Generalized Spectral Estimation of the Consumption-Based Asset Pricing Model

Jeremy Berkowitz
University of California, Irvine

(FORTHCOMING, JOURNAL OF ECONOMETRICS)

February 12, 2001

Address correspondence to:

Jeremy Berkowitz
Grad. School of Management
University of California, Irvine
Washington, D.C. 92697
jberkowitz@gsm.uci.edu

Abstract: This paper provides a framework for estimating parameters of a wide class of dynamic rational expectations models in the frequency domain. The approach is particularly useful for models that are meant to match the data only in limited ways. Specifically, this holds when interest is focused on a subset of frequencies. The estimation strategy generalizes band spectrum regression to more general settings and allows for more general loss functions. A noteworthy special case involves whitening estimators which force estimated innovations to be close to white noise. These estimators can be understood as minimizing a set of moment conditions which grows with sample size and which are typically implied by RE models.

Key Words: Estimation, Frequency Domain, Misspecification
JEL Classification: C13, C22

Acknowledgements: I thank Timothy Cogley, Valentina Corradi, Frank Diebold, Jin Hahn, Roberto Mariano, Masao Ogaki and Lee Ohanian for helpful suggestions. Any inaccuracies are mine.

1. Introduction

In recent years, there has been a resurgence of interest in evaluating economic models in the frequency domain (e.g., Canova (1993), Hansen and Sargent (1993), Watson (1993), Cogley (1995), King and Watson (1995), Diebold, Ohanian and Berkowitz (1998)). These methods allow models to be evaluated over subsets of frequencies, such as business cycles, seasonal frequencies or long horizons. Rather than simply rejecting or accepting the empirical fit of a particular model, researchers can constructively examine the frequencies along which the model does relatively poorly or well (see Eichenbaum (1996)).

Spectral methods capable of estimating coefficients in linear regression models date back to the frequency domain regression of Hannan (1963). His approach was extended and applied to econometrics by Engle (1974, 1980) who discusses estimation of linear regression models in the context of measurement error and seasonality.

More recently, Diebold, Ohanian and Berkowitz (1998) present a general strategy for estimating parameters of dynamic rational expectations models. Parameters are estimated by minimizing distance between spectra of observed data and model-generated data. Distance may be defined by the user and may focus on any relevant subset of frequencies. A shortcoming of this framework is that, in all but the simplest cases, the model must be approximated and simulated in order to carry out the estimation. To the extent that the approximate solution differs from the true model, the Diebold, Ohanian and Berkowitz (1998) estimated parameters will differ from the parameters which minimize loss.

In this paper, I formulate an alternative estimation procedure which generalizes Engle's (1974) band spectrum regression (BSR) and does not require an approximate solution of the model. Broadly speaking, I suggest frequency domain extremum estimators which minimize a function of spectra of Euler residuals.

Particular attention is paid to whitening estimators which minimize unexplained dynamics in model residuals. If a model's Euler residuals are constrained to be white noise, they have a convenient representation in the frequency domain which is readily transformed into a minimization criterion. This approach bears a close resemblance to generalized method of moments estimators based on a set of instruments that expands with sample size (see Hansen and Sargent (1982), Hayashi and Sims (1983) and Hansen (1985)).

Band-restricted estimation is appealing when the model of interest is understood to be misspecified.[1] Relative to much existing work, the present approach gives greater emphasis to *estimating* models over subsets of frequency ranges. For example, in the context of consumption-based asset pricing models (CCAPM) the model is typically calibrated or estimated using traditional methods (e.g., Hansen and Sargent (1982), Ferson and Constantinides (1991), Hansen and Singleton (1996)). Model *evaluation* then proceeds in the time domain, frequency domain or both. But this parameter configuration fits the model to all frequencies in observed data -- rather than those frequencies of most interest. In this sense, the traditional approach may ask more of model-fit than is

---

[1]In the linear context considered by Engle (1974), it is possible to write a restricted version of the regression in the frequency domain, in which the original model holds only over a subset of frequencies. In this case, since the model is *defined* in the frequency domain, band-restricted estimation is not necessarily associated with misspecification.

actually intended.

To illustrate the approach, I estimate the CCAPM under the assumption of constant relative risk aversion. As more high frequency fluctuations are included in estimation, I find that the coefficient of risk aversion increases. This finding is consistent with the conjecture that implausible parameter values are caused by high-frequency noise in observed data.

