for interest rates ($2T/q = 84$ quarters). The important point is that q defines the long-run periods of interest for the research question at hand.

2.3. *Frequency Domain Interpretation of Ω_T and* $(\rho_T, \beta_T, \sigma^2_{y|x,T})$

The covariance matrix Ω_T has a natural and familiar frequency domain interpretation. Since (X_T, Y_T) are weighted averages of the data $z_t = (x_t, y_t)'$, Ω_T is a weighted average of the variances and covariances of z_t. If the time series are covariance stationary with spectral density matrix $F_z(\phi)$, these variances and covariances are weighted averages of the spectrum over different frequencies ϕ. In fact, a straightforward calculation shows (see the Supplemental Material) that $\Omega_T = (2\pi)^{-1}\int_{-\pi}^{\pi} F_z(\phi)w_T(\phi)\,d\phi$, where $w_T(\phi) = |\sum_{j=1}^{q}\sum_{t=1}^{T}\Psi_j((t-1/2)/T)e^{-\mathrm{i}t\phi}|^2$ and $\mathrm{i} = \sqrt{-1}$.

Figure 3 plots the weights $w_T(\phi)$ for $T = 272$ quarters (the GDP-consumption sample size) and $q = 12$, where the horizontal axis shows periods (in years) instead of frequency (annual period $= 2\pi/(4\phi)$). The figure shows that Ω_T is essentially a bandpass version of the spectrum with periods between $2T$ (136 years) and $T/6$ (11.3 years) corresponding to cosine transforms with $j = 1$ through $j = 12$. Thus, Ω_T and the associated values of $(\rho_T, \beta_T, \sigma^2_{y|x,T})$ are bandpass regression parameters, as in Engle (1974), for a particular low-frequency band. If, as in Engle's analysis, the data are generated by an $I(0)$ stochastic process, the spectrum is approximately flat over this band and inference follows relatively directly from classic results on spectral estimators (e.g., Brillinger (2001), Brockwell and Davis (1991)). However, if the process is not $I(0)$ in the sense that the spectrum (or pseudo-spectrum) is not flat over this low-frequency band, $I(0)$ procedures lead to faulty inference akin to Granger and Newbold's (1974) spurious regressions. The goal of our analysis is to develop inference procedures that are robust to this $I(0)$-flat-spectrum assumption.

As discussed above, the upper frequency cutoff $2T/q$, corresponding to 11 years in Figure 3, represents the highest low-frequency (shortest period) of interest for the researcher's analysis and is problem-specific. The lower cutoff $2T$, corresponding to 136 years in Figure 3, is induced by the invariance to location shifts of the cosine transforms. Without knowledge of the population means, it is not possible to extract empirical information about arbitrarily low frequencies, and our estimand Ω_T reflects this impossibility. What is more, the fact that the weight $w_T(\phi)$ converges to zero as $\phi \to 0$ keeps our estimand Ω_T well-defined even for some (pseudo) spectra that diverge at frequency zero, such as for $I(1)$ processes.

3. ASYMPTOTIC APPROXIMATIONS AND PARAMETERIZING LONG-RUN PERSISTENCE AND COVARIABILITY

The long-run correlation and regression parameters from Ω_T are functions of Σ_T, the covariance matrix of (X_T, Y_T). This section takes up two related issues. The first is the asymptotic normality of the cosine-weighted averages (X_T, Y_T). This serves as the basis for the inference methods developed in Section 4 and provides large-sample approximation for the matrices Σ_T and Ω_T. The second issue is a parameterization of long-run persistence and comovement that determines the large-sample value of Σ_T and Ω_T.

3.1. *Large-Sample Properties of Long-Run Sample Averages*

Because (X_T, Y_T) are smooth averages of (x_t, y_t), a central limit theorem effect suggests that these averages are approximately Gaussian random variables under a range of primitive conditions about (x_t, y_t). The set of assumptions under which asymptotic normality holds turns out to be reasonably broad, and encompasses many forms of potential persistence. Specifically, with $z_t = (x_t, y_t)'$, suppose that Δz_t has moving average representation $\Delta z_t = C_T(L)\varepsilon_t$, where ε_t is a martingale difference sequence with non-singular covariance matrix, Δz_t has a spectral density $F_{\Delta z,T}$, and ε_t and $C_T(L)$ satisfy other moment and decay restrictions given in Müller and Watson (2017, Theorem 1). The dependence of C_T and $F_{\Delta z,T}$ on the sample size T accommodates many forms of persistence that require double arrays as data generating process, such as autoregressive roots of the order $1 - c/T$, for fixed c.[6]

If the spectral density converges for all frequencies close to zero

$$T^2 F_{\Delta z,T}(\omega/T) \to S_{\Delta z}(\omega) \tag{8}$$

in a suitable sense, then

$$\sqrt{T}\begin{pmatrix} X_T \\ Y_T \end{pmatrix} \Rightarrow \begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}(0, \Sigma), \tag{9}$$

and the finite-sample second moment matrix correspondingly converges to its large-sample counterpart (Müller and Watson (2017, Lemma 2)):

$$T \operatorname{Var}\begin{pmatrix} X_T \\ Y_T \end{pmatrix} = T\Sigma_T \to \Sigma. \tag{10}$$

[6]We omit the corresponding dependence of $z_t = (x_t, y_t)'$ on T to ease notation.

The scaling necessary to achieve (8) is subsumed in $C_T(L)$. For instance, if x_t is a random walk, and y_t is i.i.d., then $C_T(L) = \mathrm{diag}(1/T, 1-L)$ induces (8).

The limiting covariance matrix Σ in (9) and (10) is a function of the (pseudo) "local-to-zero spectrum" $S_z(\omega) = S_{\Delta z}(\omega)/\omega^2$ of z_t and the cosine weights $\Psi_j(s)$ that determine (X_T, Y_T); specifically, from Müller and Watson (2017),

$$\Sigma = \int_{-\infty}^{\infty} \left(I_2 \otimes \int_0^1 e^{\mathrm{i}\lambda s} \Psi(s)\, ds \right)' S_z(\omega) \left(I_2 \otimes \int_0^1 e^{-\mathrm{i}\lambda s} \Psi(s)\, ds \right) d\omega.$$

These limiting results are consistent with the results for $T = 272$ shown in Figure 3 in the previous section: (X_T, Y_T) are low-frequency weighted averages of the data and their covariance matrix depends on the (pseudo-) spectrum of (x_t, y_t) in a small band around frequency zero.

We make three comments about these large-sample results. First, they hold when the *first difference* of z_t has a spectral density; the *level* of z_t is more persistent than its first difference and may have a (pseudo-) spectrum that diverges at frequency zero. In this case, Σ remains finite because the cosine averages sum to zero ($\Psi_T' l_T = 0$), so they do not extract zero-frequency variation in the data. If the *level* of z_t has a spectral density, then this restriction on the weights is not required and, for example, the centered sample mean of z_t also has a large-sample normal limit. Second, for $z_t \sim I(d)$, the decay restrictions on $C_T(L)$ allow values of $d \in (-0.5, 1.5)$, which allows a reasonably wide range of persistent processes, but rules out some models of practical interest. For example, the first difference of an $I(0)$ process appended to a linear trend (i.e., the first difference of a "trend-stationary" process) is $I(-1)$ and is ruled out. And, of course, it does not accommodate $I(d)$ processes with $d > 1.5$. Third, because $T\Sigma_T \to \Sigma$, also $T\Omega_T \to \Omega$ where Ω is defined as in the last expression of (4) with Σ in place of Σ_T. Correspondingly, $(\rho_T, \beta_T, T\sigma^2_{y|x,T}) \to (\rho, \beta, \sigma^2_{y|x})$ with the limits defined by (5) with Ω in place of Ω_T. Thus, a solution to the problem of inference about $(\rho, \beta, \sigma^2_{y|x})$ given observations (X, Y) readily translates into a solution to large-sample valid inference about $(\rho_T, \beta_T, \sigma^2_{y|x,T})$ given (X_T, Y_T), and, by invoking the arguments in Müller (2011), efficient inference in the former problem amounts to large-sample efficient inference in the latter problem.

3.2. *Parameterizing Long-Run Persistence and Covariability*

The limiting average covariance matrix of the long-run projections, Ω, is a function of the covariance matrix of the cosine projections, Σ, which in turn is a function of the local-to-zero spectrum S_z of z_t. In this section, we discuss parameterizations of S_z, and thus Σ and Ω.

It is constructive to consider two leading examples. In the first, z_t is $I(0)$ with long-run covariance matrix Λ. In this case, while the spectrum $F_{z,T}(\phi)$ potentially varies across frequencies $\phi \in [-\pi, \pi]$, the local-to-zero spectrum is flat $S_z(\omega) \propto \Lambda$. Straightforward calculations then show that $\Sigma = \Lambda \otimes I_q$ and $\Omega = \Lambda$, so the covariance matrix associated with the long-run projections corresponds to the usual long-run $I(0)$ covariance matrix. In this model, the cosine transforms (X_{jT}, Y_{jT}) plotted in Figures 1 and 2 are, in large samples and up to a deterministic scale, i.i.d. draws from a $\mathcal{N}(0, \Lambda)$ distribution. Inference about $\Omega = \Lambda$ and $(\rho, \beta, \sigma^2_{y|x})$ thus follows from well-known small-sample inference procedures for Gaussian data (see Müller and Watson (2017)). In the second example, z_t is $I(1)$ with Λ the long-run covariance matrix for Δz_t. In this case, $S_z(\omega) \propto \omega^{-2}\Lambda$, and

TABLE I

LONG-RUN COVARIABILITY ESTIMATES AND CONFIDENCE INTERVALS USING THE $I(0)$ AND $I(1)$ MODELS: PERIODS LONGER THAN 11 YEARS

		ρ	β	$\sigma_{y\|x}$
		a. GDP and consumption		
$I(0)$	Estimate	0.93	0.76	0.36
	67% CI	0.87, 0.96	0.67, 0.85	0.30, 0.46
	90% CI	0.80, 0.97	0.60, 0.92	0.27, 0.55
$I(1)$	Estimate	0.93	0.84	0.35
	67% CI	0.88, 0.96	0.74, 0.94	0.29, 0.45
	90% CI	0.82, 0.97	0.66, 1.01	0.26, 0.54
		b. Short- and long-term interest rates		
$I(0)$	Estimate	0.98	0.96	0.60
	67% CI	0.96, 0.98	0.90, 1.03	0.50, 0.79
	90% CI	0.93, 0.99	0.84, 1.08	0.44, 0.96
$I(1)$	Estimate	0.97	0.93	0.38
	67% CI	0.93, 0.98	0.85, 1.01	0.32, 0.50
	90% CI	0.90, 0.98	0.78, 1.07	0.28, 0.61

