Measures of Fit for Calibrated Models

Mark W. Watson

Northwestern University and Federal Reserve Bank of Chicago

This paper suggests a new procedure for evaluating the fit of a dynamic structural economic model. The procedure begins by augmenting the variables in the model with just enough stochastic error so that the model can exactly match the second moments of the actual data. Measures of fit for the model can then be constructed on the basis of the size of this error. The procedure is applied to a standard real business cycle model. Over the business cycle frequencies, the model must be augmented with a substantial error to match data for the postwar U.S. economy. Lower bounds on the variance of the error range from 40 percent to 60 percent of the variance in the actual data.

I. Introduction

Economists have long debated appropriate methods for assessing the empirical relevance of economic models. The standard econometric approach can be traced back to Haavelmo (1944), who argued that an economic model should be embedded within a complete probability model and analyzed using statistical methods designed for conducting inference about unknown probability distributions. In the modern literature, this approach is clearly exemplified in work such as that of L. Hansen and Sargent (1980) or McFadden (1981). However, many economic models do not provide a realistic and complete

probability structure for the variables under consideration. To ana-
lyze these models using standard econometric methods, they must
first be augmented with additional random components. Inferences
drawn from these expanded models are meaningful only to the ex-
tent that the additional random components do not mask or change
the salient features of the original economic models.

Another approach, markedly different from the standard econo-
metric approach, has become increasingly popular for evaluating dy-
namic macroeconomic models. This approach is clearly articulated in
the work of Kydland and Prescott (1982) and Prescott (1986). In a
general sense, the approach asks whether data from a real economy
share certain characteristics with data generated by the artificial econ-
omy described by an economic model. There is no claim that the
model explains all the characteristics of the actual data, nor is there
any attempt to augment the model with additional random compo-
nents to more accurately describe the data. On the one hand, the
results from this approach are easier to interpret than the results
from the standard econometric approach since the economic model
is not complicated by additional random elements added solely for
statistical convenience. On the other hand, since the economic model
does not provide a complete probability structure, inference proce-
dures lack statistical foundations and are necessarily ad hoc. For ex-
ample, a researcher may determine that a model fits the data well
because it implies moments for the variables under study that are
"close" to the moments of the actual data, even though the metric
used to determine the distance between the moments is left unspeci-
fied.

This paper is an attempt to put the latter approach on a less ad
hoc foundation by developing goodness-of-fit measures for the class
of dynamic econometric models whose endogenous variables follow
covariance stationary processes. It is not assumed that the model accu-
rately describes data from the actual economy; the economic model
is not a null hypothesis in the statistical sense. Rather, the economic
model is viewed as an approximation to the stochastic processes gen-
erating the actual data, and goodness-of-fit measures are proposed
to measure the quality of this approximation. A standard device—
stochastic error—is used to motivate the goodness-of-fit measures.
These measures answer the question, How much random error
would have to be added to the data generated by the model so that
the autocovariances implied by the model + error match the autoco-
variances of the observed data?

The error represents the degree of abstraction of the model from
the data. Since the error cannot be attributed to a data collection
procedure or to a forecasting procedure, for instance, it is difficult a

priori to say much about its properties. In particular, its covariance with the observed data cannot be specified by a priori reasoning. Rather than make a specific assumption about the error's covariance properties, I construct a representation that minimizes the contribution of the error in the complete model. Thus, in this sense, the error process is chosen to make the model as close to the data as possible.

Many of the ideas in this paper are close to, and were motivated by, ideas in Altug (1989) and Sargent (1989). Altug (1989) showed how a one-shock real business cycle model could be analyzed using standard dynamic econometric methods, after first augmenting each variable in the model with an idiosyncratic error. This produces a restricted version of the dynamic factor analysis or unobserved index models developed by Sargent and Sims (1977) and Geweke (1977). Sargent (1989) discusses two models of measurement error: in the first the measurement error is uncorrelated with the data generated by the model, and in the second the measurement error is uncorrelated with the sample data (see also G. Hansen and Sargent 1988). While similar in spirit, the approach taken in this paper differs from that of Altug and Sargent in two important ways. First, in this paper, the error process is not assumed to be uncorrelated with the model's artificial data or with the actual data. Rather, the correlation properties of the error process are determined by the requirement that the variance of the error is as small as possible. Second, the joint data-error process is introduced to motivate goodness-of-fit measures; it is not introduced to describe a statistical model that can be used to test statistical hypotheses, at least in the standard sense. Rather, the analysis in this paper is similar to the analysis in Campbell and Shiller (1988), Durlauf and Hall (1989), Hansen and Jagannathan (1991), and Cochrane (1992). Each of these papers uses a different approach to judge the goodness of fit of an economic model by calculating a value or an upper bound on the variance of an unobserved "noise" or a "marginal rate of substitution" or a "discount factor" in observed data.

The minimum approximation error representation developed in this paper motivates two sets of statistics that can be used to evaluate the goodness of fit of the economic model. First, like the variance of the error in a regression model, the variance of the approximation error can be used to form an R^2 measure for each variable in the model. This provides an overall measure of fit. Moreover, spectral methods can be used to calculate this R^2 measure for each frequency so that the fit can be calculated over the "business cycle," "growth," or other specific frequency bands. A second set of statistics can be constructed by using the minimum error representation to form fitted values of the variables in the economic model. These fitted values

show how well the model explains specific historical episodes; for example, can a real business cycle model simultaneously explain the growth in the United States during the 1960s and the 1981–82 recession?

The paper is organized as follows. Section II develops the minimum approximation error representation and goodness-of-fit measures. Section III calculates these goodness-of-fit statistics for a standard real business cycle model using postwar U.S. macroeconomic data on output, consumption, investment, and employment. Section IV concludes the paper by providing a brief discussion of some tangential issues that arise from the analysis.

II. Measures of Fit

Consider an economic model that describes the evolution of an $n \times 1$ vector of variables $\mathbf{x}_t$. Assume that the variables have been transformed, say by first-differencing or forming ratios, so that $\mathbf{x}_t$ is covariance stationary. As a notational device, it is useful to introduce the autocovariance generating function (ACGF) of $\mathbf{x}_t$, $A_\mathbf{x}(z)$. This function completely summarizes the unconditional second-moment properties of the process. In what follows, "economic model" and "$A_\mathbf{x}(z)$" will be used interchangeably; that is, the analysis considers only the unconditional second-moment implications of the model. Nonlinearities and variation in conditional second and higher moments are ignored to help keep the problem tractable. The analysis will also ignore the unconditional first moments of $\mathbf{x}_t$; modifying the measures of fit for differences in the means of the variables is straightforward.

The empirical counterparts of $\mathbf{x}_t$ are denoted $\mathbf{y}_t$. These variables differ from $\mathbf{x}_t$ in an important way. The variables making up $\mathbf{x}_t$ correspond to the variables appearing in the theorist's simplification of reality; in a macroeconomic model they are variables such as "output," "money," and the "interest rate." The variables making up $\mathbf{y}_t$ are functions of raw data collected in a real economy; they are variables such as "per capita gross domestic product in the United States in 1987 dollars" or "U.S. M2" or "the yield on 3-month U.S. Treasury bills."

The question of interest is whether the model generates data with characteristics similar to those of the data generated by the real economy. Below, goodness-of-fit measures are proposed to help answer this question. Before I introduce these new measures, it is useful to review standard statistical goodness-of-fit measures to highlight their deficiencies for answering the question at hand.

Standard statistical goodness-of-fit measures use the size of sampling error to judge the coherence of the model with the data. They

are based on the following: First, $A_\mathbf{y}(z)$, the population ACGF for $\mathbf{y}_t$, is unknown but can be estimated from sample data. Discrepancies between the estimator $\hat{A}_\mathbf{y}(z)$ and $A_\mathbf{y}(z)$ arise solely from sampling error in $\hat{A}_\mathbf{y}(z)$, and the likely size of the error can be deduced from the stochastic process that generated the sample. Now, *if* $A_\mathbf{y}(z) = A_\mathbf{x}(z)$, sampling error also accounts for the differences between $\hat{A}_\mathbf{y}(z)$ and $A_\mathbf{x}(z)$. Standard goodness-of-fit measures show how likely it is that $A_\mathbf{y}(z) = A_\mathbf{x}(z)$, on the basis of the probability that differences between $\hat{A}_\mathbf{y}(z)$ and $A_\mathbf{x}(z)$ arise solely from sampling error. If the differences between $\hat{A}_\mathbf{y}(z)$ and $A_\mathbf{x}(z)$ are so large as to be unlikely, standard measures of fit suggest that the model fits the data poorly, and vice versa if the differences between $\hat{A}_\mathbf{y}(z)$ and $A_\mathbf{x}(z)$ are not so large as to be unlikely. The key point is that the differences between $\hat{A}_\mathbf{y}(z)$ and $A_\mathbf{x}(z)$ are judged by how informative the sample is about the population moments of $\mathbf{y}_t$. This is a sensible procedure for judging the coherence of a null hypothesis, $A_\mathbf{y}(z) = A_\mathbf{x}(z)$, with the data. It is arguably less sensible when this null hypothesis is known to be false.

Rather than rely on sampling error, the measures of fit that are developed here are based on the size of the stochastic error required to reconcile the autocovariances of $\mathbf{x}_t$ with those of $\mathbf{y}_t$. In particular, let $\mathbf{u}_t$ denote an $n \times 1$ error vector; then the importance of a difference between $A_\mathbf{x}(z)$ and $A_\mathbf{y}(z)$ will be determined by asking, How much error would have to be added to $\mathbf{x}_t$ so that the autocovariances of $\mathbf{x}_t + \mathbf{u}_t$ are equal to the autocovariances of $\mathbf{y}_t$? If the variance of the required error is large, then the discrepancy between $A_\mathbf{x}(z)$ and $A_\mathbf{y}(z)$ is large, and conversely if the variance of $\mathbf{u}_t$ is small. The vector $\mathbf{u}_t$ is the approximation error in the economic model interpreted as a stochastic process. It captures the second-moment characteristics of the observed data that are not captured by the model. Loosely speaking, it is analogous to the error term in a regression in which the set of regressors is interpreted as the economic model. The economic model might be deemed a good approximation to the data if the error term is small (i.e., the R^2 of the regression is large) and might be deemed a poor approximation if the error term is large (i.e., the R^2 of the regression is small).

To be more precise, assume that $\mathbf{x}_t$ and $\mathbf{y}_t$ are jointly covariance stationary, and define the error $\mathbf{u}_t$ by the equation

$$\mathbf{u}_t = \mathbf{y}_t - \mathbf{x}_t, \tag{1}$$

so that

$$A_\mathbf{u}(z) = A_\mathbf{y}(z) + A_\mathbf{x}(z) - A_\mathbf{xy}(z) - A_\mathbf{yx}(z), \tag{2}$$

where $A_\mathbf{u}(z)$ is the ACGF of $\mathbf{u}_t$, $A_\mathbf{xy}(z)$ is the cross ACGF between $\mathbf{x}_t$ and $\mathbf{y}_t$, and so forth. From the right-hand side of (2), three terms are

needed to calculate $A_u(z)$. The first, $A_y(z)$, can be consistently estimated from sample data; the second, $A_x(z)$, is completely determined by the model; but the third, $A_{xy}(z)$, is not determined by the model, and it cannot be estimated from the data (since this would require a sample drawn from the *joint* $(\mathbf{x}_t, \mathbf{y}_t)$ process). To proceed, an assumption is necessary.