The remainder of the paper proceeds as follows. Section 2 presents the basic framework and generalized spectral estimation. Section 3 describes existing frequency domain estimators. In section 4, I study the role of GSE in minimizing a measure of model misspecification. Section 5 illustrates GSE by estimating the consumption-based asset pricing model and section 6 concludes.

2. Framework

The Euler equation implied by a typical rational expectations model can be written as

$$\textbf{(1)} \qquad \mathrm{E}\left(g(y_t, q_0) \mid \Omega_t\right) = 0,$$

where g(A,A) is a function given by model's first order conditions, y_t is a typical element of the Tx1 vector of observable data, θ_o is a vector of parameter values, and Ω_t is the s -algebra defined by the agent's time t information set. Equation 1 states that the Euler residual, $g(y_t, \theta_o)$, has a zero conditional mean. It implies that for any rx1 instrument x_{t-1} in the agent's time t information set,

$$\textbf{(2)} \qquad \mathrm{E}\left(g(y_t, \theta_0) \otimes \begin{pmatrix} 1 \\ x_{t-1} \end{pmatrix} \right) = 0,$$

the familiar basis for GMM estimation. Each element of x_{t-1} induces a stochastic process with mean zero, reflecting that economic agents should be unable to forecast model errors.

The instrument x_{t-1} is typically composed of lags of the endogenous variables, y_t. However, lags of the Euler residuals $g(y_t, \theta_o)$ are also valid and lead to some very interesting results. The motivation for viewing lagged *residuals* as instruments is that this choice corresponds to the familiar statement that the Euler residuals form a martingale difference sequence.

Rewriting equation 2, we have that

$$\textbf{(3)} \qquad \mathrm{E}(g(y_t, \theta_0) \cdot g(y_{t-\tau}, \theta_0)) = 0,\ \tau = 1,2,\ldots$$

Let $g_t(\theta) = g(y_t, \theta)$ and denote the associated spectral density $f_{g_t(\theta)}(\omega)$. Equation 3 is equivalent to the statement,

$$f_{g_t(\theta_0)}(\omega) = k,\ \ \forall \omega \in [0, 2\pi].$$

The spectral density of $g_t(\theta_o)$ is constant over its entire support.[2] Equivalently,

[2]The spectral density function is defined as $f(\omega_j) = \sum_{\tau=-\infty}^{\tau=\infty} e^{i\omega_j\tau} \gamma(\tau)/2\pi$, where $\gamma(\tau)$ are the autocovariances at lag τ. Consistent estimates can be obtained by averaging each periodogram ordinate with neighboring ordinates, $\hat{f}(\omega_j) = \sum_{\omega} I(\omega) k(\omega - \omega_j)$, where $k(\cdot)$ is a suitable weighting function and

(4) $$F_{g_t(q_0)}(l) = \int_0^l \frac{1}{s^2} f_{g_t(q_0)}(w)\, dw = \frac{l}{p},$$

the spectral distribution function is a 45° line. Given a finite realization of length T, consistent estimates of the spectral distribution are easily calculated such that $\hat{F}_{g_T}(w) \rightarrow \frac{w}{p}$.

This suggests a class of GSE estimators which make the residuals as close as possible to a martingale difference sequence. The martingale hypothesis has a long history in economics and, as a result, a number of associated test statistics have emerged. Durlauf (1991) unifies and derives the asymptotic properties of a wide class of tests of the martingale hypothesis. The test statistics discussed in Durlauf (1991) are readily inverted into minimization criteria that can be used for estimation.

Definition: The generalized spectral estimator is defined as:

$$\hat{q}_{GSE} = \underset{q}{\operatorname{argmin}} \sum_{?_T \in v} C_w \left(\hat{F}_{g(q)}(w) \right),$$

where $C_\omega(\cdot)$ may include a wide variety of appropriate loss functions that may vary with the frequency ω. The possible frequencies, $\{?_T \in v\}$, are now indexed by T to acknowledge that given a finite set of observations, a maximum of T/2 distinct frequencies can be estimated. We may appropriately think of

---

I(ω) is the periodogram. Since the spectral density is symmetric about frequency π, we henceforth limit attention to the range [0,π]. See, for example, Brillinger(1975).

the $\hat{F}_{g_T(q)}(w_m)$ as a triangular array with $?_m = \frac{2p\,m}{T}, m = 0, \ldots, \frac{T}{2} - 1$. The sample spectrum at a given frequency is indexed by sample size, but the number of estimable frequencies also increases with sample size. For what follows, I consider only univariate $g_T(\theta)$ for notational simplicity.