Notes: Periods longer than 11 years correspond to $q = 12$ for panel (a) and $q = 11$ for panel (b). The rows labeled "Estimate" are the maximum likelihood estimates using the large-sample distribution of the cosine transforms for the $I(0)$ and $I(1)$ models and the rows labeled "CI" denote the associated confidence intervals.

a calculation shows that $\Sigma = \Lambda \otimes D$, where D is a $q \times q$ diagonal matrix with jth diagonal element $D_{jj} = (j\pi)^{-2}$. In this model, the cosine transforms (X_{jT}, Y_{jT}) plotted in Figures 1 and 2 are, in large samples and up to a deterministic scale, independent but heteroscedastic draws from $\mathcal{N}(0, (j\pi)^{-2}\Lambda)$ distributions. Thus $\Omega \propto \Lambda$, so the covariance matrix for long-run projections for z_t corresponds to the long-run covariance matrix for its first differences, Δz_t. By weighted least-squares logic, inference for $I(1)$ processes follows after reweighting the elements of (X_{jT}, Y_{jT}) by the square roots of the inverse of the diagonal elements of D and then using the same methods as in the $I(0)$ model.

GDP and consumption and short-term and long-term interest rates: Table I presents estimates and confidence sets for $(\rho_T, \beta_T, \sigma_{y|x,T})$ using (X_T, Y_T) for GDP and consumption (panel (a)) and for short- and long-term interest rates (panel (b)), where the focus is on periods longer than 11 years, so that $q = 12$ for panel (a) and $q = 11$ for panel (b). Results are presented for $I(0)$ and $I(1)$ models; a more general model of persistence is introduced below. The point estimates shown in the table are MLEs, and confidence intervals for $(\beta_T, \sigma^2_{y|x,T})$ are computed using standard finite-sample normal linear regression formulae (after appropriate weighting in $I(1)$ model), and confidence sets for ρ_T are constructed as in Anderson (1984, Section 4.2.2).

For GDP and consumption, there are only minor differences between the $I(0)$ and $I(1)$ estimates and confidence sets. The estimated long-run correlation is greater than 0.9, and the lower range of the 90% confidence interval exceeds 0.8 in both the $I(0)$ and $I(1)$ models. Thus, despite the limited long-run information in the sample (captured here by the 12 observations making up (X_T, Y_T)), the evidence points to a large long-run correlation between GDP and consumption. The long-run regression of consumption onto GDP yields a regression coefficient that is estimated to be 0.76 in the $I(0)$ model and 0.84 in the $I(1)$ model. This estimate is sufficiently accurate that $\beta = 1$ is not included in the 90% $I(0)$ confidence set. The results for long-term and short-term nominal interest rates are similarly informative: there is strong evidence that the series are highly correlated over the long run.

The validity of these confidence sets rests on the quality of the $I(0)$ and $I(1)$ models for approximating the spectral shape over the low-frequency band plotted in Figure 3. The $I(0)$ model assumes the spectrum is flat over this band, so the series behave like white noises for periods longer than 11 years, while the $I(1)$ model assumes random-walk behavior. Neither assumption is particularly compelling. Moreover, as we show in Table IV below, the $I(0)$ assumption yields confidence intervals with coverage probability far below the nominal level when, in fact, the data were generated by the $I(1)$ model, and vice versa, and both yield faulty inference when persistence is something other than $I(0)$ or $I(1)$. With this motivation, the next subsection proposes a more flexible parameterization of persistence.

3.2.1. (A, B, c, d) *Model*

The shape of the local-to-zero spectrum determines the long-run persistence properties of the data, and misspecification of this persistence leads to faulty inference about long-run covariability. Simply put, reliable inference requires a parameterization of the spectrum that yields a good approximation to the persistence patterns of the variables over the low-frequency band under study. Addressing this issue faces a familiar trade-off: the parameterization needs to be sufficiently flexible to yield reliable inference about long-run covariability for a wide range of economically relevant stochastic processes and yet be sufficiently parsimonious to allow meaningful inference with limited sample information. $I(0)$ persistence generates a flat local-to-zero spectrum, and $I(1)$ persistence generates a local-to-zero spectrum proportional to ω^{-2}. Both of these models are parsimonious, but tightly constrain the spectrum. This limits their usefulness as general models for conducting inference about long-run covariability.

With this trade-off in mind, we use a parameterization that nests and generalizes a range of models previously used to model persistence in economic time series. The parameterization is a bivariate extension of the univariate (b, c, d) model used in Müller and Watson (2016) and yields a local-to-zero spectrum of the form

$$S_z(\omega) \propto A \begin{pmatrix} (\omega^2 + c_1^2)^{-d_1} & 0 \\ 0 & (\omega^2 + c_2^2)^{-d_2} \end{pmatrix} A' + BB', \tag{11}$$

where A and B are 2×2 matrices with A unrestricted and B lower triangular.[7]

The primary motivation for this (A, B, c, d) model is as a parsimonious but flexible functional form for the local-to-zero spectrum. It combines and generalizes several standard spectral shapes. For example, with $A = 0$, it is the $I(0)$ local-to-zero spectrum. When $B = 0$, $c = 0$, and $d_1 = d_2 = 1$, it yields the $I(1)$ spectrum; $B = 0$, $d_1 = d_2 = 1$ yields a spectrum for arbitrary linear combinations of two independent local-to-unity processes with mean reversion parameters equal to c_1 and c_2; $B = 0$ and $c = 0$ yields a bivariate fractional spectrum with parameters d_1 and d_2. Other choices of (A, B, c, d) yield spectra from models that combine persistent and non-persistent components (as in cointegrated or "local-level" models) but go beyond the usual $I(0)/I(1)$ or fractional formulations.

[7]This is the spectrum of a bivariate Whittle–Matérn (cf. Lindgren (2013)) process with time series representation $z_t = A\tau_t + e_t$, where $\tau_t = (\tau_{1t}, \tau_{2t})'$ is a bivariate process with uncorrelated $\{\tau_{1t}\}$ and $\{\tau_{2t}\}$, $(1 - \phi_{i,T} L)^{d_i} \tau_{it} = T^{-d_i/2} \varepsilon_{it}$, $\phi_{i,T} = 1 - c_i/T$, $\varepsilon_t \sim I(0)$ with long-run variance equal to I_2, $e_t \sim I(0)$ with long-run variance equal to BB', and zero long-run covariance with ε_t.

The nesting of the cointegrated model in the (A, B, c, d) model is particularly interesting because it, too, focuses on long-run relationship between the variables. Formally, of course, cointegration concerns common patterns of *persistence* in the variables, not their *variance* and *covariance*: in its canonical form, x_t and y_t are cointegrated if both are $I(1)$ and yet a linear combination of the variables is $I(0)$. This implies that the variables share a single $I(1)$ trend in addition to $I(0)$ components. When the innovations of the $I(1)$ and $I(0)$ components are of the same order of magnitude, then the spectrum is dominated by the $I(1)$ component over low frequencies. Thus, in the context of the (A, B, c, d) model, the canonical cointegration model corresponds to the restrictions $B = 0$ (because the marginal processes are dominated by the stochastic trends), $c_1 = c_2 = 0$ and $d_1 = d_2 = 1$ (so the trend components are $I(1)$), and A has rank 1 (because there is a single common trend). The singularity in A in turn induces singularities in S_z and Σ. The large-T limit of the usual formulation of cointegration thus implies that long-run projections computed with a fixed value of q are perfectly correlated, and the scatterplot of projection coefficients lies on a straight line with slope corresponding to the cointegrating coefficient. Of course, this perfect correlation will not obtain with finite T, so practically useful approximations would require a non-singular value of A and/or a non-zero value of B.

3.2.2. *Beyond the (A, B, c, d) Model*

While the (A, B, c, d) parameterization encompasses many standard models, it is useful to highlight some models that are *not* encompassed by (11). We discuss two here.

One restriction of the (A, B, c, d) model is the asymptotic independence of the persistent and $I(0)$ components, as captured by A and B, respectively. One model where this independence is restrictive is in a cointegrated model with a local-to-zero cointegration coefficient and correlated $I(0)$ and trend components. Consider, for instance, a model where x_t is the stochastic trend, and y_t is a linear combination of an $I(0)$ error correction term and x_t, with a coefficient on x_t that is of order $1/T$.[8] Allowing x_t to follow a local-to-unity process, the model is

$$\begin{aligned} x_t &= (1 - c/T)x_{t-1} + u_{x,t}, \\ y_t &= \frac{\eta}{T} x_t + u_{y,t}, \end{aligned} \tag{12}$$

where $u_t = (u_{x,t}, u_{y,t})$ is $I(0)$ with long-run covariance Σ_u and elements σ_x^2, σ_y^2, and $\lambda\sigma_x\sigma_y$ in obvious notation. The corresponding local-to-zero spectral density is

$$S_z(\omega) \propto \begin{pmatrix} \sigma_x & 0 \\ \eta & \sigma_y \end{pmatrix} \begin{pmatrix} (\omega^2 + c^2)^{-1} & \lambda(c + \mathrm{i}\omega)^{-1} \\ \lambda(c - \mathrm{i}\omega)^{-1} & 1 \end{pmatrix} \begin{pmatrix} \sigma_x & \eta \\ 0 & \sigma_y \end{pmatrix}. \tag{13}$$

This "local cointegration" spectral density is outside the (A, B, c, d) model whenever $\lambda \neq 0$. The complex-valued local-to-zero spectral density indicates the presence of lead and lag relationships that span a non-trivial fraction of the sample size. Since the level of x_t depends in a non-negligible manner on values of $u_{x,s}$ with $s \ll t$, x_t is correlated with values of $u_{y,s}$ in the distant past when $\lambda \neq 0$.

In this model, a calculation shows that the population regression of Y onto X has a regression coefficient $\beta = \eta + \lambda c$, so it depends on the local cointegrating coefficient η, the

[8]Given the invariance restrictions we impose in Section 4.2.1 below, such a transformation of the cointegration model is without loss of generality for inference about β_T and $\sigma^2_{y|x,T}$.