A common assumption used in econometric analysis is that $A_{xy}(z)$ $= A_x(z)$ so that $\mathbf{x}_t$ and $\mathbf{u}_t$ are uncorrelated at all leads and lags. Equation (1) can then be interpreted as the dynamic analogue of the classical errors-in-variables model. Sargent (1989) discusses this assumption and an alternative assumption, $A_{xy}(z) = A_y(z)$. He points out that under this latter assumption, $\mathbf{u}_t$ can be interpreted as signal extraction error, with $\mathbf{y}_t$ an optimal estimate of the unobserved "signal" $\mathbf{x}_t$.[1] In many applications, these covariance restrictions follow from the way the data were collected or the way expectations are formed. For example, if $\mathbf{x}_t$ represented the true value of the U.S. unemployment rate and $\mathbf{y}_t$ the value published by the U.S. Department of Labor, then $\mathbf{y}_t$ would differ from $\mathbf{x}_t$ because of the sampling error inherent in the monthly Current Population Survey from which $\mathbf{y}_t$ is derived. The sample design underlying the survey implies that the error term, $\mathbf{u}_t$, is statistically independent of $\mathbf{x}_t$. Similarly, if $\mathbf{y}_t$ denoted a rational expectation of $\mathbf{x}_t$, then the error would be uncorrelated with $\mathbf{y}_t$. Neither of these assumptions seems appropriate in the present context. The error is not the result of imprecise measurement. It is not a forecast or signal extraction error. Rather, it represents approximation or abstraction error in the economic model. Any restriction used to identify $A_{xy}(z)$, and hence $A_u(z)$, is arbitrary.[2]

It is possible, however, to calculate a lower bound for the variance of $\mathbf{u}_t$ without imposing any restrictions on $A_{xy}(z)$. When this lower bound on the variance of $\mathbf{u}_t$ is large, then under any assumption on $A_{xy}(z)$, the model fits the data poorly. If the lower bound on the variance of $\mathbf{u}_t$ is small, then there are possible assumptions about $A_{xy}(z)$ that imply that the model fits the data well. Thus the bound is potentially useful for rejecting models on the basis of their empirical

---

[1] The reader familiar with work on data revisions will recognize these two sets of assumptions as the ones underlying the "news" and "noise" models of Mankiw, Runkle, and Shapiro (1984) and Mankiw and Shapiro (1986).

[2] It is interesting to note that it is possible to determine whether the dynamic errors-in-variables model or the signal extraction error model is consistent with the model and the data. The dynamic errors-in-variables model implies that $A_y(z) - A_x(z) \geq 0$ for $|z| = 1$, so that the spectrum of $\mathbf{y}_t$ lies everywhere above the spectrum of $\mathbf{x}_t$; the signal extraction error model implies the converse. If the spectrum of $\mathbf{x}_t$ lies anywhere above the spectrum of $\mathbf{y}_t$, the errors-in-variables model is inappropriate; if the spectrum of $\mathbf{y}_t$ lies anywhere above the spectrum of $\mathbf{x}_t$, the signal extraction model is inappropriate. If the spectra of $\mathbf{x}_t$ and $\mathbf{y}_t$ cross, neither model is appropriate.

fit. Needless to say, models that appear to fit the data well using this bound require further scrutiny.

The bound is calculated by choosing $A_{xy}(z)$ to minimize the variance of u_t subject to the constraint that the implied joint ACGF for x_t and y_t is positive semidefinite. Equivalently, since the spectrum is proportional to the ACGF evaluated at $z = e^{-i\omega}$, the cross spectrum between x_t and y_t, $(2\pi)^{-1}A_{xy}(e^{-i\omega})$, must be chosen so that the spectral density matrix of $(x_t'\ y_t')'$ is positive semidefinite at all frequencies.

Since the measures of fit proposed in this paper are based on the solution to this minimization problem, it is useful to discuss the problem and its solution in detail. Rather than move directly to the solution of the general problem, we shall first solve two simpler problems. This helps develop intuition for the general solution. In the first problem, x_t and y_t are serially uncorrelated scalars, and the representation follows by inspection. In the second problem, x_t and y_t are serially uncorrelated $n \times 1$ vectors, and the solution is slightly more difficult to derive. Finally, in the last problem, x_t and y_t are allowed to be serially correlated.

Model 1

Suppose that x_t, y_t, and u_t are scalar serially uncorrelated random variables. The problem is to choose σ_{xy} to minimize the variance of u_t, $\sigma_u^2 = \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}$, subject to the constraint that the covariance matrix of x_t and y_t remains positive semidefinite, that is, $|\sigma_{xy}| \leq \sigma_x\sigma_y$. By inspection, the solution sets $\sigma_{xy} = \sigma_x\sigma_y$ and yields $\sigma_u^2 = (\sigma_x - \sigma_y)^2$ as the minimum. Since $\sigma_{xy} = \sigma_x\sigma_y$, x_t and y_t are perfectly correlated with

$$x_t = \gamma y_t, \tag{3}$$

where $\gamma = \sigma_x/\sigma_y$. Equation (3) is important because it shows how to calculate fitted values of x_t, given data on y_t. Variants of equation (3) will hold for all the models considered. In each model, the minimum approximation error representation makes $\{x_t\}$ perfectly correlated with $\{y_t\}$. In each model, the analogue of (3) provides a formula for calculating the fitted values of the variables in the model, given data from the actual economy.

Model 2

Now suppose that x_t and y_t are serially uncorrelated random vectors with covariance matrices Σ_x and Σ_y, respectively. Let $\Sigma_u = \Sigma_x + \Sigma_y - \Sigma_{xy} - \Sigma_{yx}$ denote the covariance matrix of u_t. Since Σ_u is a matrix, there is not a unique way to judge how "small" it is. A convenient

measure of the size of $\mathbf{u}_t$ is the trace of $\boldsymbol{\Sigma}_\mathbf{u}$, $\text{tr}(\boldsymbol{\Sigma}_\mathbf{u}) = \sum_{i=1}^n \boldsymbol{\Sigma}_{\mathbf{u},ii}$, where $\boldsymbol{\Sigma}_{\mathbf{u},ij}$ denotes the ijth element of $\boldsymbol{\Sigma}_\mathbf{u}$. While convenient, this measure is not always ideal since it weights all variables equally. Below, we shall find a representation that minimizes $\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_\mathbf{u})$, where $\mathbf{W}$ is a prespecified $n \times n$ matrix. When all variables are equally important, $\mathbf{W} = I_n$, and unequal weighting can be implemented by making $\mathbf{W}$ diagonal with the desired weights as the diagonal elements. The matrix $\mathbf{W}$ can also be used to focus attention on specific linear combinations of the variables that may be particularly interesting. For example, let $\mathbf{G}$ denote an $n \times n$ matrix and suppose that the researcher is primarily interested in the variables $\mathbf{G}\mathbf{x}_t$ and $\mathbf{G}\mathbf{y}_t$. Then since $\text{tr}(\mathbf{G}\boldsymbol{\Sigma}_\mathbf{u}\mathbf{G}') = \text{tr}(\mathbf{G}'\mathbf{G}\boldsymbol{\Sigma}_\mathbf{u})$, $\mathbf{W}$ can be chosen as $\mathbf{G}'\mathbf{G}$.

The problem then is to choose $\boldsymbol{\Sigma}_{\mathbf{xy}}$ to minimize $\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_\mathbf{u})$ subject to the constraint that the covariance matrix of $(\mathbf{x}_t' \ \mathbf{y}_t')'$ is positive semidefinite. The solution is given below for the case in which $\boldsymbol{\Sigma}_\mathbf{x}$ has rank $k \le n$. This occurs, for example, in economic models in which the number of variables exceeds the number of shocks. The solution is summarized in the following proposition.

PROPOSITION. Assume (i) $\text{rank}(\boldsymbol{\Sigma}_\mathbf{x}) = k \le n$, (ii) $\text{rank}(\mathbf{W}\boldsymbol{\Sigma}_\mathbf{x}\mathbf{W}') = \text{rank}(\boldsymbol{\Sigma}_\mathbf{x})$, and (iii) $\text{rank}(\boldsymbol{\Sigma}_\mathbf{y}) = n$. Let $\mathbf{C}_\mathbf{y}$ denote an arbitrary $n \times n$ matrix square root of $\boldsymbol{\Sigma}_\mathbf{y}$ (i.e., $\boldsymbol{\Sigma}_\mathbf{y} = \mathbf{C}_\mathbf{y}\mathbf{C}_\mathbf{y}'$), and let $\mathbf{C}_\mathbf{x}$ denote an arbitrary $n \times k$ matrix square root of $\boldsymbol{\Sigma}_\mathbf{x}$ (i.e., $\boldsymbol{\Sigma}_\mathbf{x} = \mathbf{C}_\mathbf{x}\mathbf{C}_\mathbf{x}'$). Let $\mathbf{USV}'$ denote the singular value decomposition of $\mathbf{C}_\mathbf{y}'\mathbf{W}\mathbf{C}_\mathbf{x}$, where $\mathbf{U}$ is an $n \times k$ orthogonal matrix ($\mathbf{U}'\mathbf{U} = I_k$), $\mathbf{S}$ is a $k \times k$ diagonal matrix, and $\mathbf{V}$ is a $k \times k$ orthonormal matrix. Then $\boldsymbol{\Sigma}_{\mathbf{xy}} = \mathbf{C}_\mathbf{x}\mathbf{VU}'\mathbf{C}_\mathbf{y}'$ is the unique matrix that minimizes $\text{tr}(\mathbf{W}\boldsymbol{\Sigma}_\mathbf{u})$ subject to the constraint that the covariance matrix of $(\mathbf{x}_t' \ \mathbf{y}_t')'$ is positive semidefinite.

The proof is given in the Appendix.

One important implication of this solution is that, like the scalar example, the joint covariance matrix $(\mathbf{x}_t' \ \mathbf{y}_t')'$ is singular and $\mathbf{x}_t$ can be represented as

$$\mathbf{x}_t = \Gamma\mathbf{y}_t, \qquad (4)$$

where $\Gamma = \mathbf{C}_\mathbf{x}'\mathbf{VU}'\mathbf{C}_\mathbf{y}^{-1}$. (Since $\mathbf{U}'\mathbf{U} = \mathbf{VV}' = I_k$, this simplifies to the scalar result when x_t and y_t are scalars.)

Model 3

This same approach can be used in a dynamic multivariate model with slight modifications; when $\mathbf{u}_t$ is serially correlated, the weighted trace of the spectral density matrix rather than the covariance matrix can be minimized.

To motivate the approach, it is useful to use the Cramer representations for $\mathbf{x}_t$, $\mathbf{y}_t$, and $\mathbf{u}_t$ (see, e.g., Brillinger 1981, sec. 4.6). Assume

that $\mathbf{x}_t$, $\mathbf{y}_t$, and $\mathbf{u}_t$ are jointly covariance stationary with mean zero; the Cramer representation can be written as

$$\mathbf{x}_t = \int_0^{2\pi} e^{i\omega t} dz_{\mathbf{x}}(\omega),$$

$$\mathbf{y}_t = \int_0^{2\pi} e^{i\omega t} dz_{\mathbf{y}}(\omega), \tag{5}$$

$$\mathbf{u}_t = \int_0^{2\pi} e^{i\omega t} dz_{\mathbf{u}}(\omega),$$

where $dz(\omega) = (dz_{\mathbf{x}}(\omega)' \ dz_{\mathbf{y}}(\omega)' \ dz_{\mathbf{u}}(\omega)')'$ is a complex valued vector of orthogonal increments, with $E(dz(\omega)\overline{dz(\lambda)}') = \delta(\omega - \lambda)\mathbf{S}(\omega)d\omega d\lambda$, where $\delta(\omega - \lambda)$ is the dirac delta and $\mathbf{S}(\omega)$ is the spectral density matrix of $(\mathbf{x}_t' \ \mathbf{y}_t' \ \mathbf{u}_t')'$ at frequency ω. Equation (5) represents $\mathbf{x}_t$, $\mathbf{y}_t$, and $\mathbf{u}_t$ as the integral (sum) of increments $dz_{\mathbf{x}}(\omega)$, $dz_{\mathbf{y}}(\omega)$, and $dz_{\mathbf{u}}(\omega)$, which are uncorrelated across frequencies and have variances and covariances given by the spectra and cross spectra of $\mathbf{x}_t$, $\mathbf{y}_t$, and $\mathbf{u}_t$. Since the spectra are proportional to the ACGFs evaluated at $z = e^{-i\omega}$, $E(dz_{\mathbf{x}}(\omega)\overline{dz_{\mathbf{x}}(\omega)}')$ is proportional to $A_{\mathbf{x}}(e^{-i\omega})$, $E(dz_{\mathbf{x}}(\omega)\overline{dz_{\mathbf{y}}(\omega)}')$ is proportional to $A_{\mathbf{xy}}(e^{-i\omega})$, and so forth.