Generalized spectral estimation permits several conveniences relative to existing estimators. First, in the context of DSGE models the data generating process (policy function) is not typically known analytically. The innovation process $\{g_t(\theta)\}_1^T$ must first be simulated and then the rest of the model data must be generated through an approximate policy function. Maximum likelihood as well as the estimators discussed in Diebold, Ohanian and Berkowitz (1998) require an additional layer of approximation.

GSE, like GMM, does not require knowledge of policy functions. Parameters can be estimated using only observed data. Moreover, GSE utilizes a set of instruments which correspond to an economically interesting loss function. When studying representative agent models, we expect model errors to be unforecastable.

By defining the minimand in the frequency domain, the loss function can be restricted to contain only frequencies of interest to the user. Sims (1972, 1993) and Hansen and Sargent (1993), for example, discuss the utility of ignoring seasonal frequencies in the context of models which are not designed to explain the seasonality.[3] Sims (1993) demonstrates that reducing spectral power over seasonal frequencies can facilitate accurate estimates of model behavior over other frequencies. Hansen

[3]Sims (1972, 1993) and Hansen and Sargent (1993) study different estimators (different cost functions) than those of interest here.

and Sargent (1993) note that in some cases deleting seasonal frequencies can reduce asymptotic bias when the model is misspecified.

2.1 Whitening Estimators

It will be convenient to begin by defining the partial sums, $\mathrm{U}(s)=\int_0^{\pi s}\left(f_{\gamma_T}(\omega)/\sigma^2_{\gamma_T} - \frac{1}{\pi}\right)d\omega$,

and the sample counterparts, $\hat{\mathrm{U}}(s)=\int_0^{\pi s}\left(\hat{f}_{\gamma_T}(\omega)/\hat{\sigma}^2_{\gamma_T} - \frac{1}{\pi}\right)d\omega$, $s\in(0,1)$. From equation 4, $\hat{\mathrm{U}}(s)$ has zero expectation when γ_t is a martingale difference sequence. In order to construct estimators based on the deviations, it remains only to specify a convex loss function. For example, the L^2 metric yields,

$$\hat{\theta} = \arg\min_{\theta} \int_0^1 \hat{\mathrm{U}}(s)^2\, ds .$$

This estimator minimizes the Cramér-Von Mises (CVM) test statistic described in Durlauf (1991).

The "moment conditions" associated with this estimator are $E\left(\int_0^{\pi s}(\hat{f}_{\gamma}(\omega)/\hat{\sigma}^2_{\gamma}-1/\pi)\,d\omega\right)=0$ for $s\in(0,1)$; the spectral distribution function should be a straight line. Interestingly, these moments are equivalent to the conditions,

$$\textbf{(5)} \qquad E\,\frac{1}{p}\sum_{j=1}^{\infty}\hat{r}(j)\sin(j\,p\,s)/j = 0,\ s \in (0,1)$$

where $\hat{r}(j)$ is the sample autocorrelation function.[4]

This equivalence makes clear that the CVM estimator could be formulated as a GMM estimator. In particular, consider the moment conditions $E\,\hat{r}(j)=0$, for j=1,..,T and a weighting matrix with typical element $[W]_{mj} = \frac{1}{p\,j}\sin(?_m j)$, $?_m = p\,s_m/T$, s =1,...,T and j=1,...,T. As $T\rightarrow\infty$, the first order condition associated with GMM converges to that of CVM. The equivalent time-domain estimator sets the autocorrelations of residuals as close as possible to zero.

Of course, an infinite number of valid functions, C(⋅), are uniquely minimized when g_t is a martingale. Durlauf (1991) considers a number of metrics that penalize deviations of the cumulative periodogram. The same metrics can be used to define estimators. For example, we could invert the Kolmogorov-Smirnov statistic,

$$\hat{q} = \arg\min_{q}\ \sup_{s}\ |\,\hat{U}(s)\,|\,.$$

Such an estimator penalizes the maximal deviation across frequencies.