$I(0)$ correlation coefficient λ, and the persistence parameter c. In contrast, previous analyses of the model (12) (see, for instance, Elliott (1998) or Jansson and Moreira (2006)) focused on the cointegration parameter η. This difference emphasizes the distinct goals of the analyses: The cointegration parameter η yields the linear combination of (x, y) with minimum *persistence*, while the regression parameter β yields the linear combination with minimum *variance*. In general, there is no reason to expect that these are the same, and this makes it is surprising that these two goals yield the same estimand ($\beta = \eta$) in the canonical cointegration model with $c = 0$, even for $\lambda \neq 0$.

There are many other ways of modeling the joint long-run properties of z_t. As already mentioned, since the (A, B, c, d) model has a real-valued local-to-zero spectrum, it rules out long-span leads and lags between the series. One simple way to generate such lags is via

$$\begin{aligned} x_t &= A_{11}u_{1,t} + A_{12}u_{2,t-\lfloor \delta T \rfloor}, \\ y_t &= A_{21}u_{1,t} + A_{22}u_{2,t-\lfloor \delta T \rfloor}, \end{aligned} \tag{14}$$

where $u_t \sim I(0)$ as above. The parameter $\delta \in \mathbb{R}$ measures the lead of $u_{1,t}$ relative to $u_{2,t-\lfloor \delta T \rfloor}$ as a fraction of the sample size. In this "$I(0)$ long-lag" model, the local-to-zero spectral density satisfies

$$S_z(\omega) \propto A \begin{pmatrix} \sigma_x & \lambda \sigma_x \sigma_y \exp(\mathrm{i}\delta\omega) \\ \lambda \sigma_x \sigma_y \exp(-\mathrm{i}\delta\omega) & \sigma_y \end{pmatrix} A' \tag{15}$$

with $[A] = A_{ij}$. When $\lambda \neq 0$, the long lags relating x and y yield a complex local-to-zero cross spectrum which is outside the (A, B, c, d) class.

Our construction does not guarantee that confidence sets have their desired coverage probabilities outside the (A, B, c, d) model; we present numerical results in Section 4.3 below that investigate the magnitude of inference errors associated with models (13) and (15).

4. CONSTRUCTING CONFIDENCE INTERVALS FOR ρ, β, AND $\sigma_{y|x}$

4.1. *An Overview*

There are several approaches one might take to construct confidence intervals for the parameters ρ, β, and $\sigma_{y|x}$ from the observations (X, Y). As a general matter, the goal is to compute confidence intervals that are as informative ("short") as possible, subject to the coverage constraint that they contain the true value of the parameter of interest with a pre-specified probability. We construct confidence intervals by explicitly solving a version of this problem.

Generically, let θ denote the vector of parameters characterizing the probability distribution of (X, Y), and let Θ denote the parameter space. (In our context, θ denotes the (A, B, c, d) parameters.) Let $\gamma = g(\theta)$ denote the parameter of interest. ($\gamma = \rho$, β, or $\sigma_{y|x}$ for the problem we consider.) Let $H(X, Y)$ denote a confidence interval for γ and $\text{lgth}(H(X, Y))$ denote the length of the interval. The objective is to choose H so it has small expected length, $E[\text{lgth}(H(X, Y)]$, subject to coverage, $P(\gamma \in H(X, Y)) \geq 1 - \alpha$, where α is a pre-specified constant. Because the probability distribution of (X, Y) depends on θ, so will the expected length of $H(X, Y)$ and the coverage probability. By definition, the coverage constraint must be satisfied for all values of $\theta \in \Theta$, but one has freedom in choosing the value of θ over which expected length is to be minimized. Thus, let W denote a distribution that puts weight on different values of θ, so the problem becomes

$$\min_H \int E_\theta\big[\mathrm{lgth}\big(H(X,Y)\big)\big]\,dW(\theta) \tag{16}$$

subject to

$$\inf_{\theta\in\Theta} P_\theta\big(\gamma \in H(X,Y)\big) \geq 1-\alpha, \tag{17}$$

where the objective function (16) emphasizes that the expected volume depends on the value of θ, with different values of θ weighted by W, and the coverage constraint (17) emphasizes that the constraint must hold for all values of θ in the parameter space Θ.

As noted by Pratt (1961), the expected length of confidence set for γ can be expressed in terms of the power of hypothesis tests of $H_0 : \gamma = \gamma_0$ versus $H_1 : \theta \sim W$. The solution to (16)–(17) thus amounts to the determination of a family of most powerful hypothesis tests, indexed by γ_0. Elliott, Müller, and Watson (2015) suggested a numerical approach to compute corresponding approximate "least favorable distributions" for θ. We implement a version of those methods here; details are provided in the Supplemental Material. A key feature of the solution is that, conditional on the weighting function W and the least favorable distribution, the confidence sets have the familiar Neyman–Pearson form with a version of the likelihood ratio determining the values of γ included in the confidence interval.

While the resulting confidence intervals have smallest weighted expected length (up to the bounds used in the numerical approximation of the least favorable distributions), they can have unreasonable properties for particular realizations of (X,Y). Indeed, for some values of (X,Y), the confidence intervals might be empty, with the uncomfortable implication that, conditional on observing these values of (X,Y), one is certain that the confidence interval excludes the true value. To avoid this, we follow Müller and Norets (2016) and restrict the confidence sets to be supersets of $1-\alpha$ equal-tailed Bayes credible sets.

4.2. *Some Specifics*

4.2.1. *Invariance and Equivariance*

Correlations are invariant to the scale of the data. The linear regression of y_t onto x_t is the same as the regression of $y_t + bx_t$ onto x_t after subtracting b from the latter's regression coefficient. It is sensible to impose the same invariance/equivariance on the confidence intervals. Thus, letting H^ρ, H^β, and H^σ denote confidence sets for ρ, β, and $\sigma_{y|x}$, we restrict these sets as follows:

$$\rho \in H^\rho(X,Y) \quad \Leftrightarrow \quad \rho \in H^\rho(b_xX, b_yY) \text{ for } b_xb_y > 0, \tag{18}$$

$$\begin{aligned} &\beta \in H^\beta(X,Y) \\ &\quad \Leftrightarrow \quad \frac{b_y\beta + b_{yx}}{b_x} \in H^\beta(b_xX, b_yY + b_{yx}X) \text{ for } b_x, b_y \neq 0 \text{ and all values of } b_{yx}, \end{aligned} \tag{19}$$

$$\begin{aligned} &\sigma_{y|x} \in H^\sigma(X,Y) \\ &\quad \Leftrightarrow \quad |b_y|\sigma_{y|x} \in H^\sigma(b_xX, b_yY + b_{yx}X) \text{ for } b_x, b_y \neq 0 \text{ and all values of } b_{yx}. \end{aligned} \tag{20}$$

These invariance/equivariance restrictions lead to two modifications to the solution to (16)–(17). First, they require the use of maximal invariants in place of the original (X, Y). The density of the maximal invariants for each of these transformations is derived in the Supplemental Material. Second, because the objective function (16) is stated in terms of (X, Y), minimizing expected length by inverting tests based on the maximal invariant leads to a slightly different form of optimal test statistic. Müller and Norets (2016) developed these modifications in a general setting, and the Supplemental Material derives the resulting form of confidence sets for our problem.

4.2.2. *Parameter Space*

We use the following parameter space for $\theta = (A, B, c, d)$: A and B are real, with B lower-triangular and (A, B) chosen so that Ω is non-singular, $c_i \geq 0$, and $-0.4 \leq d_i \leq 1$, for $i = 1, 2$.[9] Thus, the confidence intervals control coverage over a wide range of persistence patterns including processes less persistent than $I(0)$, as persistent as $I(1)$, local-to-unity autoregressions, and where different linear combinations of x_t and y_t may have markedly different persistence. Even though our theoretical development would allow for values of d_i up to 1.5, we consider values of $d_i > 1$ reasonably rare in empirical analysis of economic time series. In order to obtain more informative inference, we therefore restrict the parameter space to $-0.4 \leq d_i \leq 1$, for $i = 1, 2$.

The confidence sets we construct require three distributions over θ: the weighting function W for computing the average length in the objective (16), the Bayes prior associated with the Bayes credible sets that serve as subsets for the confidence sets (Müller and Norets (2016)), and the least favorable distribution for θ that enforces the coverage constraint. The latter is endogenous to the program (16)–(17) and is approximated using numerical methods similar to those discussed in Elliott, Müller, and Watson (2015), with details provided in the Supplemental Material. In our baseline analysis, we use the same distribution for W and the Bayes prior. Specifically, this distribution is based on the bivariate $I(d)$ model (so that $c_1 = c_2 = 0$, $B = 0$) with d_1 and d_2 independently distributed $U(-0.4, 1.0)$. Because of the invariance/equivariance restrictions, the scale of the matrix A is irrelevant and we set $A = R(\lambda_1)G(s)R(\lambda_2)$, where $R(\lambda)$ is a rotation matrix indexed by the angle λ, with λ_1 and λ_2 independently distributed $U[0, \pi]$. The relative eigenvalues of A are determined by the diagonal matrix $G(s)$, with $G_{11}/G_{22} = 15^s$ with s independently distributed $U[0, 1]$. We investigate the robustness of this choice in Section 4.3 below.

4.2.3. *Empirical Results for GDP, Consumption, and Interest Rates*

Table II shows estimates for $(\rho_T, \beta_T, \sigma_{y|x,T})$ and confidence sets using the (A, B, c, d) model. The estimated value of $(\rho_T, \beta_T, \sigma_{y|x,T})$ is the median of the posterior using the $I(d)$-model prior, and the table also shows Bayes credible sets for this prior for comparison with the frequentist confidence intervals. For GDP and consumption, the (A, B, c, d) results look much like the results obtained for the $I(0)$ model shown in Table I. For most entries, the Bayes credible sets are slightly larger than the $I(0)$ sets, presumably reflecting the possibility of persistence greater than $I(0)$. The frequentist confidence intervals often coincide with Bayes intervals, but occasionally are somewhat wider. The results indicate that GDP and consumption are highly correlated in the long run (the 90% confidence set is $0.71 \leq \rho \leq 0.97$) and the long-run regression coefficient of consumption onto GDP

[9]See the Appendix for additional details.