Now consider the problem of choosing $A_{\mathbf{xy}}(z)$ to minimize the variance of $\mathbf{u}_t$. Since $\mathbf{u}_t$ can be written as the integral of the uncorrelated increments $dz_{\mathbf{u}}(\omega)$, the variance of $\mathbf{u}_t$ can be minimized by minimizing the variance of $dz_{\mathbf{u}}(\omega)$ for each ω. Since the increments are uncorrelated across frequency, the minimization problems can be solved independently for each frequency. Thus the analysis carried out for model 2 carries over directly, with spectral density matrices replacing covariance matrices. The minimum trace problem for model 2 is now solved frequency by frequency using the spectral density matrix.

Like models 1–2, the solution yields

$$dz_{\mathbf{x}}(\omega) = \Gamma(\omega)dz_{\mathbf{y}}(\omega), \tag{6}$$

where $\Gamma(\omega)$ is the complex analogue of Γ from (4). Equation (6) implies

$$A_{\mathbf{xy}}(e^{-i\omega}) = \Gamma(\omega)A_{\mathbf{y}}(e^{-i\omega}) \tag{7}$$

and

$$A_{\mathbf{u}}(e^{-i\omega}) = A_{\mathbf{x}}(e^{-i\omega}) + A_{\mathbf{y}}(e^{-i\omega}) - A_{\mathbf{xy}}(e^{-i\omega}) - A_{\mathbf{xy}}(e^{i\omega})'. \tag{8}$$

The autocovariances of $\mathbf{u}_t$ follow directly from (8). Moreover, since $dz_{\mathbf{x}}(\omega)$ and $dz_{\mathbf{y}}(\omega)$ are perfectly correlated from (7), $\mathbf{x}_t$ can be represented as a function of leads and lags of $\mathbf{y}_t$:

$$\mathbf{x}_t = \beta(L)\mathbf{y}_t, \tag{9}$$

where $\beta(L) = \sum_{-\infty}^{\infty} \beta_j L^j$, with $\beta_j = \int_{-\pi}^{\pi} \Gamma(\omega) e^{i\omega j} d\omega$. Thus fitted values of $\mathbf{x}_t$ can be calculated from leads and lags of $\mathbf{y}_t$.

An Example

The model considered in the next section describes the dynamic properties of output, consumption, investment, and employment as functions of a single productivity shock. To demonstrate the mechanics of the minimum approximation error representation for that model, assume that $\mathbf{x}_t$ and $\mathbf{y}_t$ are $n \times 1$ vectors and that $\mathbf{x}_t$ is driven by a single iid(0, 1) shock ϵ_t:

$$\mathbf{x}_t = \boldsymbol{\alpha}(L)\epsilon_t, \tag{10}$$

where $\boldsymbol{\alpha}(L)$ is an $n \times 1$ matrix polynomial in the lag operator, L. Assume that the Wold representation for the data is given by

$$\mathbf{y}_t = \boldsymbol{\Theta}(L)\mathbf{e}_t, \tag{11}$$

where $\boldsymbol{\Theta}(L)$ is an $n \times n$ matrix polynomial in L, and $\mathbf{e}_t$ is an $n \times 1$ serially uncorrelated vector with mean zero and identity covariance matrix.

The minimum error representation can then be computed directly from the matrix expressions given in the proposition. From (10), $A_{\mathbf{x}}(z) = \boldsymbol{\alpha}(z)\boldsymbol{\alpha}(z^{-1})'$ and, from (11), $A_{\mathbf{y}}(z) = \boldsymbol{\Theta}(z)\boldsymbol{\Theta}(z^{-1})'$. Suppose that the weighting matrix is $\mathbf{W} = I_n$, so that the trace of the spectral density of $\mathbf{u}_t$ is to be minimized for each frequency. In terms of the matrices in the proposition, $\mathbf{C}_{\mathbf{x}}(\omega) = \boldsymbol{\alpha}(e^{-i\omega})$ and $\mathbf{C}_{\mathbf{y}}(\omega) = \boldsymbol{\Theta}(e^{-i\omega})$. Thus the cross spectrum/cross ACGF for $\mathbf{x}_t$ and $\mathbf{y}_t$ is chosen as $A_{\mathbf{xy}}(e^{-i\omega}) = \boldsymbol{\alpha}(e^{-i\omega})\mathbf{V}(\omega)\mathbf{U}(\omega)'\boldsymbol{\Theta}(e^{i\omega})'$, where $\mathbf{U}(\omega)\mathbf{S}(\omega)\mathbf{V}(\omega)'$ is the singular value decomposition of $\boldsymbol{\Theta}(e^{i\omega})'\boldsymbol{\alpha}(e^{-i\omega})$. (Since $\mathbf{U}(\omega)$ and $\mathbf{V}(\omega)$ are complex matrices, $\mathbf{V}(\omega)'$ and $\mathbf{U}(\omega)'$ denote the transpose conjugates of $\mathbf{V}(\omega)$ and $\mathbf{U}(\omega)$, respectively.) The ACGF for $\mathbf{u}_t$ follows from $A_{\mathbf{u}}(e^{-i\omega}) = A_{\mathbf{x}}(e^{-i\omega}) + A_{\mathbf{y}}(e^{-i\omega}) - A_{\mathbf{xy}}(e^{-i\omega}) - A_{\mathbf{yx}}(e^{-i\omega})$. Finally, to compute fitted values of $\mathbf{x}_t$ from the $\mathbf{y}_t$ realization, note that $dz_{\mathbf{x}}(\omega) = \Gamma(\omega)dz_{\mathbf{y}}(\omega)$, where $\Gamma(\omega) = \boldsymbol{\alpha}(e^{-i\omega})\mathbf{V}(\omega)\mathbf{U}(\omega)'\boldsymbol{\Theta}(e^{-i\omega})^{-1}$.

Relative Mean Square Approximation Error

A bound on the relative mean square approximation error for the economic model can be calculated directly from (8). The bound— analogous to a lower bound on $1 - R^2$ from a regression—is

$$r_j(\omega) = \frac{[A_{\mathbf{u}}(z)]_{jj}}{[A_{\mathbf{y}}(z)]_{jj}}, \quad z = e^{-i\omega}, \tag{12}$$

where $[A_u(z)]_{jj}$ and $[A_y(z)]_{jj}$ are the jth diagonal elements of $A_u(z)$ and $A_y(z)$, respectively. Thus $r_j(\omega)$ is the variance of the jth component of $dz_u(\omega)$ relative to the jth component of $dz_y(\omega)$, that is, the variance of the error relative to the variance of the data for each frequency. A plot of $r_j(\omega)$ against frequency shows how well the economic model fits the data over different frequencies. Integrating the numerator and denominator of $r_j(\omega)$ provides an overall measure of fit. (Note that since $\mathbf{u}_t$ and $\mathbf{x}_t$ are correlated, $r_j(\omega)$ can be larger than one; i.e., the R^2 of the model can be negative.)[3]

One advantage of $r_j(\omega)$ is that it is unaffected by time-invariant linear filters applied to the variables. Filtering merely multiplies both the numerator and denominator of $r_j(\omega)$ by the same constant, the squared gain of the filter. So, for example, $r_j(\omega)$ is invariant to "Hodrick-Prescott" filtering (see Hodrick and Prescott 1980; King and Rebelo 1993) or standard linear seasonal adjustment filters.[4] The integrated version of the relative mean square approximation error is not invariant to filtering since it is a ratio of averages of both the numerator and denominator across frequencies. When the data are filtered, the integrated version of $r_j(\omega)$ changes because the weights implicit in the averaging change. Frequencies for which the filter has a large gain are weighted more heavily than frequencies with a small gain.

III. Measures of Fit for a Real Business Cycle Model

In this section, a standard real business cycle model is evaluated using the measures of fit developed in the last section. The model, which derives from Kydland and Prescott (1982), is the "baseline" model of King, Plosser, and Rebelo (1988b). It is a one-sector neoclassical growth model driven by an exogenous stochastic trend in technology.[5]

---

[3] The measure $r_j(\omega)$ is not technically a metric since it does not satisfy the triangle inequality.

[4] Standard seasonal adjustment filters such as the linear approximation to Census X-11 have zeros at the seasonal frequencies, so that $r_j(\omega)$ is undefined at these frequencies for the filtered data.

[5] This model is broadly similar to the model analyzed in Kydland and Prescott (1982). While the baseline model does not include the complications of time to build, inventories, time-nonseparable utility, and a transitory component to technology contained in the original Kydland and Prescott model, these complications have been shown to be reasonably unimportant for the empirical predictions of the model (see Hansen 1985). Moreover, the King, Plosser, and Rebelo baseline model appears to fit the data better at the very low frequencies than the original Kydland and Prescott model since it incorporates a stochastic trend rather than the deterministic trend present in the Kydland and Prescott formulation.

This baseline model is analyzed, rather than a more complicated variant, for several reasons. First, the calibration/simulation exercises reported in King, Plosser, and Rebelo suggest that the model explains the relative variability of aggregate output, consumption, and investment, and it produces series with serial correlation properties broadly similar to the serial correlation properties of postwar U.S. data. Second, King, Plosser, Stock, and Watson (1991) show that the low-frequency/cointegration implications of the model are broadly consistent with similar postwar U.S. data. Finally, an understanding of where this baseline model fits the data and where it does not fit may suggest how the model should be modified.

Only a brief sketch of the model is presented; a thorough discussion is contained in King, Plosser, and Rebelo (1988a, 1988b). The details of the model are as follows:

$$\text{preferences:} \quad E_t \sum_{t=0}^{\infty} \beta^t u(C_t, L_t), \quad \text{with } u(C_t, L_t) = \log(C_t) + \theta \log(L_t);$$

$$\text{technology:} \quad Q_t = K_t^{1-\alpha}(A_t N_t)^{\alpha},$$

$$\log(A_t) \equiv a_t = \gamma_a + a_{t-1} + \epsilon_t, \quad \epsilon_t \text{ iid}(0, \sigma_\epsilon^2),$$

$$K_t = (1 - \delta)K_{t-1} + I_t;$$

$$\text{constraints:} \quad Q_t = C_t + I_t,$$

$$1 = N_t + L_t,$$

where C_t denotes consumption, L_t is leisure, Q_t is output, K_t is capital, N_t is employment, I_t is investment, and A_t is the stock of technology, with $\log(A_t)$ assumed to follow a random walk with drift γ_a and innovation ϵ_t.