For estimators of the form, $\arg\min_{q}\ C(\hat{U}(s))$, the periodograms can be used in place of

[4]See Grenander and Rosenblatt (1957) for a formal derivation.

consistent estimates of the spectrum. As emphasized by Durlauf (1991), the cumulated deviations of periodogram ordinates will converge to the cumulated deviations of spectra due to the law of large numbers which arises from averaging.

2.2 Asymptotic Behavior

This section delineates sufficient conditions for the consistency of whitening estimators when the model being estimated is correctly specified. It is of interest to verify that our estimation procedure would be asymptotically valid if we did, in fact, know the true model. This allows the estimators to be understood in terms of standard asymptotic criteria rather than depending on misspecification for meaning.

We impose the following regularity conditions on the data-generating process*:*

Assumption B1: The unknown parameter vector that we wish to estimate θ_0 is an element of Θ, a compact subset of R^k.

Assumption B2: $\{g(y_t,\theta):t=1,2,...\}$ is a sequence of random variables defined on a probability space (O, õ, P) for each $\theta\in\Theta$ and $g(y_t,\theta)$ is continuous in θ for all real y_t.

The function g(A,A) is given by the first-order conditions of the econometric model.

Assumption B3: $g(y_t,\theta)$ is strong mixing of size -r/(r-1) for some r>1, and possesses a finite variance, for all t, all $\theta\in\Theta$.

Assumption B4: The sequence $g(y_t,\theta)$ has unconditional mean zero and a continuous spectral density function on $\theta\in\Theta$.

A sufficient condition for $g(y_t,\theta)$ to possess continuous spectra is that it have absolutely summable autocovariances (e.g. Priestley (1981)).

Assumption B5: There exists a unique θ_0 such that the spectral density function of $g(y_t,\theta_0)$ is a constant.

This is the identification condition underlying whitening estimators. It is the GSE analog of Hansen's (1982) requirement that the population moments implied by the model have a unique zero. This is not overly restrictive in the context of many dynamic models. However, some time series models such as ARCH models must be handled with care. Consider for example, an AR(1) with ARCH innovations. Even under correct specification, the innovations are not uniquely white noise for any set of ARCH parameters. The parameters may be estimated, however, by noting that the squared innovations have a conditional homoskedastic ARMA representation.

Assumption B6: $E(g(y_t,\theta))^4$ are finite for all t, all θ.

Assumption B6 is a standard condition needed for establishing the asymptotic properties of sample spectral distributions.

Three additional assumptions are made to restrict the class of permissible estimators:

Assumption B7: $C(\cdot)$ is continuous.

Assumption B8: The loss function, $C(\cdot)$, must have the property that $\underset{q}{\arg\min} \int_v C\left(U_{g(q)}(s)\right) ds = \theta_0$ with $0 \leq \int_v C\left(U_{g(q)}(s)\right) ds \leq \infty$.

Assumption B8 is an identification condition on the cost function. If C is quadratic, for example, as in GMM, the condition is trivially satisfied. We can trivially verify the assumption for the CVM estimator since

$$\int C(F_{g(q)}(w))\,dw = \int \left(p\,F_{g(q)}(w) - w\right)^2 dw$$

which achieves a minimum at $p\ \mathrm{F}_{g(q)}(w) = w$ and this holds only for θ_0.

Assumption B9: $\mathrm{EC}(\hat{\mathrm{U}}_{g_T(q)}(s))^4$ are finite for all $\theta \in \Theta$, $s \in (0,1)$, $T \geq 1$.

If the estimated model is correctly specified, we can now show convergence in probability.

Theorem 1(Consistency): Let $\hat{q}_T = \arg\min_{q} \frac{1}{T} \sum_{s \in v} \mathrm{C}\left(\hat{\mathrm{U}}_{g(q)}(s)\right)$. *Under assumptions B1-B9,* $\hat{q}_T \rightarrow q_0$ in probability.

Proof: See appendix.

With two additional assumptions, it is possible to show that the generalized spectrum estimator is asymptotically Normally distributed. The estimator's asymptotic variance will be seen to depend on the loss function C(·) as well as the underlying data generating process.

Assumption C1: C(·) is twice continuously differentiable.