TABLE II

LONG-RUN COVARIABILITY ESTIMATES, CONFIDENCE INTERVALS AND CREDIBLE SETS USING THE (A, B, C, D) MODELS: PERIODS LONGER THAN 11 YEARS

		ρ	β	$\sigma_{y\|x}$
		a. GDP and consumption		
(A, B, c, d)	Estimate	0.91	0.77	0.41
	67% CI	0.83, 0.96	0.66, 0.87	0.33, 0.53
	90% CI	0.71, 0.97	0.48, 0.96	0.29, 0.66
	67% Bayes CS	0.83, 0.96	0.66, 0.87	0.33, 0.53
	90% Bayes CS	0.71, 0.97	0.58, 0.96	0.29, 0.66
		b. Short- and long-term interest rates		
(A, B, c, d)	Estimate	0.96	0.95	0.63
	67% CI	0.92, 0.98	0.87, 1.07	0.49, 0.97
	90% CI	0.89, 0.99	0.76, 1.16	0.42, 1.27
	67% Bayes CS	0.92, 0.98	0.87, 1.03	0.49, 0.82
	90% Bayes CS	0.89, 0.99	0.81, 1.09	0.42, 1.02

Notes: Periods longer than 11 years correspond to $q = 12$ for panel (a) and $q = 11$ for panel (b). The rows labeled "Estimate" are the posterior median based on the $I(d)$ model. "CI" denotes confidence interval, which is calculated as described in the text. "Bayes CS" are Bayes equal-tailed credible sets based on the posterior from the $I(d)$ model.

is large, but less than unity (the 90% confidence set is $0.48 \leq \beta \leq 0.95$). The results for interest rates indicate that long-run movements in short and long rates are highly correlated, and that these data are consistent with a unit long-run response of long rates to short rates.

4.3. *Asymptotic Power and Size*

In this subsection, we investigate some aspects of power and size that govern the average length and coverage of the confidence intervals. For power, we investigate the choice of the weighting function W and the restriction of our confidence sets to be supersets of $1-\alpha$ Bayes credible sets. For size, we investigate the non-coverage probability of confidence intervals constructed for misspecified models. To keep the discussion concise, we focus exclusively on tests of $H_0 : \rho = 0$ of level $\alpha = 10\%$ using $q = 12$ in this subsection.

We first consider variations in weighting functions W and the cost of imposing the Bayes superset restriction. We investigate three weighting functions: the baseline weighting function $W = W_{\text{base}}$ described above, the weighting function $W_{I(0)}$ that is equivalent to W_{base} except that $d_1 = d_2 = 0$, and $W_{I(1)}$ obtained by setting $d_1 = d_2 = 1$. For each W, we compute the weighted average power maximizing test, both with and without the restriction that the implied confidence set is a superset of the $1-\alpha$ Bayes credible set under the prior W_{base}. This results in six tests. For each test, we then compute its weighted average power, under each of the three weighting functions. In other words, we compute the power of the six tests of $H_0 : \rho = 0$ against the alternative that the data were generated by θ randomly drawn from W, for each of the three W. Table III shows the results. Consider first tests that do not impose the Bayes superset restriction. By construction, the test that is constructed to maximize weighted average power against a given W has the highest weighted average power against that W among all level-α tests; these power-envelope values are shown in the diagonal entries in the table. The table indicates that the optimal test for $W_{I(0)}$ has substantially less power under $W_{I(1)}$ than this envelope, and vice versa. The test constructed under W_{base}, in contrast, is essentially on the envelope under $W_{I(0)}$,

TABLE III

ASYMPTOTIC WEIGHTED AVERAGE POWER OF 10% LEVEL EFFICIENT TEST OF H_0: $\rho = 0$ IN (A, B, c, d) MODEL ($q = 12$)

	WAP efficient test for					
	$W_{I(0)}$	$W_{I(1)}$	W_{base}	$W_{I(0)}$	$W_{I(1)}$	W_{base}
WAP computed for	without Bayes superset constraint			with Bayes superset constraint		
$W_{I(0)}$	0.69	0.34	0.69	0.64	0.34	0.64
$W_{I(1)}$	0.38	0.66	0.60	0.38	0.65	0.60
W_{base}	0.57	0.41	0.61	0.55	0.41	0.58

Notes: The entries are the asymptotic weighted average power over alternatives shown in rows using WAP efficient tests for alternatives given in columns.

and loses only about 6 percentage points under $W_{I(1)}$. Turning to the comparison with tests that impose the Bayes superset restriction, the table suggests that its cost in terms of weighted average power is fairly small, especially under W_{base}.

We now turn to studying the size of tests in misspecified models. Results are shown in Table IV. The rows indicate the model used to generate the data, the columns show the model used to construct the test, and the entries are null rejection probabilities, maximized over the parameters used to generate the data under the constraint of $\rho = 0$. We consider five tests that all maximize weighted average power against W_{base}, but that control size only in restricted versions of the (A, B, c, d) model by construction. None impose the Bayes superset constraint. The five models used to construct the tests are the $I(0)$ model ($S_z(\omega) \propto BB'$), the $I(1)$ model ($S_z(\omega) = \omega^{-2}AA'$), a bivariate "local level" model that includes $I(0)$ and $I(1)$ components ($S_z(\omega) \propto \omega^{-2}AA' + BB'$), the fractional $I(d)$ model ($S_z(\omega) \propto ADA'$, D diagonal with $D_{jj} = \omega^{-2d_j}$), and the general (A, B, c, d) model with $S_z(\omega)$ given by (11), where for the last two, $-0.4 \leq d_1, d_2 \leq 1.0$, as in our baseline. We then compute the size of these tests under eight models. The first five are just the same as were used in the construction of the tests. The other three models are

TABLE IV

ASYMPTOTIC SIZE OF 10% LEVEL EFFICIENT TEST OF H_0: $\rho = 0$ ($q = 12$)

	Efficient test constructed for				
Data generated by	$I(0)$	$I(1)$	$I(0) + I(1)$	$I(d)$	(A, B, c, d)
			(A, B, c, d) models		
$I(0)$	0.10	1.00	0.10	0.09	0.09
$I(1)$	0.99	0.10	0.10	0.10	0.10
$I(0) + I(1)$	0.99	0.99	0.10	0.12	0.10
$I(d)$ with $-0.4 \leq d \leq 1.0$	0.99	1.00	0.31	0.10	0.10
(A, B, c, d)	0.99	1.00	0.31	0.13	0.10
			non-(A, B, c, d) models		
Local cointegration	0.84	1.00	0.15	0.09	0.10
Long-lag $I(0)$	0.30	1.00	0.13	0.09	0.10
$I(d)$ with $-0.4 \leq d \leq 1.4$	1.00	1.00	0.46	0.42	0.47

Notes: Entries in the table are maximal null rejection frequencies for stochastic processes shown in rows for efficient tests shown in columns. The stochastic processes in the top panel are special cases of the (A, B, c, d) model with the parametric restrictions discussed in the paper. The stochastic processes in the bottom panel are not included in the paper's parameterization of the (A, B, c, d) model.

the local cointegration model (13), the long-lag $I(0)$ model (15), and the fractional model with $-0.4 \leq d_1, d_2 \leq 1.4$. The table therefore shows sizes computed from 40 experiments composed of five different tests and data generated from eight different stochastic processes.

The top panel of the table shows results for data generated by each of the five models used to construct the tests, so the diagonal entries of the table are equal to 0.10 by construction. The off-diagonal entries larger than 0.10 indicate size distortions. For example, the 10-percent level $I(0)$ test mistakenly rejects for 99 percent of the draws when the data are generated by other models. The $I(1)$ test has similarly large size distortions when the data are not generated by the $I(1)$ model. These results mirror previous findings of the fragility of inference based on the assumption of exact $I(0)$ or $I(1)$ persistence patterns (e.g., den Haan and Levin (1997) for HAC inference in $I(0)$ models and Elliott (1998) for inference in cointegrated models). The $I(0) + I(1)$ model encompasses both the $I(0)$ and $I(1)$ models, so the associated test has good size control for these models, but has size equal to 31% in the $I(d)$ and (A, B, c, d) models. The $I(d)$ model encompasses the $I(0)$ and $I(1)$ models, and so controls size there by construction. It does not encompass the $I(0) + I(1)$ or (A, B, c, d) models, but exhibits only a relatively small size distortion in these cases.[10] Since the models in the top panel are special cases of our baseline (A, B, c, d) model, the corresponding entries in the fifth column cannot be larger than 0.10 by construction.

The bottom panel of the table shows sizes for data generated by data outside our parameterization of the (A, B, c, d) model. The (A, B, c, d) model based inference seems robust to the long leads and lag patterns induced by the local cointegration model, and the $I(0)$ long lag model. Even though the Σ matrices induced by these models are quite distinct from those induced by the (A, B, c, d) model, this misspecification does not induce substantial overrejections. In contrast, allowing more persistence in the form of fractional stochastic trends with $d \geq 1.0$ can induce severe overrejections. Apparently, for purposes of inference about Ω, it is essential to allow for the correct persistence pattern of the marginals of z_t, while misspecifications of intertemporal dependence seem to play a lesser role.

5. EMPIRICAL ANALYSIS

The last section showed results for the long-run covariation between GDP and consumption and between short- and long-term nominal interest rates. In this section, we use the same methods to investigate other important long-run correlations. We focus on two questions: first, how much information does the sample contain about the long-run covariability, and second, what are the values of the long-run covariability parameters. A knee-jerk reaction to investigating long-run propositions in economics using, say, 68-year spans of data is that little can be learned, particularly so using analysis that is robust to a wide range of persistence patterns. In this case, even efficient methods for extracting relevant information from the data will yield confidence intervals that are so wide that they rule out few plausible parameter values. We find this to be true for some of the long-run relationships investigated below. But, as we have seen from the consumption-income and interest rate data, confidence intervals about long-run parameters can be narrow and informative, and this holds for several of the relationships that we now investigate.

[10] Müller and Watson (2016) showed that the $I(d)$ model yields long-run prediction sets with significant undercoverage when data are generated by a univariate analogue of the (A, B, c, d) model, however.

Throughout the first two subsections, we focus on periods longer than 11 years. For data available over the entire post-WWII period, this entails setting $q = 12$. For shorter sample periods, smaller values of q are used, and these values are noted in context. The last subsection investigates the robustness of the empirical conclusions to focusing on periods longer than 20 years (so that $q = 6$ for data available over the entire post-WWII period).