To analyze the model's empirical predictions, the equilibrium of the model must be calculated as a function of the parameters β, θ, α, γ_a, σ_ϵ^2, and δ. This equilibrium implies a stochastic process for the variables C_t, L_t, N_t, K_t, I_t, and Q_t, and these stochastic processes can then be compared to the stochastic processes characterizing U.S. postwar data. As is well known, the equilibrium can be calculated by maximizing the representative agent's utility function subject to the technology and the resource constraints. In general, a closed-form expression for the equilibrium does not exist, and numerical methods must be used to calculate the stochastic process for the variables corresponding to the equilibrium. A variety of numerical approximations have been proposed (see Taylor and Uhlig [1990] for a survey); here I use the log linearization of the Euler equations proposed by King, Plosser, and Rebelo (1987). A formal justification for approximating

the equilibrium of this stochastic nonlinear model near its deterministic steady state using linear methods is provided in Woodford (1986, theorem 2).

The approximate solution yields a vector autoregression (VAR) for the logarithms of Q_t, C_t, K_t, I_t, and N_t. (As in the standard convention, these logarithms will be denoted by lowercase letters.) All of the variables except n_t are nonstationary but can be represented as stationary deviations about a_t, the logarithm of the stock of technology, which by assumption follows an integrated process. Thus q_t, c_t, i_t, and k_t are cointegrated with a single common trend, a_t. Indeed, not only are the variables in the VAR cointegrated, they are singular; the singularity follows since ϵ_t is the only shock to the system. The coefficients in the VAR are complicated functions of the structural parameters β, θ, α, γ_a, σ_ϵ^2, and δ. Values for these parameters are the same as those used by King, Plosser, and Rebelo (1988b): when the variables are measured quarterly, the parameter values are $\alpha = .58$, $\delta = .025$, $\gamma_a = .004$, $\sigma_\epsilon = .010$, and $\beta = .988$, and θ is chosen so that the steady-state value of N is .20. These parameter values were chosen so that the model's steady-state behavior matches postwar U.S. data.[6] With these values for the parameters, the VAR describing the equilibrium can be calculated and the ACGF of $\mathbf{x}_t = (\Delta q_t\ \Delta c_t\ \Delta i_t\ n_t)'$ follows directly.[7]

These autocovariances will be compared to the autocovariances of postwar data for the United States. The data used here are the same data used by King, Plosser, Stock, and Watson (1991). The output measure is total real private GNP, defined as total real GNP less government purchases of goods and services. The measure of consumption is total real consumption expenditures, and the measure of investment is total real fixed investment. The measure of employment is total labor hours in private nonagricultural establishments. All variables are expressed in per capita terms using the total civilian noninstitutional population over the age of 16.[8] Let $\tilde{q}_t$ denote the logarithm

[6] The choice of parameter values is described in King, Plosser, and Rebelo (1988a). The value of α was chosen to equal the average value of labor's share of gross national product over 1948–86. The value of γ_a was chosen as the common average quarterly rate of growth of per capita values of real GNP, consumption of nondurables and services, and gross fixed investment. The depreciation rate was chosen to yield a gross investment share of GNP of approximately 30 percent. The parameter θ was chosen so that the model's steady-state value of N matched the average workweek as a fraction of total hours over 1948–86. The discount rate β was chosen so that the model's steady-state annual interest rate matched the average rate of return on equity over 1948–81. The value of $\sigma_\epsilon = .01$ appears to have been chosen as a convenient normalization. This value is used here because it does a remarkably good job matching the very low frequency movements in output, consumption, and investment.

[7] Of course, this not the only possible definition of $\mathbf{x}_t$. The only restriction on $\mathbf{x}_t$ is covariance stationarity, so, e.g., $c_t - q_t$ and $i_t - q_t$ could be included as elements.

[8] All data are taken from Citibase. With the Citibase labels, the precise variables used were gnp82 − gge82 for output, gc82 for consumption, and gif82 for investment. The

of per capita private output, $\bar{c}_t$ the logarithm of per capita consumption expenditures, and so forth. Then the data used in the analysis can be written as $\mathbf{y}_t = (\Delta \bar{q}_t \, \Delta \bar{c}_t \, \Delta \bar{\imath}_t \, \bar{n}_t)'$.

The analysis presented in the last section assumed that the ACGF/spectrum of $\mathbf{y}_t$ was known. In practice, of course, this is not the case, and the spectrum must be estimated. In this work, the spectrum of $\mathbf{y}_t$ was estimated in two different ways. First, an autoregressive spectral estimator was used, calculated by first estimating a VAR for the variables and then forming the implied spectral density matrix. As in King, Plosser, Stock, and Watson (1991), the VAR was estimated imposing a cointegration constraint between output, consumption, and investment. Thus the VAR was specified as the regression of $\mathbf{w}_t = (\Delta \bar{q}_t, \, \bar{q}_t - \bar{c}_t, \, \bar{q}_t - \bar{\imath}_t, \, \bar{n}_t)'$ onto a constant and six lags of $\mathbf{w}_t$. The parameters of the VAR were estimated using data for 1950–88. (Values before 1950 were used as lags in the regression for the initial observations.) Second, a standard nonparametric spectral estimator was also calculated. The spectrum was estimated by a simple average of 10 periodogram ordinates after prewhitening employment with the filter $1 - .95L$. These two estimators yielded similar values for the measures of fit, and to conserve space only the results for the autoregressive spectral estimator are reported.

For each variable, figure 1 presents the spectrum implied by the model, the spectrum of the data, and the spectrum of the error required to reconcile the model with the data.[9] The error process was chosen to minimize the unweighted trace of the error spectral density matrix, subject to the positive semidefiniteness constraint discussed in the last section. Thus the objective function weighted all the variables equally. For output, consumption, and investment, the model and data spectra differ little for very low frequencies (periods greater than 50 quarters) and for output and investment at high frequencies (periods less than five quarters). There are significant differences between the model and data spectra for periods typically associated with the business cycle; the largest differences occur at a frequency corresponding to approximately 20 quarters. The spectra of Δn_t and $\Delta \bar{n}_t$ are quite different. In addition to large differences at business cycle frequencies, the spectra are also very different at low frequencies. The model implies that employment is stationary so that its

---

measure of total labor hours was constructed as total employment in nonagricultural establishments (lhem) less total government employment (lpgov) multiplied by average weekly hours (lhch). The population series was P16.

[9] Figure 1 is reminiscent of figures in Howrey (1971, 1972), who calculated the spectra implied by the Klein-Goldberger and Wharton models. A similar exercise is carried out in Soderlind (1993), who compares the spectra of variables in the Kydland-Prescott model to the spectra of postwar U.S. data.

growth rate has a spectrum that vanishes at frequency zero. In contrast, the data suggest significant low-frequency variation in postwar U.S. employment.[10]

The figure shows that relatively little error is needed to reconcile the model and the data for output, consumption, and investment over the very low frequencies. On the other hand, error with a variance on the order of 40–50 percent of the magnitude of the variance of the series is necessary for the components of output, consumption, and investment with periods in the 6–32-quarter range. At higher frequencies, the model is able to match the stochastic process describing investment, but not the processes describing the other series.

Table 1 provides a summary of the relative mean square approximation error (RMSAE) for a variety of weighting functions and filters. Each panel shows the RMSAE for the variables for five different minimum error representations. Column 1 presents results for the representation that obtains when the unweighted trace of the spectrum is minimized; this is the representation used to construct the error spectra shown in figure 1. Column 2 summarizes the results for the representation that minimizes the output error, with no weight placed on the other variables. Column 3 summarizes results for the representation that minimizes the consumption error, and so forth. Panel A presents the results for the differences of the data integrated across all frequencies, panel B shows results for the levels of the series detrended by the Hodrick-Prescott filter integrated across all frequencies, and panel C presents results for the levels of the series integrated over business cycle frequencies (6–32 quarters). The trade-off inherent in the different representations is evident in all panels. For example, in panel A, with the minimum output error representation, the RMSAE for output is 26 percent, and the RMSAE for consumption is 78 percent; when the minimum consumption error representation is chosen, the RMSAE for consumption falls to 21 percent but the RMSAE for output rises to 75 percent. When all the variables are equally weighted, the RMSAE is 52 percent for output and 66 percent for consumption. Panel C shows that most of this trade-off occurs at the high frequencies, at least for output, consumption, and investment; over the business cycle frequencies their RMSAEs are in the 40–60 percent range.

As explained in Section II, fitted values of the model's variables can be constructed using the minimum error representation together

---

[10] The figures do not include standard errors for the spectra estimated from these data. These standard errors are large—approximately one-third the size of the estimated spectra. The standard errors for the RMSAE, averaged across frequencies, are considerably smaller. These are included in tables 1 and 2 below.
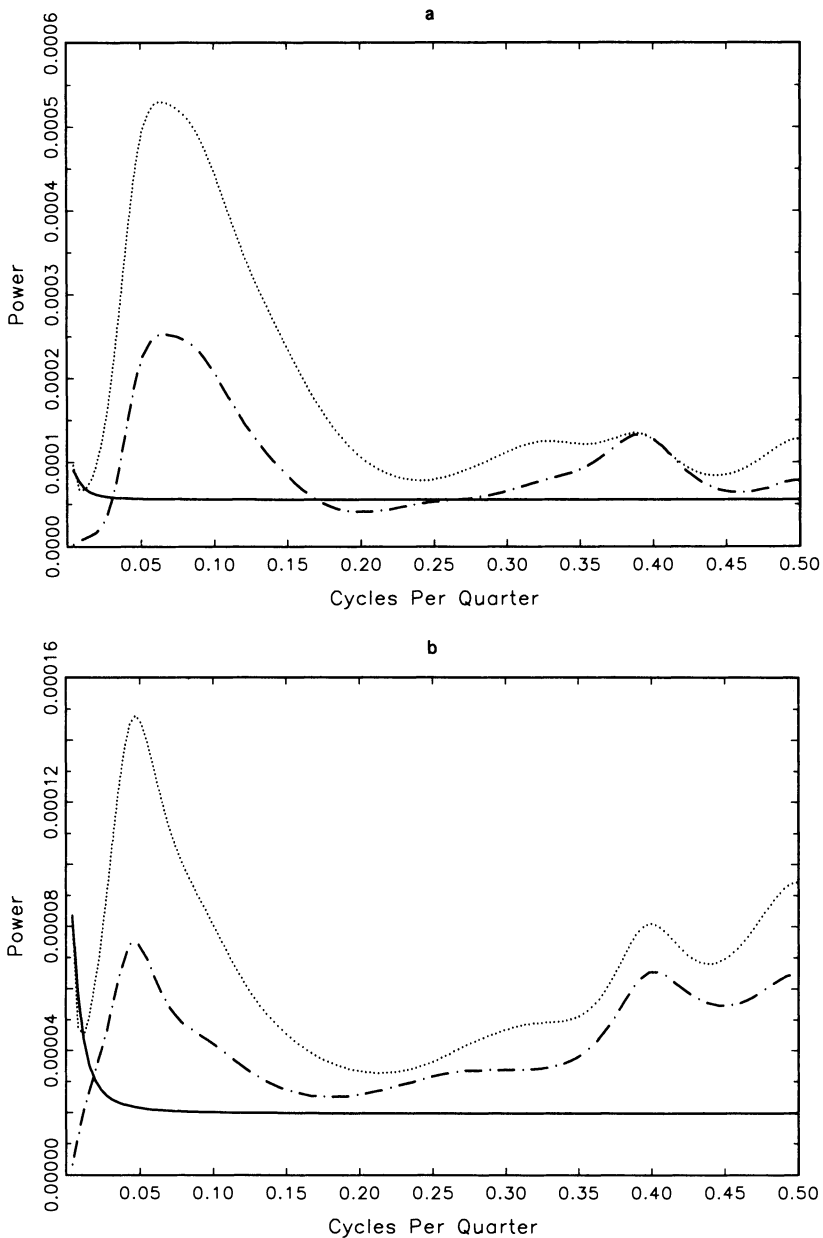
FIG. 1.—Decomposition of spectra: *a*, output; *b*, consumption; *c*, investment; *d*, employment. Dotted lines refer to the data, dashed lines to the model, and solid lines to approximation error.
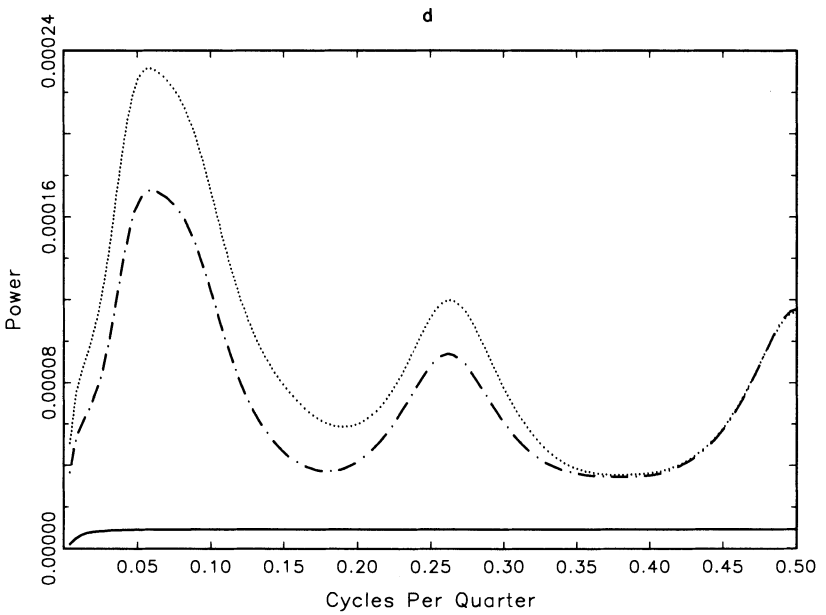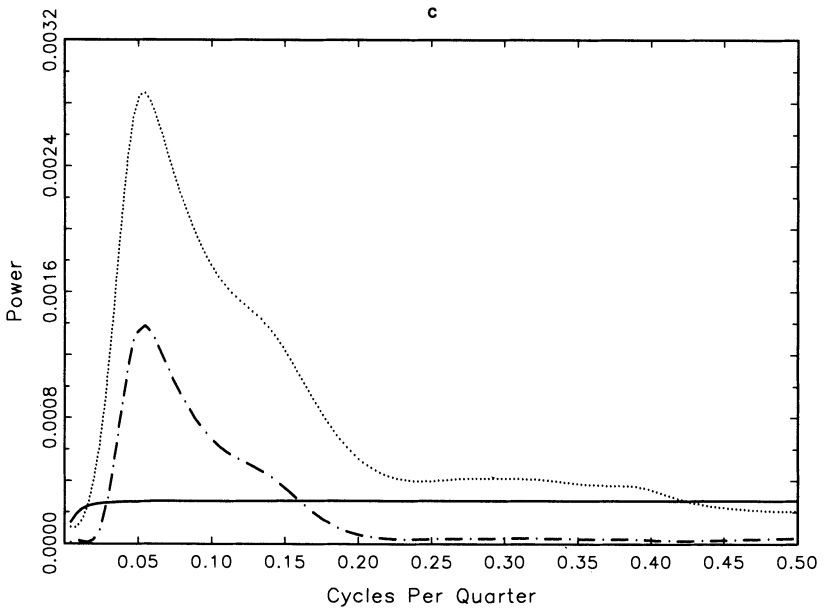
c