Assumption C2: Let $\lim \mathrm{T}^{-1}(\partial \mathrm{Q}_T(q)/\partial q\,|_{q_0})(\partial \mathrm{Q}_T(q)/\partial q'|_{q_0}) = \Sigma$ be a finite nonstochastic matrix, where $\mathrm{Q}_T(q) = \sum_{s \in v} \mathrm{C}\left(\hat{\mathrm{U}}_{g(q)}(s)\right)$.

Theorem 2: Under C1-C2 and the assumptions of Theorem 1,

$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, H(\theta_0)^{-1}\Sigma H(\theta_0)^{-1})$, where $H(\theta_0)$ *is the Hessian of the minimand.*

Proof: See appendix

3. Existing Frequency Domain Estimators

This section describes the relationship between GSE, Engle's (1974) band spectrum regression and maximum likelihood estimation. Consider the fixed distributed lag model,

$$y_t = \sum_{j=0}^{\infty} \theta_j x_{t-j} + u_t, \tag{6}$$

where y_t is a scalar, x_t a vector and the coefficients θ_j are parameters to be estimated. We can stack the observations to write Y=θX+U. Let W be the Fourier matrix with typical element

$[W]_{mj} = \frac{1}{\sqrt{2\pi T}}[e^{i\omega_{m-1} j}]$, $\omega_m = \frac{2\pi m}{T}$, m=0,..., T-1 and j=1,...,T. Premultiplication of a matrix of data by W yields the Fourier transform of the data.

Multiplying both sides of the equation, WY = Wxθ + Wu or $\tilde{Y} = \tilde{X}\theta + \tilde{u}$, where $\tilde{Y}$ and $\tilde{X}$ are Fourier transformed data. A restricted version of the model in which some frequencies are ignored may then be written as $A\tilde{Y} = A\tilde{X}\theta + A\tilde{u}$, where the TxT matrix A contains suitable zeros or ones on its diagonal.

Band spectrum regression sets $\hat{\theta}_{\mathrm{BSR}} = (\tilde{\mathrm{X}}'\mathrm{A}'\mathrm{A}\tilde{\mathrm{X}})^{-1}\tilde{\mathrm{X}}'\mathrm{A}'\mathrm{A}\tilde{\mathrm{Y}}$. This is simply the least squares estimator of the Fourier transformed data. Band spectrum regression can equivalently be written as

$$\textbf{(7)} \qquad \hat{\theta}_{\mathrm{BSR}} = \arg\min_{\theta} \sum_{\omega \in \varpi} |\, \tilde{\mathrm{u}}_{\omega}(\theta) \,|^2,$$

where $| A |^2$ denotes the conjugate square and ϖ denotes the set of frequencies which are included in estimation. But equation 7 can be further rewritten as

$$\textbf{(8)} \qquad \hat{\theta}_{\mathrm{BSR}} = \arg\min_{\theta} \sum_{\omega \in \varpi} \mathrm{I}_{\mathrm{u}(\theta)}(\omega),$$

where $\mathrm{I}_{\mathrm{u}(\theta)}(\omega)$ is the periodogram of the residual $u_t(\theta)$.

The frequency domain approximation to maximum likelihood estimation can also be written in this framework. It is,

$$\textbf{(9)} \qquad \hat{\theta}_{\mathrm{MLE}} = \arg\min_{\theta} \sum_{\omega} (\log \hat{\mathrm{f}}_{\mathrm{u}(\theta)}(\omega) + \hat{\mathrm{f}}^{-1}_{\mathrm{u}(\theta)}(\omega)\mathrm{I}_{\mathrm{u}(\theta)}(\omega)),$$

where $\hat{\mathrm{f}}_{\mathrm{u}(\theta)}(\omega)$ is a consistent estimate of the spectral density of $u_t(\theta)$. Under normality and circularity of the residual this estimator can be derived from the familiar time domain likelihood function (see, for example, Harvey (1989)).[5] Diebold, Ohanian and Berkowitz (1998) advocate band-restricted maximum likelihood, which differs from maximum likelihood by taking the summation over only those frequencies of foremost interest,

$$\textbf{(10)} \qquad \hat{q} = \underset{q}{\operatorname{argmin}} \sum_{w \in v} (\log \hat{\mathrm{f}}_{\mathrm{u}(q)}(w) + \hat{\mathrm{f}}^{-1}_{\mathrm{u}(q)}(w) \mathrm{I}_{\mathrm{u}(q)}(w)),$$

If, for example, we would like to focus on fluctuations of business cycle frequency we might restrict ϖ = [$\pi/16$, $\pi/4$]. Using quarterly data, this band would isolate cycles of length between 2 and 8 years.