5.1. *Balanced Growth Correlations*

In the standard one-sector growth model, variations in per-capita GDP, consumption, investment, and in real wages arise from variations in total factor productivity (TFP). Balanced growth means that the consumption-to-income ratio, the investment-to-income ratio, and labor's share of total income are constant over the long run. This implies perfect pairwise long-run correlations between the logarithms of income, consumption, investment, labor compensation, and TFP. In this model, the long-run regression of the logarithm of consumption onto the logarithm of income has a unit coefficient, as do the same regressions with consumption replaced by investment or labor income. A long-run one-percentage-point increase in TFP leads to a long-run increase of $1/(1-\alpha)$ percentage points in the other variables, where $(1-\alpha)$ is labor's share of income. Of course, these implications involve the evolution of the variables over the untestable infinite long run. That said, empirical analysis can determine how well these implications stand up as approximations to below business cycle frequency variation in data spanning the post-WWII period. We use data for the United States and the methods discussed above to investigate these long-run balance growth propositions. The Supplemental Material contains a description of the data that are used.

Figure 4 plots the long-run projections of the growth rates of GDP, consumption, investment, labor income, and TFP. (The long-run projections for consumption and GDP were shown previously in Figure 1(b).) The figure indicates substantial long-run covariability over the post-WWII period, but less so for investment than the other variables. Table V summarizes the results on the long-run correlations. The values above the main

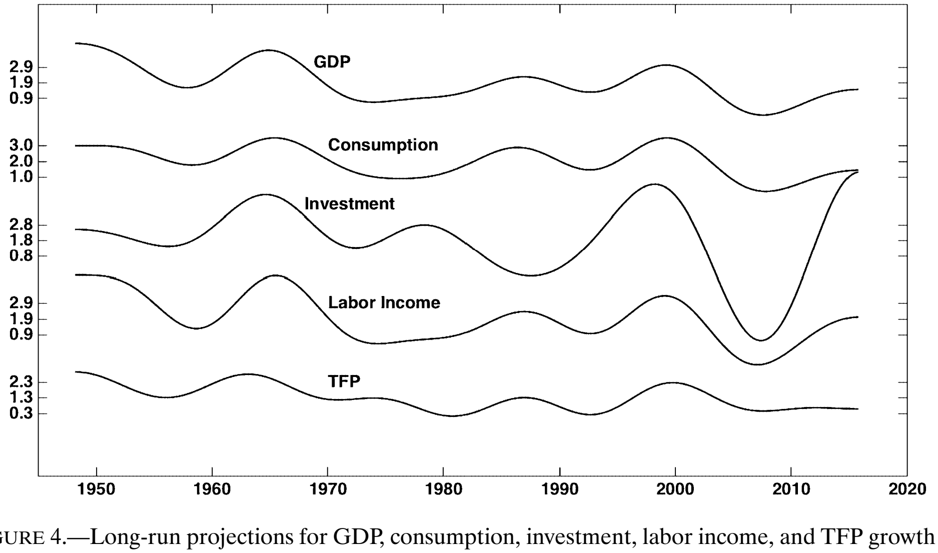

FIGURE 4.—Long-run projections for GDP, consumption, investment, labor income, and TFP growth rates: periods longer than 11 years.

TABLE V

LONG-RUN CORRELATIONS OF GDP, CONSUMPTION, INVESTMENT, LABOR COMPENSATION, AND TFP: PERIODS LONGER THAN 11 YEARS

	GDP	Cons.	Inv.	$w \times n$	TFP
GDP		0.91 (0.83, 0.96)	0.53 (0.29, 0.72)	0.98 (0.96, 0.99)	0.78 (0.64, 0.89)
Cons.	(0.71, 0.97)		0.53 (0.30, 0.72)	0.90 (0.83, 0.96)	0.70 (0.49, 0.82)
Inv.	(0.02, 0.81)	(0.02, 0.81)		0.57 (0.34, 0.74)	0.38 (0.05, 0.60)
$w \times n$	(0.94, 0.99)	(0.55, 0.97)	(0.06, 0.82)		0.71 (0.53, 0.84)
TFP	(0.46, 0.95)	(0.29, 0.91)	(−0.08, 0.72)	(0.36, 0.93)	

Notes: All variables are measured in growth rates. The entries above the diagonal show the median of the posterior distribution followed by the 67% confidence interval. The entries below the diagonal show the 90% confidence interval.

diagonal show point estimates constructed as the posterior median using the $I(d)$ model with prior discussed above, together with 67% confidence intervals (shown in parentheses) using the general (A, B, c, d) model. The values below the main diagonal are the corresponding 90% confidence intervals using the (A, B, c, d) model. Table VI reports results from selected long-run regressions.

As reported in the previous section, the long-run correlation between GDP and consumption is large. Labor income and GDP are highly correlated with a tightly concentrated 90% confidence interval of 0.94 to 0.99. The estimated long-run correlation of TFP and GDP is also high, although the correlation of TFP and the other variables appears to be somewhat lower. Investment and GDP are less highly correlated; the *upper* bound of the 90% confidence interval is only 0.81 and the lower bound is close to zero.

Table VI shows results from long-run regressions of the growth rates of consumption, investment, and labor income onto the growth rate of GDP, and the corresponding regression of GDP onto TFP. Labor compensation appears to vary more than one-for-one with GDP and (as reported above) consumption less than one-for-one. The long-run investment-GDP regression coefficient is imprecisely estimated. Disaggregating consumption into nondurables, durables, and services, suggests that durable consumption

TABLE VI

SELECTED LONG-RUN REGRESSIONS INVOLVING GDP, CONSUMPTION, INVESTMENT, LABOR COMPENSATION, AND TFP: PERIODS LONGER THAN 11 YEARS

		β			
Y	X	$\hat{\beta}$	67% CI	90% CI	$\hat{\sigma}_{y\mid x}$
Consumption	GDP	0.77	0.66, 0.87	0.48, 0.96	0.41
Investment	GDP	1.23	0.68, 1.77	0.12, 2.24	2.19
Labor comp. ($w \times n$)	GDP	1.29	1.20, 1.36	1.14, 1.42	0.32
GDP	TFP	1.22	0.94, 1.49	0.72, 1.72	0.73
Cons. (Nondurable)	GDP	0.36	0.12, 0.59	−0.08, 0.76	0.89
Cons. (Services)	GDP	0.83	0.67, 0.99	0.54, 1.25	0.61
Cons. (Durables)	GDP	1.85	1.47, 2.26	1.19, 2.59	1.52
Inv. (Nonresidential)	GDP	0.97	0.39, 1.51	−0.05, 1.93	2.18
Inv. (Residential)	GDP	2.19	0.81, 3.57	−0.23, 4.69	5.64
Inv. (Equipment)	GDP	0.87	0.14, 1.55	−0.41, 2.12	2.77

Notes: All variables are measured in growth rates, in percentage points at an annual rate. The entries were constructed from the long-run regression of the variable labeled Y onto the variable labeled X.

responds more to long-run variations in GDP than do services and nondurables. These long-run regression results are reminiscent of results using business cycle covariability, and in Section 5.3 we investigate their robustness to the periods longer than 20 years.

In summary, what has the 68-year post-WWII sample been able to say about the balanced-growth implications of the simple growth model? First, that several of the variables are highly correlated over the long run, defined as periods between 11 and 136 years, and second, that the long-run regression coefficient on GDP is different from unity for some variables (consumption and labor income). There is less information about the long-run covariability of investment with the other variables, although even here there are things to learn, such as the long-run correlation of investment and GDP is unlikely to be much larger than 0.8. Section 5.3 shows that similar results obtain using only periods longer than 20 years.

5.2. *Other Long-Run Relations*

Figure 5 and Table VII summarize long-run covariation results for an additional dozen pair of variables, using post-WWII U.S. data. (See the Supplemental Material for description and sources of the data.) We discuss each in turn.

CPI and PCE inflation. We begin with two widely-used measures of inflation, the first based on the consumer price index (CPI) and the second based on the price deflator for personal consumption expenditures (PCE). The Boskin Commission Report and related research (Boskin et al. (1996), Gordon (2006)) highlights important methodological and quantitative differences in these two measures of inflation. For example, the CPI is a Laspeyres index, while the PCE deflator uses chain weighting, and this leads to greater substitution bias in the CPI. Differences in these inflation measures may change over time both because of the variance of relative prices (which affects substitution bias) and because measurement methods for both price indices evolved over the sample period.

Panel (a) of Figure 5 presents two plots; the first shows a time series plot of the long-run projections for PCE and CPI inflation, and the second shows the corresponding scatterplot of the projection coefficients, where the scatterplot symbols are the periods (in years) associated with the coefficients. For instance, the outlier "68.8" corresponds to the large negative coefficient on the first cosine function $\cos(\pi(t-1/2)/T)$, which has a U-shape, and both inflation rates have a pronounced inverted U-shape in the sample. Long-run movements in PCE and CPI inflation track each other closely and the 90% confidence interval shown in Table VII suggests that the long correlation is greater than 0.95. The long-run regression of CPI inflation on PCE inflation yields an estimated slope coefficient that is 1.13 (90% confidence interval: $0.98 \leq \beta \leq 1.24$), suggesting a larger bias in the CPI during periods of high trend inflation.