d

FIG. 1.—*Continued*

TABLE 1

RELATIVE MEAN SQUARE APPROXIMATION ERROR

VARIABLE	Equal Weight (1)	ERROR MINIMIZED WITH RESPECT TO			
		Output (2)	Consumption (3)	Investment (4)	Employment (5)
A. First Differences: All Frequencies					
Output	.52 (.04)	.26 (.06)	.75 (.06)	.64 (.05)	.79 (.07)
Consumption	.66 (.06)	.78 (.08)	.21 (.05)	.75 (.08)	1.00 (.10)
Investment	.29 (.06)	.63 (.05)	.75 (.07)	.27 (.06)	.79 (.07)
Employment	.78 (.04)	.81 (.04)	.92 (.04)	.81 (.04)	.50 (.04)
B. Hodrick-Prescott Detrended Levels: All Frequencies					
Output	.43 (.05)	.37 (.07)	.58 (.04)	.46 (.04)	.63 (.05)
Consumption	.53 (.04)	.55 (.05)	.29 (.09)	.57 (.05)	.81 (.09)
Investment	.40 (.08)	.48 (.05)	.61 (.04)	.39 (.08)	.63 (.04)
Employment	.73 (.03)	.75 (.03)	.85 (.04)	.74 (.03)	.59 (.06)
C. Levels: 6–32 Quarters					
Output	.43 (.06)	.39 (.07)	.55 (.05)	.44 (.05)	.61 (.05)
Consumption	.51 (.05)	.51 (.06)	.32 (.09)	.53 (.05)	.77 (.08)
Investment	.42 (.08)	.46 (.06)	.59 (.04)	.42 (.08)	.61 (.05)
Employment	.72 (.03)	.74 (.03)	.84 (.04)	.73 (.03)	.61 (.06)

NOTE.—Relative mean square approximation error is the lower bound of the variance of the approximation error divided by the variance of the series. Each column represents the relative mean square approximation error of the row variable constructed from the representation that minimizes the weighted trace of the error spectrum. The weights are summarized by the column headings. For example, col. 1 is the equally weighted trace, col. 2 puts all the weight on the output error, etc. The numbers in parentheses are standard errors based on the sampling error in the estimated VAR coefficients used to estimate the data spectrum.

with the observed data. Since the measurement error model repre-
sents $\mathbf{y}_t$ as $\mathbf{x}_t$ plus error, the standard signal extraction formula can be
used to extract $\{\mathbf{x}_t\}$ from $\{\mathbf{y}_t\}$. In general, of course, signal extraction
methods will yield an estimate of $\mathbf{x}_t$, say $\hat{\mathbf{x}}_t$, that is not exact in the
sense that $E[(\mathbf{x}_t - \hat{\mathbf{x}}_t)^2] \neq 0$. In the present context, the estimate will
be exact since the measurement error process is chosen so that $dz_\mathbf{x}(\omega)$
and $dz_\mathbf{y}(\omega)$ are perfectly correlated for all ω.[11] Figure 2 shows the
realizations of the data and the realizations of the variables in the
model calculated from the data using the equally weighted minimum
output error representation.[12]

   In figure 2a, which shows the results for output, the model seems
capable of capturing the long swings in the postwar U.S. data but not
capable of capturing the cyclical variability in the data. Private per
capita GNP fell by 8.4 percent from the cyclical peak in 1973 to the
trough in 1975 and by 6.8 percent from the peak in 1979 to the
trough in 1982. In contrast, the corresponding drops in Q_t—output
in the model—were 3.1 percent and 3.0 percent, respectively. The
dampened cyclical swings in consumption and investment, shown in
figure 2b and c, are even more dramatic. Finally, figure 2d shows that
the model predicts changes in employment that have little to do with
the changes observed in the United States during the postwar
period.[13]

   One possible explanation for the relatively poor fit of the model is
that the "calibrated" values of the parameters are wrong. In particu-
lar, Christiano and Eichenbaum (1990) show that the model's predic-
tions change in an important way when the technology process
changes from a random walk to a stationary AR(1). Table 2 shows
how the model fares for a range of values of the AR(1) coefficient
for technology, denoted by ρ_a. Panel A of the table shows the results
for first differences of the variables across all frequencies, panel B
presents results for the Hodrick-Prescott detrended levels of the se-
ries, and panel C shows the results for the levels of the series over
the "business cycle" frequencies. From panel A, the value of ρ_a has

---

[11] More precisely, the estimate is exact in the sense that $\mathrm{Proj}(\mathbf{x}_t|\mathbf{y}_{t-j}, \ldots, \mathbf{y}_{-1}, \mathbf{y}_0,$
$\mathbf{y}_1, \ldots, \mathbf{y}_{t+j})$ converges in mean square to $\mathbf{x}_t$ as $j \rightarrow \infty$.
[12] As shown in eq. (9), $\mathbf{x}_t$ can be calculated as $\beta(L)\mathbf{y}_t$, where $\beta(L)$ is the inverse Fourier
transform of $\Gamma(\omega)$. To calculate the estimates shown in the figure, $\Gamma(\omega)$ was calculated
at 128 equally spaced frequencies between zero and π. Since $\beta(L)$ is two-sided, pre-
and postsample values of $\mathbf{y}_t$ are required to form $\beta(L)\mathbf{x}_t$. These pre- and postsample
values were replaced with the sample means of the $\mathbf{y}_t$ data. The first differences, $\mathbf{x}_t$
and $\mathbf{y}_t$, were then accumulated to form the levels series shown in the figure.
[13] The calculations required to construct figs. 1 and 2 and the results in table 1 are
easily carried out. For this example, the model spectrum, data spectrum, RMSAEs,
and fitted values were calculated in less than a minute on a standard desktop computer.
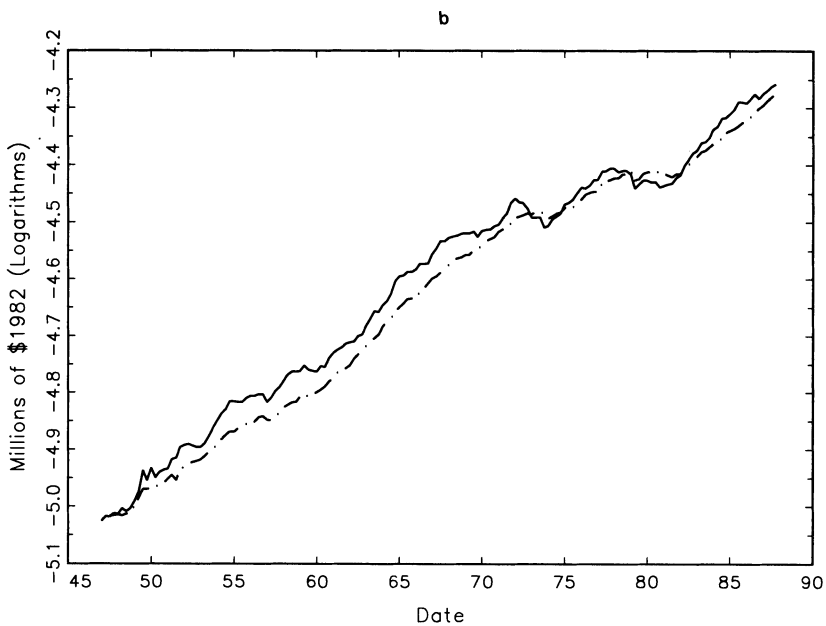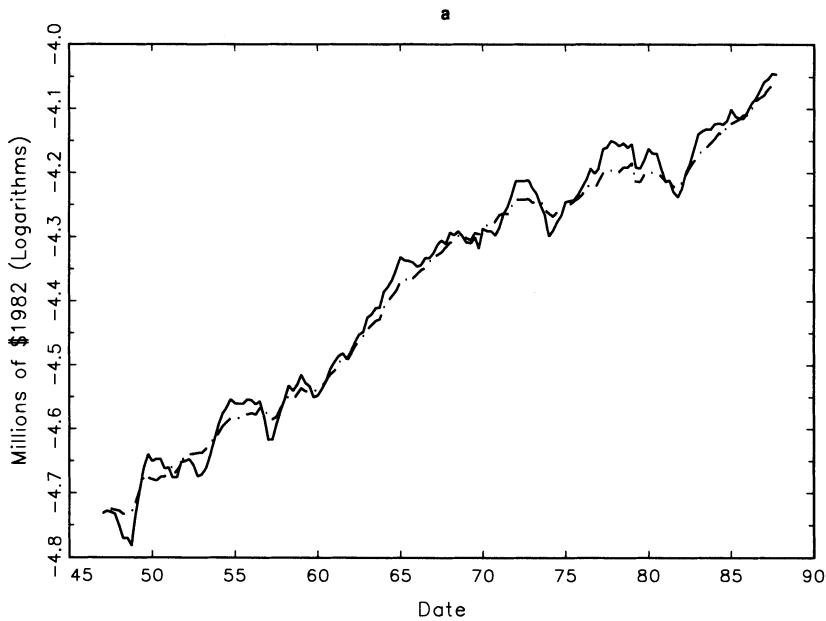A GAUSS program for these calculations is available from the author.