4. Minimization of Specification Error

In the presence of some forms of misspecification, Sims (1972) shows that least squares estimation is equivalent to choosing the parameter vector which minimizes a measure of specification error. In this section, GSE is shown to minimizing a somewhat different metric of misspecification than that of Sims. Again, consider the fixed distributed lag model,

$y_t = \Sigma\theta_j \, x_{t-j} + u_t$. Assume that x_{t-1} is orthogonal to the innovations, u_t . In this context, misspecification may be finite truncation of the infinite sum or approximation of the true lag structure by a ratio of

[5]A process is said to be circular if its autocovariance matrix has the form of a circulant. A circulant has the property that $\gamma(\tau)=\gamma(T-\tau)$, for τ = 1,...,T-1. Circularity does not hold in infinite moving average processes, in general, but even without circularity Whittle's derivation holds asymptotically.

polynomials. The discussion is limited to GSE estimators of the form, $\mathrm{argmin}_\theta\, C(\hat{U}_\theta(s))$, where C(·) is a sensible convex function and, as before,

$$\hat{U}_\theta(s) = \int_0^{\pi s}\left(\frac{\hat{f}_{u(\theta)}(\omega)}{\hat{\sigma}^2_{u(\theta)}} - \frac{1}{\pi}\right)d\omega. \qquad \textbf{(11)}$$

It is apparent that for an arbitrary parameter vector, $u_t(\theta)$, the implied residual can be written as $u_t(\theta_0)$ - $(\theta_0(L) - \theta(L))x_{t-1}$ .This immediately implies that $f_{u(\theta)}(\omega) = f_{u(\theta_0)}(\omega) + f_x(\omega)|\theta_0(e^{i\omega}) - \theta_1(e^{i\omega})|^2$. The spectrum of estimated residuals equals the spectrum of the true innovation plus a misspecification term. Misspecification is manifested in the frequency domain as the spectrum of the data, weighted by the transfer function of the model approximation error. Integrating the relation over frequencies [0,π] gives,

$$\sigma^2_{u(\theta)} = \sigma^2_{u(\theta_0)} + \int f_x(\omega)|\theta_0(e^{i\omega}) - \theta_1(e^{i\omega})|^2 d\omega \qquad \textbf{(12)}$$

Sims (1972) shows that the least squares estimator, when it exists, minimizes the quantity $\int f_x(\omega)|\theta_0(e^{i\omega}) - \theta_1(e^{i\omega})|^2 d\omega$. This motivates Sims(1993) argument that if misspecification is concentrated at specific frequencies, such as seasonals, it makes sense to pre-filter the data in such a way as to reduce the power of $f_x(\omega)$ at those frequencies. Hansen and Sargent (1993) note the asymmetric advantages of doing so: if the model is correctly specified consistency is preserved despite deletion of seasonal frequencies; if the model is misspecified deletion can reduce asymptotic bias.

A similar measure of misspecification is implicitly minimized by GSE's. The minimand of GSE is by definition a function of,

$$\textbf{(13)} \qquad \int_0^{p\,s} \left(\frac{\mathrm{f}_{\mathrm{u}(q_0)}(w) + \mathrm{m}(w)}{s^2_{\mathrm{u}(q_0)} + \int \mathrm{m}(w)\,\mathrm{d}w} - \frac{1}{p} \right) \mathrm{d}w,$$

where $\mathrm{m}(w) = \mathrm{f}_{\mathrm{x}}(w)\,|q_0(\mathrm{e}^{\mathrm{i}w}) - q_1(\mathrm{e}^{\mathrm{i}w})|^2$ for notational convenience. Rearranging terms,

$$\int_0^{p\,s} \left(f_{u(q_0)}(w) + m(w) - \frac{s^2_{u(q_0)} + \int m(w)dw}{p} \right) \left[s^2_{u(q_0)} + \int m(w)dw \right]^{-1} dw$$