Long-run Fisher correlation and the real term structure. The next two entries in the figure and table show the long-run covariation of inflation and short- and long-term nominal interest rates. (As above, long-term rates are for 10-year U.S. Treasury bonds available only since 1953, and the analysis with this rate uses $q = 11$.) The well-known Fisher relation (Fisher (1930)) decomposes nominal rates into an inflation and real interest rate component, making it interesting to gauge how much of the long-run variation in nominal rates can be explained by long-run variation in inflation. The long-run correlation of nominal interest rates and inflation is estimated to be approximately 0.5, although the confidence intervals indicate substantial uncertainty. A unit long-run regression coefficient of nomi-

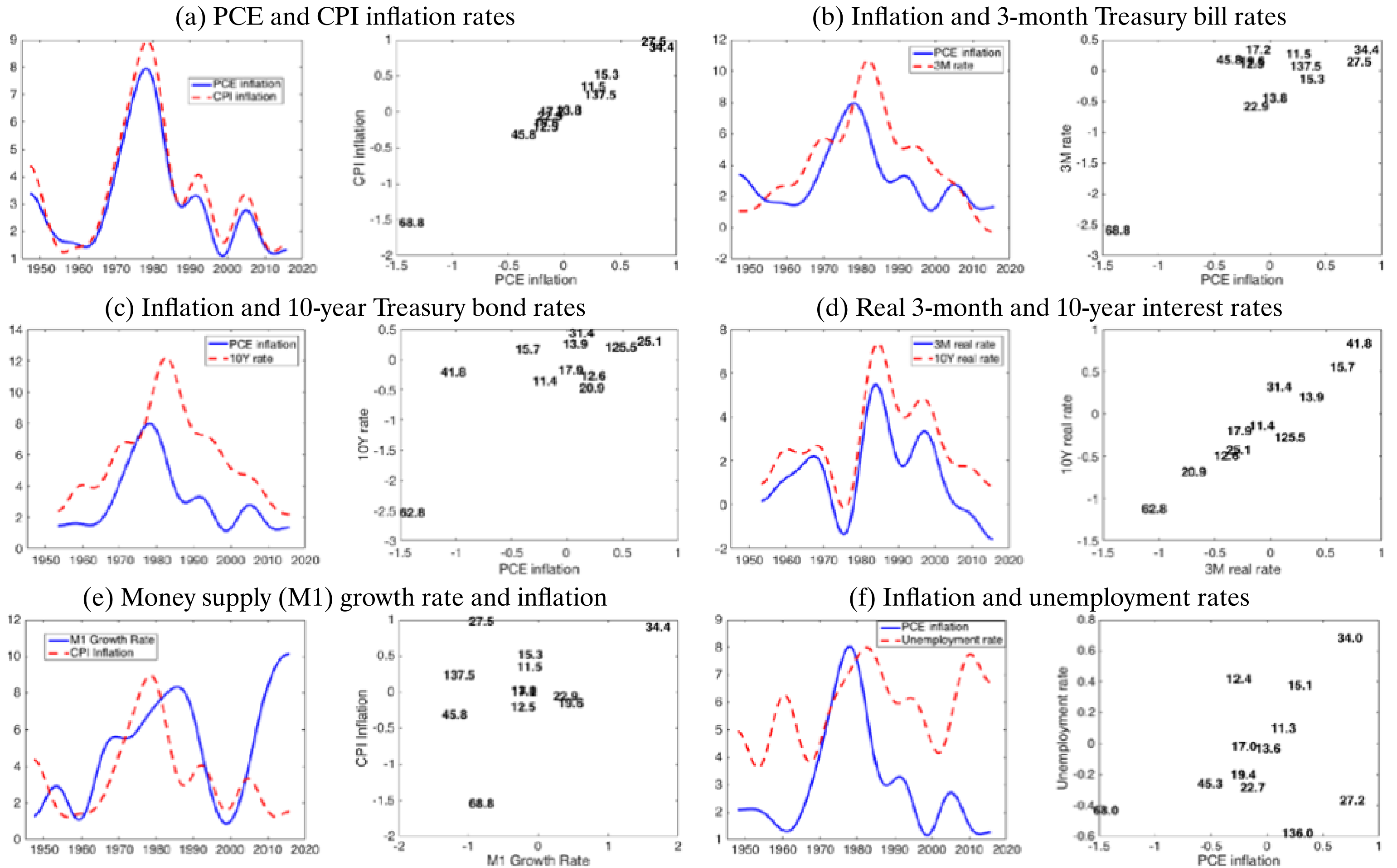

FIGURE 5.—Long-run projections and projection coefficients: periods longer than 11 years. *Notes*: The first plot in each panel shows the long-run projections of the time series. The second plot is a scatterplot of the long-run projection coefficients where the plot symbols indicate the period of the associated cosine function.

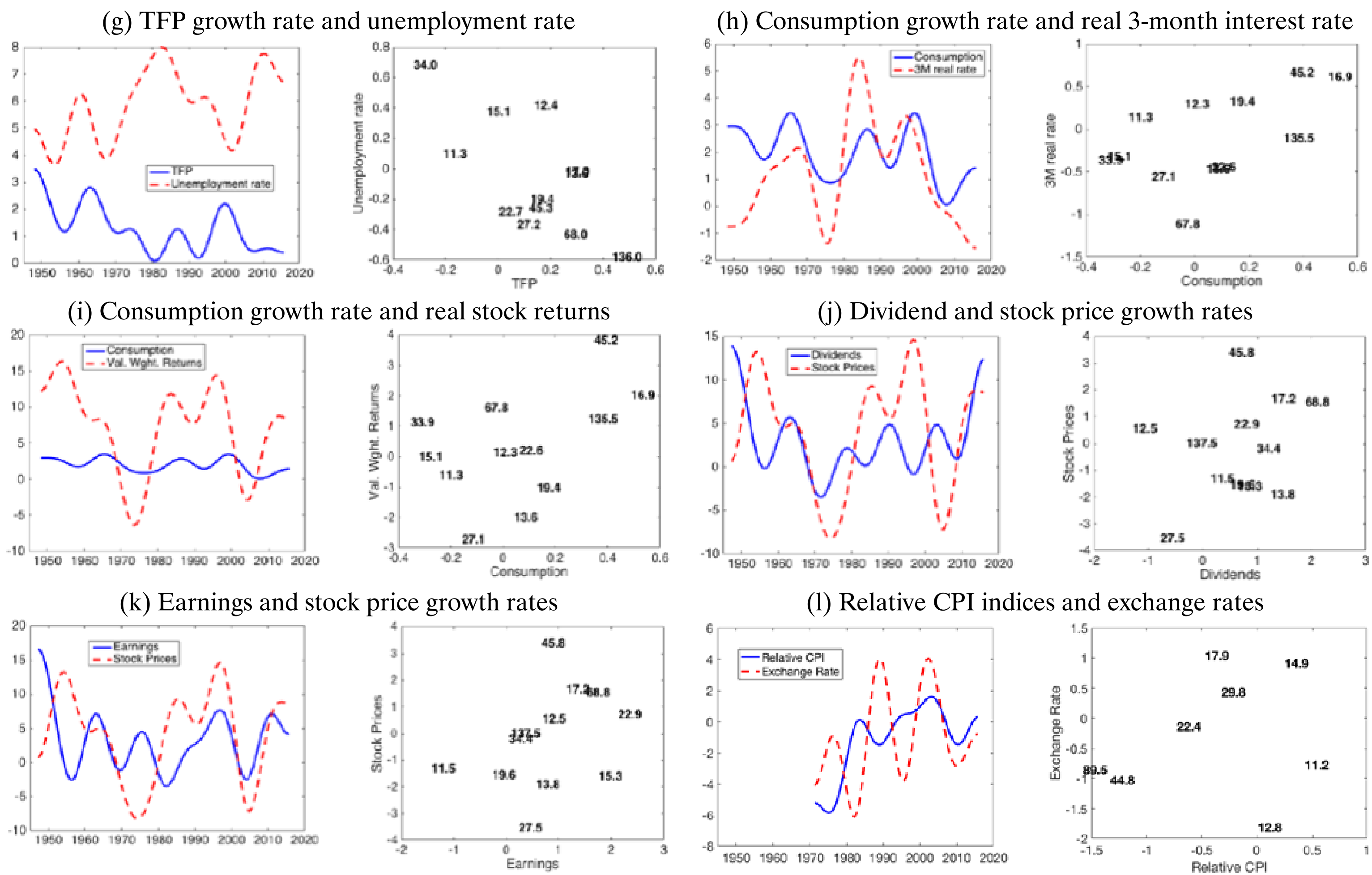

FIGURE 5.—*Continued*.

TABLE VII

LONG-RUN COVARIATION MEASURES FOR SELECTED VARIABLES: PERIODS LONGER THAN 11 YEARS

			ρ			β			
Y	X	Trans.	$\hat{\rho}$	67% CI	90% CI	$\hat{\beta}$	67% CI	90% CI	$\hat{\sigma}_{y\|x}$
CPI Infl.	PCE Infl.	L, L	0.98	0.96, 0.99	0.95, 0.99	1.13	1.07, 1.19	0.98, 1.24	0.36
3M rates	PCE Infl.	L, L	0.47	0.21, 0.80	−0.00, 0.91	0.74	0.34, 1.51	−0.02, 1.89	2.24
10Y rates	PCE Infl.	L, L	0.48	0.21, 0.85	−0.01, 0.92	0.71	0.33, 1.46	0.02, 1.82	2.19
10Y real rates	3M real rates	L, L	0.96	0.90, 0.97	0.75, 0.98	0.97	0.87, 1.11	0.67, 1.31	0.62
CPI Inflation	Money Supply	L, G	0.12	−0.17, 0.55	−0.55, 0.75	0.12	−0.15, 0.50	−0.55, 0.91	2.50
Un. Rate	PCE Infl.	L, L	0.26	−0.03, 0.55	−0.27, 0.80	0.21	−0.04, 0.45	−0.24, 0.83	1.47
Un. Rate	TFP	L, G	−0.65	−0.75, −0.34	−0.91, −0.13	−0.99	−1.38, −0.61	−1.65, −0.27	1.07
3M real rates	Consumption	L, G	0.42	0.08, 0.60	−0.06, 0.75	0.92	0.34, 1.46	−0.35, 2.59	1.84
Stock returns	Consumption	L, G	0.40	0.07, 0.60	−0.08, 0.80	3.00	0.99, 4.99	−0.47, 8.40	6.21
Stock prices	Dividends	G, G	0.20	−0.05, 0.43	−0.30, 0.75	0.45	−0.16, 1.10	−0.61, 1.76	7.28
Stock prices	Earnings	G, G	0.21	−0.04, 0.42	−0.27, 0.57	0.38	−0.14, 0.93	−0.51, 1.35	7.18
Exchange rates	Rel. price ind.	G, G	0.14	−0.13, 0.50	−0.41, 0.75	0.24	−0.37, 0.81	−0.99, 1.26	3.32

Notes: The column labeled "Trans." indicates the transformation applied to the data with "L" denoting level (no transformation) and "G" denoting growth rate (in percentage points at annual rate using scaled first-differences of logarithms). Thus, the levels of inflation, interest rates, and the unemployment were used, and other variables were transformed to growth rates. The long-run regressions were computed from the regression of the variable labeled Y onto the variable labeled X.

nal rates onto inflation is consistent with data, but the confidence intervals are wide.[11] The next entry in the figure and table shows the long-run covariation in short- and long-term real interest rates (constructed as nominal rates minus the PCE inflation rate). Like their nominal counterparts, short- and long-term real rates are highly correlated over the long run (90% confidence interval: $0.75 \leq \rho \leq 0.98$) with a near unit regression coefficient of long rates onto short rates.