FIG. 2.—Data: *a*, output; *b*, consumption; *c*, investment; *d*, employment. Solid lines refer to U.S. data and dashed lines to realizations from the model.
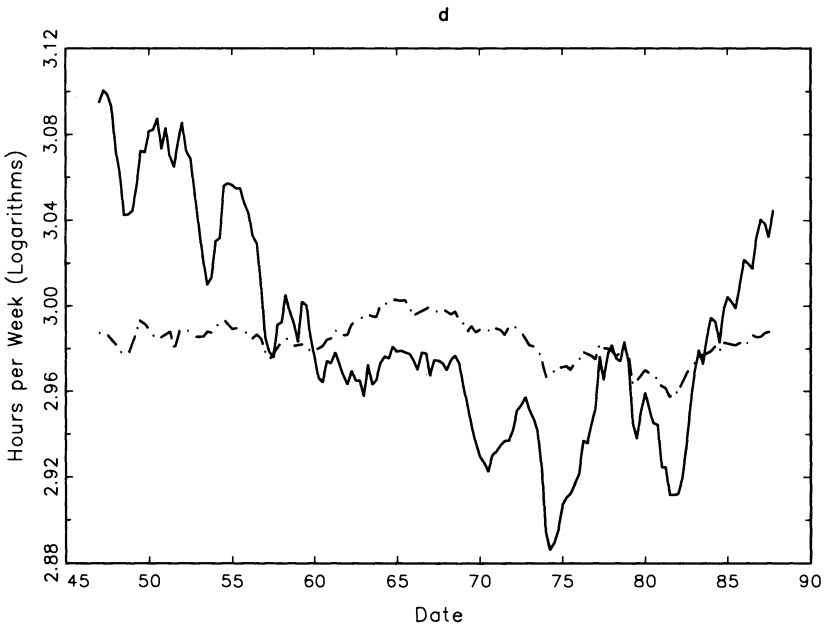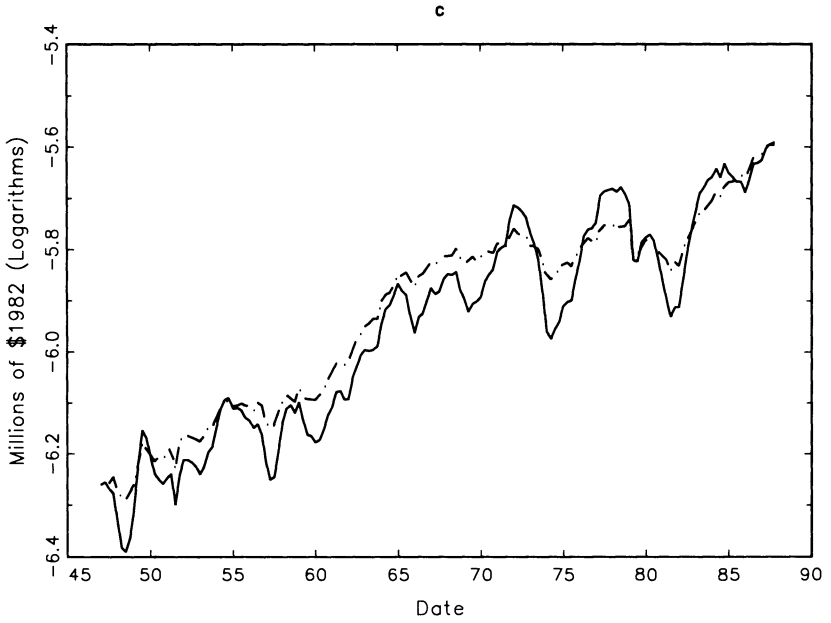
c

d

FIG. 2.—*Continued*

TABLE 2

RELATIVE MEAN SQUARE APPROXIMATION ERROR AS A FUNCTION OF THE AR(1)
COEFFICIENT FOR TECHNOLOGY

	ρ_a			
VARIABLE	1.0 (1)	.98 (2)	.95 (3)	.90 (4)
A. First Differences: All Frequencies				
Output	.52 (.04)	.52 (.04)	.53 (.05)	.54 (.06)
Consumption	.66 (.06)	.69 (.06)	.71 (.05)	.74 (.05)
Investment	.29 (.06)	.21 (.05)	.18 (.05)	.20 (.04)
Employment	.78 (.04)	.74 (.05)	.73 (.07)	.75 (.09)
B. Hodrick-Prescott Detrended Levels: All Frequencies				
Output	.43 (.05)	.40 (.05)	.37 (.05)	.36 (.05)
Consumption	.53 (.04)	.58 (.04)	.63 (.04)	.67 (.04)
Investment	.40 (.08)	.29 (.08)	.21 (.08)	.17 (.08)
Employment	.73 (.03)	.64 (.04)	.58 (.04)	.55 (.04)
C. Levels: 6–32 Quarters				
Output	.43 (.06)	.38 (.06)	.35 (.06)	.33 (.06)
Consumption	.51 (.05)	.56 (.05)	.61 (.05)	.66 (.05)
Investment	.42 (.08)	.30 (.09)	.22 (.09)	.17 (.09)
Employment	.72 (.03)	.63 (.04)	.56 (.05)	.52 (.05)

NOTE.—Relative mean square approximation error is the lower bound of the variance of the approximation error divided by the variance of the series. Each column represents the relative mean square approximation error of the row variable constructed from the representation that minimizes the weighted trace of the error spectrum. The column headings represent the AR(1) coefficient for the process of the logarithm of productivity in the model. For example, col. 1 represents results for the model with random walk technological progress. The numbers in parentheses are standard errors based on the sampling error in the estimated VAR coefficients used to estimate the data spectrum.

little effect on the fit of the model averaged across all frequencies. In particular, as ρ_a falls from 1.0 to .90, the RMSAE increases slightly for consumption, falls for investment, and changes little for output and employment. In contrast, the value of ρ_a has a significant effect on the fit of the model over business cycle frequencies. For example, panel C shows that as ρ_a falls from 1.0 to .90, the RMSAE falls for output (.43 to .33), for investment (.42 to .17), and for employment (.72 to .52); it increases for consumption (.52 to .66).

The source of the changes in the RMSAEs can be seen in figure 3, which plots the spectra of the variables in models with $\rho_a = 1$ and $\rho_a = .90$. The largest difference between the spectra of the models is the increase in variance in output, investment, and employment as ρ_a falls from 1.0 to .90. The economic mechanism behind this increased variance is the increase in intertemporal substitution in response to a technology shock. When $\rho_a = .90$, technology shocks are transitory, and they can be exploited only by large transitory increases

in employment and investment. It is interesting to note that while this mechanism increases the variance of the growth rates in employment, investment, and output, it has little effect on their autocorrelations. That is, as ρ_a changes from 1.0 to .90, the shape of the spectra changes little.

Before we leave this section, six additional points deserve mention. First, the fitted values in figure 2 are quantitatively and conceptually similar to figures presented in Christiano (1988) and Plosser (1989). They calculated the Solow residual from actual data and then simulated the economic model using this residual as the forcing process. Implicitly, they assumed that the model and data were the same in the terms of their Solow residual, and then asked whether the model and data were similar in other dimensions. Figure 2 is constructed by making the model and data as close as possible in one dimension (in this case the trace of the variance of the implied approximation error) and then asking whether the model and data are similar in other dimensions. The difference between the two approaches can be highlighted by considering the circumstances in which they would produce exactly the same figure. If the Solow residual computed from the actual data followed exactly the same stochastic process as the change in productivity in the model, and if the approximation error representation was constructed by minimizing the variance of the difference between the Solow residual in the data and productivity growth in the model, then the two figures would be identical. Thus the figures will differ if the stochastic process for the empirical Solow residual is not the same as assumed in the model, or the approximation error representation is chosen to make the model and data close in some dimension other than productivity growth.

Second, the inability of the model to capture the business cycle properties of the data is not an artifact of the minimum measurement error representation used to form the projection of $\mathbf{x}_t$ onto $\mathbf{y}_\tau$, $\tau = 1, \ldots, n$. Rather, it follows directly from a comparison of the spectra of $\mathbf{x}_t$ and $\mathbf{y}_t$. The fitted values are constrained to have an ACGF/spectra given by the economic model. Figure 1 shows that, for all the variables, the spectral power over the business cycle frequencies is significantly less for the model than for the data. Therefore, fitted values from the model are constrained to have less cyclical variability than the data.

Third, the ability of the model to mimic the behavior of the data depends critically on the size of the variance of the technology shock. The value of σ_ϵ used in the analysis is two and one-half times larger than the drift in the series. Thus if the ϵ_t were approximately normally distributed, the stock of technology A_t would fall in one out of three quarters on average. Reducing the standard deviation of the
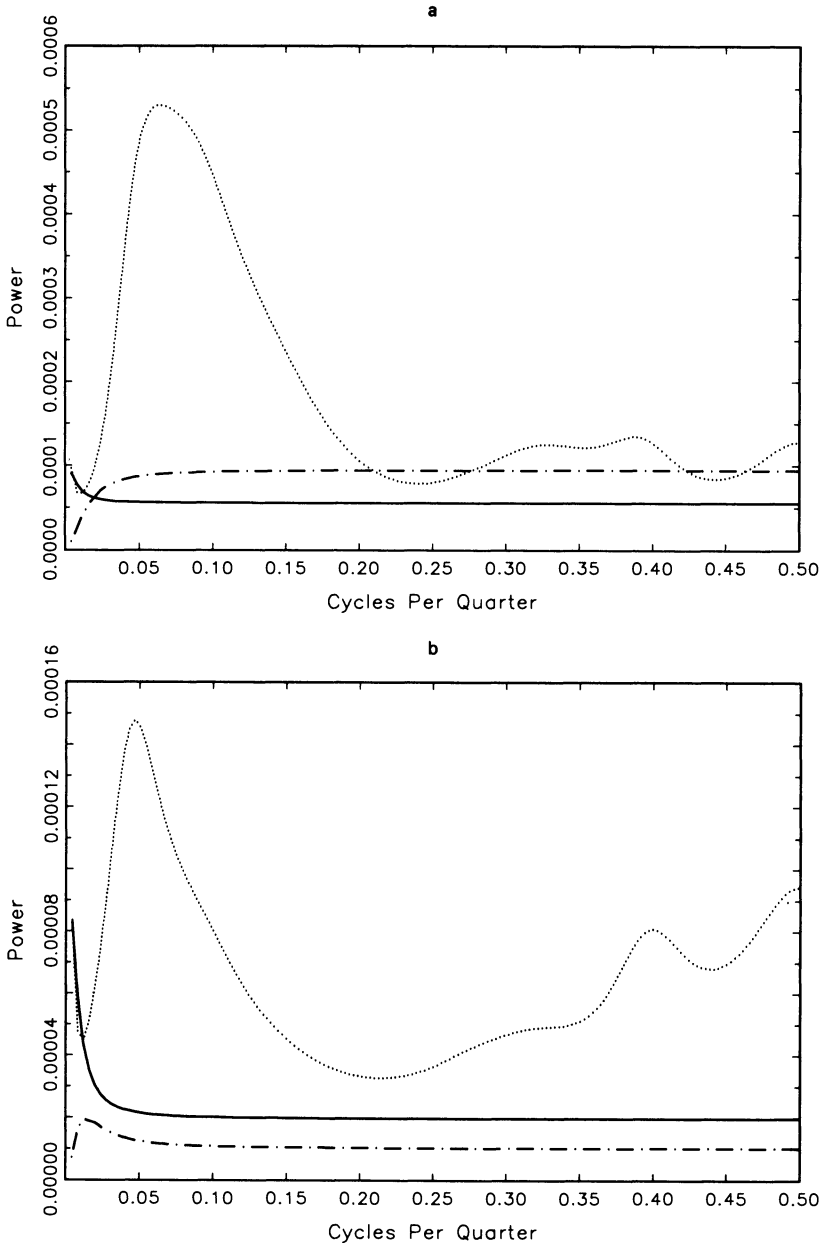
FIG. 3.—Data and model spectra: *a*, output; *b*, consumption; *c*, investment; *d*, employment. Dotted lines refer to the data, solid lines to the model with $\rho_a = 1.00$, and dashed lines to the model with $\rho_a = .90$.
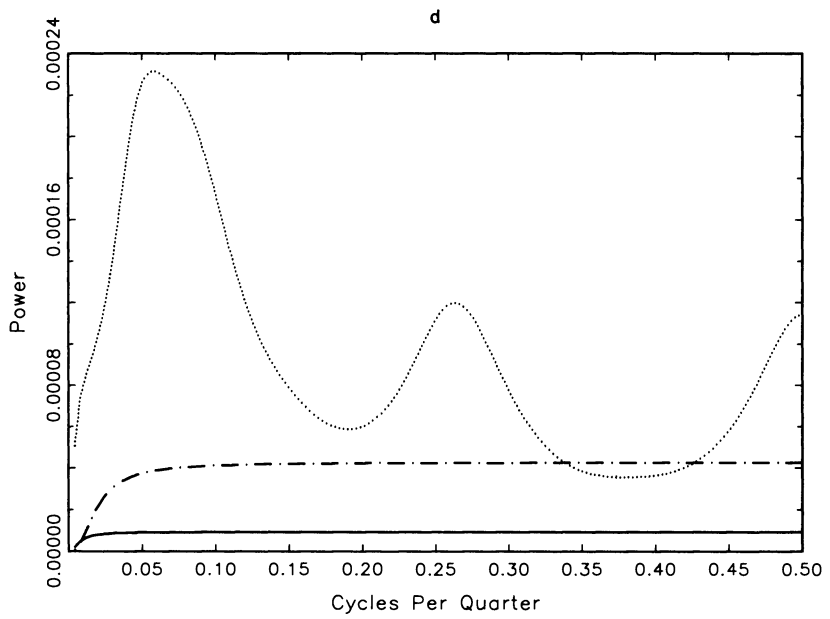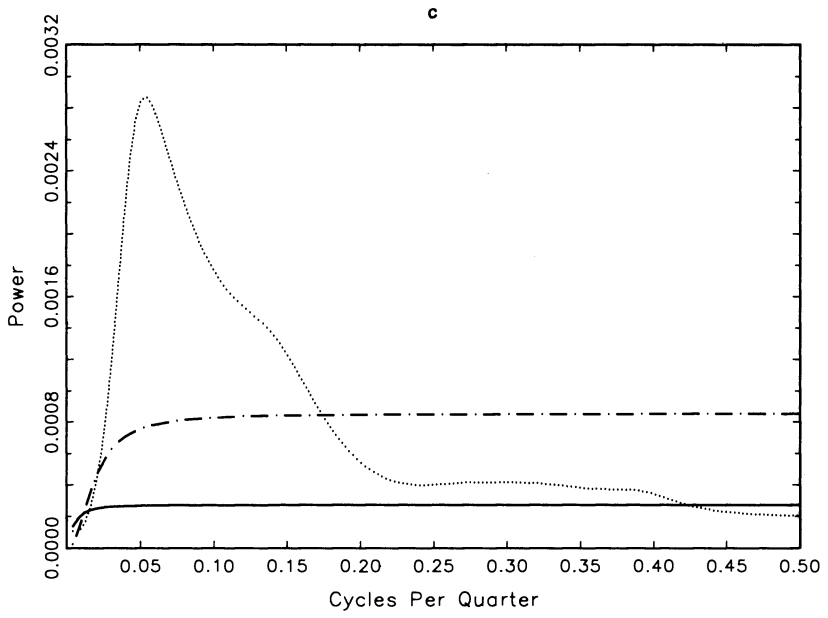
c

Power

Cycles Per Quarter

d

Power

Cycles Per Quarter

FIG. 3.—*Continued*

technology shock so that it equals the average growth in a_t drastically increases the size of the measurement error necessary to reconcile the model with the data. For example, integrated across all frequencies, the RMSAE for output increases from 52 percent to 74 percent.

Fourth, there is nothing inherent in the structure of the model that precludes the use of classical statistical procedures. Altug (1989) used maximum likelihood methods to study a version of the model that is augmented with serially correlated classical measurement errors. Singleton (1988) and Christiano and Eichenbaum (1992) pointed out that generalized method of moments procedures can be used to analyze moment implications of models like the one presented above. In the empirical work of Christiano and Eichenbaum the singularity in the probability density function of the data that is implied by the model was finessed in two ways. First, limited information estimation and testing methods were used, and second, the authors assumed that their data on employment were measured with error.

Fifth, many if not all of the empirical shortcomings of this model have been noted by other researchers. King, Plosser, and Rebelo clearly show that the model is not capable of explaining the variation in employment that is observed in the actual data. The implausibility of the large technology shocks is discussed in detail in Summers (1986), Mankiw (1989), and McCallum (1989).

Finally, the analysis above has concentrated on the ability of the model to explain the variability in output, consumption, investment, and employment across different frequencies. While it is possible to analyze the covariation of these series using the cross spectrum of the measurement error, such an analysis has not been carried out here. This is a particularly important omission since this is the dimension in which the baseline real business cycle model is typically thought to fail. For example, Christiano and Eichenbaum (1992) and Rotemberg and Woodford (1992) use the model's counterfactual implication of a high correlation between average productivity and output growth as starting points for their analysis, and the empirical literature on the intertemporal capital asset pricing model beginning with Hansen and Singleton (1982) suggests that the asset pricing implications of the model are inconsistent with the data. It would be useful to derive simple summary statistics based on the cross spectra of the measurement error and the data to highlight the ability of the model to explain covariation among the series.

IV. Discussion

The discussion thus far has assumed that the parameter values of the economic model are known. A natural question is whether the

measures of fit discussed in this paper can form the basis for estima-
tors of these parameters. Does it make sense, for example, to estimate
unknown parameters by minimizing some function of the relative
mean square error, $r_j(\omega)$ given in equation (12)? This certainly seems
sensible. For example, a researcher may want to "calibrate" his model
with a value of $\rho_a = .90$ rather than 1.0, because this value produces
spectra closer to the estimated spectra of data over the business cycle
frequencies. Yet dropping the standard statistical assumption that the
economic model is correctly specified raises a number of important
issues. Foremost among these is the meaning of the parameters. If
the model does not necessarily describe the data, then what do the
parameters measure? Presumably, the model is meant to describe
certain characteristics of the data's stochastic process (the business
cycle or the growth properties, for example), while ignoring other
characteristics. It then makes sense to define the model's parameters
as those that minimize the differences between the model and the
data's stochastic process in dimensions that the model is attempting
to explain. So, for example, it seems sensible to define the parameters
of a growth model as those that minimize $r_j(\omega)$ over very low frequen-
cies, or to define the parameters of a business cycle model as those
that minimize $r_j(\omega)$ over business cycle frequencies. Given this defini-
tion of the parameters, constructing an analog estimator (see Manski
1988) by minimizing $\hat{r}_j(\omega)$ corresponds to a standard statistical
practice.

   Of course, the parameters may also be defined using other charac-
teristics of the model and the stochastic process describing the data.
For example, in standard "calibration" estimation exercises, many of
the parameters are implicitly defined in terms of first moments of
the data. Parameters are chosen so that the first moments of the
variables in the model's steady state match the first moments of the
data.

   Two final points deserve mention. First, since the measures of fit
developed in this paper are based on a representation that minimizes
the discrepancy between the model and the data, they serve only as
a bound on the fit of the model. Models with large RMSAEs do not
fit the data well. Models with small RMSAEs fit the data well given
certain assumptions about the correlation properties of the noise that
separates the model and the data, but may fit the data poorly given
other assumptions about the noise.

   Finally, while this paper has concentrated on measures of fit moti-
vated by a model of measurement error, other measures are certainly
possible. For example, one measure, which like the measures in this
paper uses only the autocovariances implied by the model and the
data, is the expected log likelihood ratio using the normal probability

density function (pdf) of the data and the model. More precisely, if $g(x)$ denotes the normal pdf constructed from the autocovariances of the data, $f(x)$ denotes the normal pdf constructed for the autocovariances implied by the model, and E_g is the expectation operator taken with respect to $g(x)$, the expected log likelihood ratio $I(g, f) = E_g\{\log[g(x)/f(x)]\}$ can be used to measure the distance between the densities $f(\cdot)$ and $g(\cdot)$; $I(g, f)$ is the Kullback-Leibler information criterion (KLIC), which plays an important role in the statistical literature on model selection (e.g., Akaike 1973) and quasi–maximum likelihood estimation (White 1982). Unfortunately, the KLIC will not be defined when $f(x)$ is singular and $g(x)$ is not; the KLIC distance between the two densities is infinite. Thus, for example, it would add no additional information on the fit of the real business cycle model analyzed in Section III beyond pointing out the singularity.

Arguably, one of the most informative diagnostics presented in this paper is the plot of the model and data spectra. For example, figures 1 and 2 show that the data spectra have mass concentrated around the business cycle frequencies, but the model spectra do not. Any metric comparing the data and model spectra may serve as a useful measure of fit. The RMSAE proposed here has the advantage that it can be interpreted like $1 - R^2$ from a regression, but any summary statistic discards potentially useful information contained in plots such as figures 1 and 2. Some practical advice, therefore, is to present both model and data spectra as a convenient way of comparing their complete set of second moments.

Appendix

To prove the proposition, first parameterize Σ_x, Σ_y, and Σ_{xy} as

$$\Sigma_x = C_x C_x', \tag{A1}$$

$$\Sigma_y = GG' + \Sigma, \tag{A2}$$

$$\Sigma_{xy} = C_x G', \tag{A3}$$

where C_x is $n \times k$ with full column rank, G is $n \times k$, and Σ is positive semidefinite. Since $\Sigma_u = \Sigma_x + \Sigma_y - \Sigma_{xy} - \Sigma_{yx}$, minimizing $\text{tr}(W\Sigma_u)$ with Σ_x and Σ_y given is equivalent to maximizing $\text{tr}(W\Sigma_{xy}) = \text{tr}(WC_x G')$. Given an arbitrary factorization of Σ_x of the form (A1), the problem is to find the $n \times k$ matrix G to maximize $\text{tr}(WC_x G')$ subject to the constraint that $\Sigma_y - GG' = \Sigma$ is positive semidefinite.