Money growth and inflation. An important implication of the quantity theory of money is the close relationship between money growth and price inflation over the long run. Lucas (1980) investigated this implication using time series data on money (M1) growth and (CPI) inflation for the United States over 1953–1977. After using an exponential smoothing filter to isolate long-run variation in the series, he found a nearly one-for-one relationship between money growth and inflation. The next entry in the figure and table examines this long-run relation using the same M1 and CPI data used by Lucas, but over the longer sample period, 1947–2015. Figure 5 shows the close long-run relationship between money growth and inflation from the mid-1950s through late-1970s documented by Lucas, but shows a much weaker (or non-existent) relationship in the post-1980 sample period, and over the entire sample period the estimated long-run correlation is only 0.12 with a 67% confidence interval that ranges from −0.17 to 0.55.

Long-run Phillips correlation. The next entry summarizes the long-run correlation between the unemployment and inflation. The estimated long-run Phillips correlation and slope coefficient are positive, but $\rho = \beta = 0$ is contained in the 67% confidence interval. That said, the confidence intervals are wide so that, like the Fisher correlation, the data are not very informative about the long-run Phillips correlation.

Unemployment and productivity. Panel (g) of the figure investigates the long-run covariation of the unemployment rate and productivity growth. The large negative in-sample long-run correlation evident in the figure has been noted previously (e.g., Staiger, Stock, and Watson (2001)); the confidence intervals reported in Table VII show that the correlation is unlikely to be spurious. There is a statistically significant negative long-run relationship between the variables. A long-run one-percentage-point increase in the rate of growth of productivity is associated with an estimated one-percentage-point decline in the long-run unemployment rate. We are unaware of an economically compelling theoretical explanation for the large negative correlation.

Real returns and consumption growth. Consumption-based asset pricing models (e.g., Lucas (1978)) draw a connection between consumption growth (as an indicator of the intertemporal marginal rate of substitution) and asset returns. A large literature has followed Hansen and Singleton (1982, 1983) investigating this relationship, with varying degrees of success. Rose (1988) discussed the puzzling long-run implications of the model when consumption growth follows an $I(0)$ process and real returns are $I(1)$ (also see Neely and Rapach (2008)), but moving beyond the $I(0)$ and $I(1)$ models, it is clear from the empirical results reported above that both consumption growth and real interest rates exhibit substantial long-run variability. The next two entries in the figure and table investigate the long-run covariability between consumption growth and and real returns, first using real returns on short-term Treasury bills and then using real returns on stocks. Both suggest a moderate positive long-run correlation between real returns and consumptions

[11]These estimates measure the long-run Fisher "correlation," not the long-run Fisher "effect". The long-run Fisher correlation considers variation from all sources, while the Fisher effect instead considers variation associated with exogenous long-run nominal shocks (e.g., Fisher and Seater (1993), King and Watson (1997)). A similar distinction holds for the Phillips correlation and the Phillips curve (see King and Watson (1994)).

growth rates, although the confidence interval is wide (90% confidence range from just below zero to 0.80).

Stock prices, dividends, and earnings. Present value models of stock prices imply a close relationship between long-run values of prices, dividends, and earnings (e.g., Campbell and Shiller (1987)). An implication of this long-run relation in a cointegration framework is that dividends, earnings, and stock prices share a common $I(1)$ trend, so that their growth rates are perfectly correlated in the long run and the dividend-price or price-earning ratio is useful for predicting future stock returns. This latter implication has been widely investigated (see Campbell and Yogo (2006) for analysis and references). The next two entries show the long-run correlation of stock prices with dividends and with earnings.[12] While there is considerable uncertainty about the value of the long-run correlation between stock prices and dividends or earnings, the data suggest that the correlation is not strong. For example, values above $\rho = 0.43$ are ruled out by the 67% confidence set and values above 0.75 are ruled out by the 90% sets.

Long-run PPP. The final entry shows results on the long-run correlation between nominal exchange rates (here the U.S. dollar/British pound exchange rate from 1971 to 2015) and the ratio of nominal prices (here the ratio of CPI indices for the two countries). With the shortened sample, $q = 8$ captures periods longer than 11 years. Long-run PPP implies that the nominal exchange rate should move proportionally with the price ratio over long time spans, so the long-run growth rates of the nominal exchange rate and price ratios should be perfectly correlated. A large literature has tested this proposition in a unit-root and cointegration framework and obtained mixed conclusions. (See Rogoff (1996) and Taylor and Taylor (2004) for discussion and references.) From the final row of Table VII, the growth rates of nominal exchange rates and relative nominal prices are only weakly correlated over periods longer than 11 years, although the confidence interval is very wide.

5.3. *Longer-Run Periods*

The empirical results shown above relied on projections capturing periods longer than 11 years. While 11 years is longer than typical business cycles, it does incorporate periods corresponding to what some researchers refer to as the "medium run" (Blanchard (1997), Comin and Gertler (2006)). This subsection investigates the robustness of the empirical results to restricting the data to periods longer than 20 years. For the series available over the entire 68-year post-WWII sample period, this entails only the first $q = 6$ cosine transforms ($2T/6 \approx 23$ years); for the 1971–2015 sample for exchange rates and relative prices, this entails using $q = 4$ ($2T/4 \approx 22$ years). Results are summarized in Table VIII.

The first four sets of entries in Table VIII involve consumption, investment, labor compensation, TFP, and GDP. These results are remarkably similar to the results shown in earlier tables. Covariability over periods longer than 20 years is similar to covariability of periods longer than 11 years, although the reduction in information moving from $q = 12$ to $q = 6$ leads to somewhat wider confidence intervals. The other results summarized in Table VII show much of the same stability, but there are some notable differences. For example, the point estimates suggest a somewhat larger Fisher correlation over longer periods than over shorter periods, and the same holds for stock prices and dividends. In both cases, however, the confidence intervals remain wide. Exchange rates and relative nominal prices also appear more highly correlated using these longer periods. And, the

[12]The data are for the S&P, and are updated versions of the data used in Shiller (2000) available on Robert Shiller's webpage.

TABLE VIII

LONG-RUN COVARIATION MEASURES FOR SELECTED VARIABLES: PERIODS LONGER THAN 20 YEARS

		ρ			β			
Y	X	$\hat{\rho}$	67% CI	90% CI	$\hat{\beta}$	67% CI	90% CI	$\hat{\sigma}_{y\|x}$
Cons.	GDP	0.89	0.71, 0.95	0.50, 0.97	0.73	0.56, 0.89	0.36, 1.07	0.32
Inv.	GDP	0.56	0.21, 0.79	−0.04, 0.90	1.08	0.50, 1.68	−0.15, 2.29	1.16
$w \times n$	GDP	0.97	0.92, 0.98	0.83, 0.99	1.26	1.15, 1.38	1.04, 1.50	0.24
GDP	TFP	0.71	0.41, 0.94	0.08, 0.96	1.15	0.72, 1.54	0.32, 2.02	0.60
10Y nom. rates	3M nom. rates	0.96	0.92, 0.98	0.84, 0.99	0.97	0.85, 1.09	0.74, 1.20	0.61
10Y real rates	3M real rates	0.94	0.86, 0.97	0.70, 0.98	1.02	0.87, 1.18	0.62, 1.41	0.61
CPI infl.	PCE infl.	0.98	0.96, 0.99	0.93, 0.99	1.11	1.05, 1.17	0.97, 1.23	0.28
3M rates	PCE infl.	0.65	0.30, 0.85	0.00, 0.96	1.02	0.56, 1.58	0.10, 2.18	2.05
10Y rates	PCE infl.	0.57	0.23, 0.90	−0.03, 0.95	0.92	0.47, 1.44	−0.05, 2.07	2.11
CPI inflation	Money supply	0.27	−0.08, 0.59	−0.37, 0.85	0.30	−0.07, 0.67	−0.43, 1.04	2.39
Un. rate	PCE infl.	0.27	−0.08, 0.57	−0.36, 0.85	0.23	−0.07, 0.50	−0.39, 0.88	1.34
Un. rate	TFP	−0.89	−0.95, −0.72	−0.97, −0.57	−1.84	−2.49, −1.41	−2.85, −0.99	0.54
3M real rates	Consumption	0.30	−0.05, 0.59	−0.45, 0.85	0.98	−0.10, 2.07	−1.54, 3.43	1.79
Stock returns	Consumption	0.41	0.00, 0.70	−0.23, 0.90	4.32	0.85, 7.76	−2.88, 11.57	5.75
Stock prices	Dividends	0.42	0.02, 0.71	−0.20, 0.85	1.16	0.30, 1.97	−0.48, 2.75	5.58
Stock prices	Earnings	0.23	−0.08, 0.57	−0.37, 0.73	0.64	−0.20, 1.49	−0.95, 2.26	6.12
Exchange rates	Rel. price ind.	0.70	0.27, 0.93	−0.05, 0.96	0.57	0.30, 0.83	−0.02, 1.14	0.99

Notes: This table summarizes the long-run covariance confidence sets for (X, Y) for periods longer than 20 years ($q = 6$ for all entries except exchange rates, which uses $q = 4$). See notes to Table VII for units of the variables.

puzzling negative correlation between the unemployment rate and TFP appears to be stronger when the sample is restricted to periods longer than 20 years.[13]

6. CONCLUDING REMARKS

This paper began by highlighting two fundamental problems for inference about long-run variability and covariability: the first was the paucity of sample inference, and the second was that inference critically depends on the data's long-run persistence. Both play an important role in our analysis. The limited sample information is captured by focusing attention on a small number of low-frequency weighted averages that summarize the data's long-run variability and covariability. In large samples and for a reasonably wide range of persistent processes, these weighted averages are approximately normally distributed. Long-run persistence is important, but as we show, only through its effect on the spectral density in a narrow low-frequency band. Using a flexible parameterization of the spectrum for these low frequencies, small-sample normal inference methods allow us to construct asymptotically efficient confidence intervals for long-run variance and covariance parameters.