Let $\hat{G} = C_y^{-1}G$. Then $\Sigma_y - GG'$ is positive semidefinite if and only if $I_n - \hat{G}\hat{G}'$ is positive semidefinite, which in turn is true if and only if all the eigenvalues of $\hat{G}\hat{G}'$ are less than or equal to one. Since the eigenvalues of $\hat{G}\hat{G}'$ are the same as those of $\hat{G}'\hat{G}$, the problem can be written as

$$\max_{\hat{G}} \text{tr}[\hat{G}'(C_y'WC_x)] \quad \text{subject to} \quad \lambda_i(\hat{G}'\hat{G}) \le 1, \quad i = 1, \ldots, k, \tag{A4}$$

where $\lambda_i(\hat{\mathbf{G}}'\hat{\mathbf{G}})$ denotes the ith eigenvalue of $\hat{\mathbf{G}}'\hat{\mathbf{G}}$, and I have used the fact that $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ for conformable matrices $\mathbf{A}$ and $\mathbf{B}$.

Let $\mathbf{QDR}'$ denote the singular value decomposition of $\hat{\mathbf{G}}$, where $\mathbf{Q}$ is an $n \times k$ orthogonal matrix, $\mathbf{R}$ is a $k \times k$ orthonormal matrix, and $\mathbf{D}$ is a $k \times k$ diagonal matrix with elements d_{ij}. Since $\text{tr}[\hat{\mathbf{G}}'(\mathbf{C}_y'\mathbf{WC}_x)] = \text{tr}[\mathbf{RDQ}'(\mathbf{C}_y'\mathbf{WC}_x)]$ $= \text{tr}[\mathbf{DQ}'(\mathbf{C}_y'\mathbf{WC}_x)\mathbf{R}]$, and since $\lambda_i(\hat{\mathbf{G}}'\hat{\mathbf{G}}) = d_{ii}^2$, the solution to (A4) is seen to require that $\lambda_i(\hat{\mathbf{G}}'\hat{\mathbf{G}}) = 1$, $i = 1, \ldots, k$. This implies that $\hat{\mathbf{G}}'\hat{\mathbf{G}} = I$. Write the singular value decomposition of $\mathbf{C}_y'\mathbf{WC}_x$ as $\mathbf{USV}'$; then $\text{tr}[\hat{\mathbf{G}}'(\mathbf{C}_y'\mathbf{WC}_x)] = \text{tr}(\hat{\mathbf{G}}'\mathbf{USV}') = \text{tr}(\mathbf{V}'\hat{\mathbf{G}}'\mathbf{US}) = \text{tr}(\overline{\mathbf{G}}'\mathbf{US})$, where $\overline{\mathbf{G}} = \hat{\mathbf{G}}\mathbf{V}$. Since $\overline{\mathbf{G}}'\overline{\mathbf{G}} = \mathbf{V}'\hat{\mathbf{G}}'\hat{\mathbf{G}}\mathbf{V} = I_k$, the maximization problem can be written as

$$\max_{\overline{\mathbf{G}}} \text{tr}(\overline{\mathbf{G}}'\mathbf{US}) \quad \text{subject to} \quad \overline{\mathbf{G}}'\overline{\mathbf{G}} = I_k. \tag{A5}$$

Assumptions i–iii of the proposition imply that $\mathbf{C}_y'\mathbf{WC}_x$ has full column rank so that $\mathbf{S}$ is a diagonal matrix with strictly positive diagonal elements. Thus since $\mathbf{U}'\mathbf{U} = I_k$, the maximization is achieved by $\overline{\mathbf{G}} = \mathbf{U}$. Working backward, we see that $\mathbf{G} = \mathbf{C}_y\mathbf{UV}'$, so that $\Sigma_{xy} = \mathbf{C}_x\mathbf{G}' = \mathbf{C}_x\mathbf{VU}'\mathbf{C}_y'$.

Uniqueness follows since this choice of Σ_{xy} does not depend on the (arbitrary) choice of the matrix square roots, $\mathbf{C}_x$ and $\mathbf{C}_y$. To see this, let $\hat{\mathbf{C}}_y$ and $\hat{\mathbf{C}}_x$ denote other matrix square roots of Σ_y and Σ_x. Then $\hat{\mathbf{C}}_y = \mathbf{C}_y\mathbf{R}_y$ and $\hat{\mathbf{C}}_x = \mathbf{C}_x\mathbf{R}_x$, where $\mathbf{R}_y$ and $\mathbf{R}_x$ are orthonormal matrices. From the analysis above, this yields $\hat{\Sigma}_{xy} = \hat{\mathbf{C}}_x\hat{\mathbf{V}}\hat{\mathbf{U}}'\hat{\mathbf{C}}_y'$, where $\hat{\mathbf{U}}\hat{\mathbf{S}}\hat{\mathbf{V}}'$ is the singular value decomposition of $\hat{\mathbf{C}}_y'\mathbf{W}\hat{\mathbf{C}}_x$. By inspection, $\hat{\mathbf{U}} = \mathbf{R}_y'\mathbf{U}$, $\hat{\mathbf{S}} = \mathbf{S}$, and $\hat{\mathbf{V}} = \mathbf{R}_x'\mathbf{V}$, so that $\hat{\Sigma}_{xy} = \Sigma_{xy}$.

References

Akaike, H. "Information Theory and an Extension of the Maximum Likelihood Principle." In *Proceedings of the Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki. Budapest: Akademiai Kaido, 1973.

Altug, Sumru. "Time-to-Build and Aggregate Fluctuations: Some New Evidence." *Internat. Econ. Rev.* 30 (November 1989): 889–920.

Brillinger, David R. *Time Series: Data Analysis and Theory.* San Francisco: Holden-Day, 1981.

Campbell, John Y., and Shiller, Robert J. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Rev. Financial Studies* 1, no. 3 (1988): 195–228.

Christiano, Lawrence J. "Why Does Inventory Investment Fluctuate So Much?" *J. Monetary Econ.* 21 (March/May 1988): 247–80.

Christiano, Lawrence J., and Eichenbaum, Martin. "Unit Roots in Real GNP: Do We Know and Do We Care?" *Carnegie-Rochester Conf. Ser. Public Policy* 32 (Spring 1990): 7–61.

———. "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations." *A.E.R.* 82 (June 1992): 430–50.

Cochrane, John H. "Explaining the Variance of Price-Dividend Ratios." *Rev. Financial Studies* 5, no. 2 (1992): 243–80.

Durlauf, Steven N., and Hall, Robert E. "Measuring Noise in Stock Prices." Manuscript. Stanford, Calif.: Stanford Univ., 1989.

Geweke, John. "The Dynamic Factor Analysis of Economic Time-Series Models." In *Latent Variables in Socio-economic Models*, edited by Dennis J. Aigner and Arthur S. Goldberger. Amsterdam: North-Holland, 1977.

Haavelmo, Trygve. "The Probability Approach in Econometrics." *Economet-rica* 12 (suppl.; July 1944): 1–115.

Hansen, Gary D. "Indivisible Labor and the Business Cycle." *J. Monetary Econ.* 16 (November 1985): 309–27.

Hansen, Gary D., and Sargent, Thomas J. "Straight Time and Overtime in Equilibrium." *J. Monetary Econ.* 21 (March/May 1988): 281–308.

Hansen, Lars Peter, and Jagannathan, Ravi. "Implications of Security Market Data for Models of Dynamic Economies." *J.P.E.* 99 (April 1991): 255–62.

Hansen, Lars Peter, and Sargent, Thomas J. "Formulating and Estimating Dynamic Linear Rational Expectations Models." *J. Econ. Dynamics and Con-trol* 2 (February 1980): 7–46.

Hansen, Lars Peter, and Singleton, Kenneth J. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econo-metrica* 50 (September 1982): 1269–86.

Hodrick, Robert J., and Prescott, Edward C. "Post-War U.S. Business Cycles: An Empirical Investigation." Manuscript. Pittsburgh: Carnegie Mellon Univ., 1980.

Howrey, E. Philip. "Stochastic Properties of the Klein-Goldberger Model." *Econometrica* 39 (January 1971): 73–87.

———. "Dynamic Properties of a Condensed Version of the Wharton Model." In *Econometric Models of Cyclical Behavior*, vol. 2, edited by Bert G. Hickman. New York: Columbia Univ. Press (for NBER), 1972.

King, Robert G.; Plosser, Charles I.; and Rebelo, Sergio T. "Production, Growth, and Business Cycles: Technical Appendix." Manuscript. Roches-ter, N.Y.: Univ. Rochester, 1987.

———. "Production, Growth and Business Cycles: I. The Basic Neoclassical Model." *J. Monetary Econ.* 21 (March/May 1988): 195–232. (*a*)

———. "Production, Growth and Business Cycles: II. New Directions." *J. Monetary Econ.* 21 (March/May 1988): 309–41. (*b*)

King, Robert G.; Plosser, Charles I.; Stock, James H.; and Watson, Mark W. "Stochastic Trends and Economic Fluctuations." *A.E.R.* 81 (September 1991): 819–40.

King, Robert G., and Rebelo, Sergio T. "Low Frequency Filtering and Real Business Cycle Models." *J. Econ. Dynamics and Control* (1993), in press.

Kydland, Finn E., and Prescott, Edward C. "Time to Build and Aggregate Fluctuations." *Econometrica* 50 (November 1982): 1345–70.

McCallum, Bennett T. "Real Business Cycle Models." In *Modern Business Cycle Theory*, edited by Robert J. Barro. Cambridge, Mass.: Harvard Univ. Press, 1989.

McFadden, Daniel. "Econometric Models of Probabilistic Choice." In *Struc-tural Analysis of Discrete Data with Econometric Applications*, edited by Charles Manski and Daniel McFadden. Cambridge, Mass.: MIT Press, 1981.

Mankiw, N. Gregory. "Real Business Cycles: A New Keynesian Perspective." *J. Econ. Perspectives* 3 (Summer 1989): 79–90.

Mankiw, N. Gregory; Runkle, David E.; and Shapiro, Matthew D. "Are Pre-liminary Announcements of the Money Stock Rational Forecasts?" *J. Mone-tary Econ.* 14 (July 1984): 15–27.

Mankiw, N. Gregory, and Shapiro, Matthew D. "News or Noise: An Analysis of GNP Revisions." *Survey Current Bus.* 66 (May 1986): 20–25.

Manski, Charles F. *Analog Estimation Methods in Econometrics*. New York: Chap-man and Hall, 1988.

Plosser, Charles I. "Understanding Real Business Cycle Models." *J. Econ. Perspectives* 3 (Summer 1989): 51–77.

Prescott, Edward C. "Theory Ahead of Business-Cycle Measurement." *Carnegie-Rochester Conf. Ser. Public Policy* 25 (Autumn 1986): 11–44.

Rotemberg, Julio J., and Woodford, Michael. "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity." *J.P.E.* 100 (December 1992): 1153–1207.

Sargent, Thomas J. "Two Models of Measurements and the Investment Accelerator." *J.P.E.* 97 (April 1989): 251–87.

Sargent, Thomas J., and Sims, Christopher A. "Business Cycle Modeling without Pretending to Have Too Much *a Priori* Economic Theory." In *New Methods in Business Cycle Research,* by Christopher A. Sims et al. Minneapolis: Fed. Reserve Bank, 1977.

Singleton, Kenneth J. "Econometric Issues in the Analysis of Equilibrium Business Cycle Models." *J. Monetary Econ.* 21 (March/May 1988): 361–86.

Soderlind, Paul. "Cyclical Properties of a Real Business Cycle Model." Manuscript. Princeton, N.J.: Princeton Univ., Dept. Econ., 1993.

Summers, Lawrence H. "Some Skeptical Observations on Real Business Cycle Theory." *Fed. Reserve Bank Minneapolis Q. Rev.* 10 (Fall 1986): 23–27.

Taylor, John B., and Uhlig, H. "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods." *J. Bus. and Econ. Statis.* 8 (January 1990): 1–17.

White, Halbert. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50 (January 1982): 1–25.

Woodford, Michael. "Stationary Sunspot Equilibria: The Case of Small Fluctuations around a Deterministic Steady State." Manuscript. Chicago: Univ. Chicago, 1986.