This paper has focused on inference about long-run covariability of two time series. Just as with previous frameworks, it is natural to consider a generalization to a higher-dimensional setting. For example, this would allow one to determine whether the significant long-run correlation between the unemployment rate and productivity is robust to including a control for, say, some measure of human capital accumulation.

Many elements of our analysis generalize to n time series in a straightforward manner: The analogous definition of Ω_T is equally natural as a second-moment summary of the covariability of n series, and gives rise to corresponding regression parameters, such as coefficients from an $n-1$-dimensional multiple regression, corresponding residual standard deviations, and population R^2's.[14] Multivariate versions of Ω_T can also be used for long-run instrumental variable regressions. As shown in Müller and Watson (2017), the Central Limit Theorem that reduces the inference question to one about the covariance matrix of a multivariate normal holds for arbitrary fixed n. The (A, B, c, d) model of persistence naturally generalizes to an n-dimensional system.

Having said that, our numerical approach for constructing (approximate) minimal-length confidence sets faces daunting computational challenges in a higher-order system: The quadratic forms that determine the likelihood require $O(n^2q^2)$ floating point operations. Worse still, even for n as small as $n = 3$, the number of parameters in the (A, B, c, d) model is equal to 21. So even after imposing invariance or equivariance, ensuring coverage requires an exhaustive search over a high-dimensional nuisance parameter space.

At the same time, it would seem to be relatively straightforward to determine Bayes credible sets also for larger values of n: Under our asymptotic approximation, the (A, B, c, d) parameters enter the likelihood through the covariance matrix of an $nq \times 1$ multivariate normal, so with some care, modern posterior samplers should be able to reliably determine the posterior for any function of interest. Of course, such an approach

[13]This analysis has focused on long-run covariability associated with the first q cosine transforms, that is, the projections for periods between $2T$ and $2T/q$. More generally, a researcher might be interested in the covariance for periods between, say, $2T/q_1$ and $2T/q_2$ associated with the projections of the data onto Ψ_{q_1} through Ψ_{q_2}. See the Appendix for discussion.

[14]Müller and Watson (2017) provided the details of inference in the $I(0)$ model.

does not guarantee frequentist coverage, and the empirical results will depend on the choice of prior in a non-trivial way. In this regard, our empirical results in the bivariate system show an interesting pattern: Especially at a lower nominal coverage level, for many realizations, there is no need to augment the Bayes credible set computed from the bivariate fractional model. This suggests that the frequentist coverage of the unaltered Bayes intervals may be close to the nominal level, so these Bayes sets would not be too misleading even from a frequentist perspective.[15] While this will be difficult to exhaustively check, this pattern might well generalize also to larger values of n.

REFERENCES

ANDERSON, T. W. (1984): *An Introduction to Multivariate Statistical Analysis* (Second Ed.). New York: John Wiley & Sons Inc. [784]

BAXTER, M., AND R. G. KING (1999): "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, 81 (4), 575–593. [776]

BLANCHARD, O. J. (1997): "The Medium Run," *Brooking Panel on Economic Activity*, 28, 89–158. [800]

BOSKIN, M. J., E. DULBERGER, R. GORDON, Z. GRILICHES, AND D. JORGENSON (1996): "Toward a More Accurate Measure of the Cost of Living," Final Report to the Senate Finance Committee. [795]

BRILLINGER, D. R. (2001): *Time Series: Data Analysis and Theory*. SIAM. [781]

BROCKWELL, P. J., AND R. A. DAVIS (1991): *Time Series: Theory and Methods* (Second Ed.). New York: Springer. [781]

CAMPBELL, J. Y., AND R. J. SHILLER (1987): "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1062–1088. [800]

CAMPBELL, J. Y., AND M. YOGO (2006): "Efficient Tests of Stock Return Predictability," *Journal of Financial Economics*, 81, 27–60. [800]

CHEN, W. W., AND C. M. HURVICH (2003): "Semiparametric Estimation of Multivariate Fractional Cointegration," *Journal of the American Statistical Association*, 98 (463), 629–642. [776]

COMIN, D., AND M. GERTLER (2006): "Medium Term Business Cycles," *American Economic Review*, 96, 523–551. [800]

DEN HAAN, W. J., AND A. T. LEVIN (1997): "A Practitioner's Guide to Robust Covariance Matrix Estimation," in *Handbook of Statistics*, Vol. 15, ed. by G. S. Maddala and C. R. Rao. Amsterdam: Elsevier, 299–342. [792]

ELLIOTT, G. (1998): "The Robustness of Cointegration Methods When Regressors Almost Have Unit Roots," *Econometrica*, 66, 149–158. [787,792]

ELLIOTT, G., U. K. MÜLLER, AND M. W. WATSON (2015): "Nearly Optimal Tests When a Nuisance Parameter Is Present Under the Null Hypothesis," *Econometrica*, 83, 771–811. [788,789]

ENGLE, R. F. (1974): "Band Spectrum Regression," *International Economic Review*, 15, 1–11. [776,781]

FISHER, I. (1930): *The Theory of Interest*. New York: MacMillan. [795]

FISHER, M. E., AND J. D. SEATER (1993): "Long-Run Neutrality and Superneutrality in an ARIMA Framework," *American Economic Review*, 83, 402–415. [799]

GORDON, R. J. (2006): "The Boskin Commission Report: A Retrospective One Decade Later," NBER Working Paper 12311. [795]

GRANGER, C. W. J., AND P. NEWBOLD (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 111–120. [781]

HANSEN, L. P., AND K. J. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50 (5), 1269–1286. [799]

——— (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91 (2), 249–265. [799]

HODRICK, R. J., AND E. C. PRESCOTT (1997): "Postwar US Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29, 1–16. [776]

JANSSON, M., AND M. J. MOREIRA (2006): "Optimal Inference in Regression Models With Nearly Integrated Regressors," *Econometrica*, 74, 681–714. [787]

[15] In fact, a calculation analogous to those in Table IV shows that the 67% Bayes set contains the true value of $\rho = 0$ at least 64% of the time in the bivariate (A, B, c, d) model, and the 95% Bayes set has coverage of 83%.

KING, R. G., AND M. W. WATSON (1994): "The Post-War U.S. Phillips Curve: A Revisionist Econometric History," *Carnegie-Rochester Conference on Public Policy*, 41, 157–219. [799]

——— (1997): "Testing Long-Run Neutrality," *Federal Reserve Bank of Richmond Economic Quarterly*, 83, 69–101. [799]

KLEIN, L. R., AND R. F. KOSOBUD (1961): "Some Econometrics of Growth: Great Ratios of Economics," *Quarterly Journal of Economics*, LXXV, 173–198. [777]

LINDGREN, G. (2013): *Stationary Stochastic Processes: Theory and Applications*. Chapman & Hall/CRC. [785]

LUCAS, R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46 (6), 1429–1445. [799]

——— (1980): "Two Illustrations of the Quantity Theory of Money," *American Economic Review*, 70 (5), 1005–1014. [799]

MARINUCCI, D., AND P. M. ROBINSON (2001): "Semiparametric Fractional Cointegration Analysis," *Journal of Econometrics*, 105 (1), 225–247. [776]

MÜLLER, U. K. (2011): "Efficient Tests Under a Weak Convergence Assumption," *Econometrica*, 79, 395–435. [783]

MÜLLER, U. K., AND A. NORETS (2016): "Credibility of Confidence Sets in Nonstandard Econometric Problems," *Econometrica*, 84, 2183–2213. [788,789]

MÜLLER, U. K., AND M. W. WATSON (2008): "Testing Models of Low-Frequency Variability," *Econometrica*, 76, 979–1016. [776,778]

——— (2016): "Measuring Uncertainty About Long-Run Predictions," *Review of Economic Studies*, 83, 1711–1740. [785,792]

——— (2017): "Low-Frequency Econometrics," in *Advances in Economics: Eleventh World Congress of the Econometric Society*, ed. by B. Honoré and L. Samuelson. Cambridge: Cambridge University Press. [782,783, 802]

——— (2018): "Supplement to 'Long-Run Covariability'," *Econometrica Supplemental Material*, 86, https://doi.org/10.3982/ECTA15047. [777]

NEELY, C. J., AND D. E. RAPACH (2008): "Real Interest Rate Persistence: Evidence and Implications," *Federal Reserve Bank of St. Louis Review*, 90 (6), 609–641. [799]

PHILLIPS, P. C. B. (1991): "Spectral Regression for Cointegrated Time Series," in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, ed. by W. A. Barnett and G. Tauchen. Cambridge: Cambridge University Press, 4. [776]

PRATT, J. W. (1961): "Length of Confidence Intervals," *Journal of the American Statistical Association*, 56, 549–567. [788]

ROBINSON, P. M. (2008): "Multiple Local Whittle Estimation in Stationary Systems," *The Annals of Statistics*, 36 (5), 2508–2530. [776]

ROBINSON, P. M., AND J. HUALDE (2003): "Cointegration in Fractional Systems With Unknown Integration Orders," *Econometrica*, 71 (6), 1727–1766. [776]

ROGOFF, K. (1996): "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 34, 647–668. [800]

ROSE, A. K. (1988): "Is the Real Interest Rate Stable?," *Journal of Finance*, 43 (5), 1095–1112. [799]

SHILLER, R. J. (2000): *Irrational Exuberance*. Princeton: Princeton University Press. [800]

SHIMOTSU, K. (2012): "Exact Local Whittle Estimation of Fractionally Cointegrated Systems," *Journal of Econometrics*, 169 (2), 266–278. [776]

STAIGER, D., J. H. STOCK, AND M. W. WATSON (2001): "Prices, Wages, and the U.S. NAIRU in the 1990s," in *The Roaring Nineties*, ed. by A. B. Krueger and R. Solow. New York: The Russell Sage Foundation, 3–60. [799]

TAYLOR, A. M., AND M. P. TAYLOR (2004): "The Purchasing Power Parity Debate," *Journal of Economic Perspectives*, 18, 135–158. [800]

Co-editor Kenneth D. West handled this manuscript as an invited Fisher–Schultz lecture. The invitation to deliver the Fisher–Schultz lecture is also an invitation to publish a suitable version of the lecture in Econometrica.

Manuscript received 24 January, 2017; final version accepted 19 December, 2017; available online 11 January, 2018.