$$\textbf{(14)} \qquad = (\mathrm{U(s)} + \mathrm{B(s)})\,\mathrm{n}^{-1}$$

where U(s) is the usual deviation, $\int_0^{p\,s} \left(\mathrm{f}_{\mathrm{u}(q_0)}(w) - \frac{s^2_{\mathrm{u}(q_0)}}{p} \right) \mathrm{d}w,$

$$\textbf{(15)} \qquad \mathrm{B(s)} = \int_0^{p\,s} \mathrm{m}(w)\,\mathrm{d}w - \mathrm{s} \int_0^{p} \mathrm{m}(w)\,\mathrm{d}w$$

and n is a normalization term, $\mathrm{n} = \sigma^2_{u(\theta_0)} + \int \mathrm{m}(\omega)d\omega$. The term U(s) is only a function of population parameters, leaving GSE to minimize a function of B(s)n^{-1}. The basic building block of misspecification is $\mathrm{m}(\omega) = f_x(\omega)|\theta_0(e^{i\omega}) - \theta_1(e^{i\omega})|^2$ exactly as in Sims (1972).

Unlike least squares, however, GSE forces the spectrum of m(ω) to be as flat as possible rather than minimizing its integral. This presents an argument for pre-filtering the data (reducing $f_x(\omega)$) over misspecified frequencies only to the extent that there are *spectral peaks.* Suppose that seasonal fluctuations display greater spectral power in some range than neighboring frequencies. Then B(s) is minimized by reducing the power over these frequencies to the average spectral power. Unlike with least squares, it is not necessary in our context to filter the data any further and, in doing so, create a trough in the spectrum.

5. The Consumption-Based Asset Pricing Model

The landmark consumption-based asset pricing model of Lucas (1978) and Breeden (1979) has been subjected to innumerable empirical examinations. Such studies have, for the most part, cast doubt on the model's consistency with observed data. The calibration exercises of Mehra and Prescott (1985) suggest implausibly high risk aversion, while formal testing procedures lead to convincing rejections of the model (e.g., Hansen and Singleton (1982, 1983)). In a similar vein, Cochrane and Hansen (1992), Burnside (1994) and Cecchetti, Lam and Mark (1994) show that the CCAPM implies a stochastic discount factor that violates the Hansen-Jagannathan volatility bounds.

In response, several authors have suggested that model rejections can be at least partly attributed to unmodeled frictions or measurement error. Grossman, Melino and Shiller (1987) consider the implications of time-aggregation in observed consumption data, while Wheatley (1988) studies the effect of measurement error in consumption. In both cases, the result is to introduce short-horizon predictability into the pricing errors. For this reason, Breeden, Gibbons and Litzenberger (1989) argue that at short-horizons one should replace consumption with a portfolio whose return is highly correlated with longer-run consumption movements. Campbell and Mankiw (1990) argue that consumption-based models are not adequate even over the long-run because of liquidity constrained consumers.

The frequency domain presents a natural framework for studying whether the CCAPM attains a better fit over long-horizons. To address this question, we estimate a simple version of the CCAPM over various frequency ranges.

Assume perfect asset markets and homogenous agents who maximize the expected time-separable utility of consumption,

$$\textbf{(16)} \qquad E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t} U(C_t)$$

where ß is a constant discount factor, C_t is time-t consumption and the expectation is conditioned on information available up to time t_0. The representative agent is assumed to display constant relative risk aversion. Along with the usual budget constraint, this implies the familiar first-order condition:

(17) $$E_t\left[\beta R_{t+1}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]=1$$

where R_{t+1} is the one-period real return on unconsumed wealth and γ is the coefficients of risk aversion. Typically, ß is estimated to be near 1 and γ between -5 and 2. However, calibration exercises, such as Campbell, Lo and MacKinlay (1997), suggest that γ is closer to 20. That γ must be so large is essentially a statement of the equity premium puzzle.

Using GSE, we can estimate the model over long horizons or any subset of frequencies we desire. To do so, I collected quarterly data from January 1947 to December 1998 for the U.S. Asset returns are constructed from the *Center for Research in Securities Prices* (CRSP) value-weighted and equally-weighted NYSE indexes. Consumption and price data are from the *Survey of Current Business*. Expenditure on nondurables and nondurables plus services are converted into constant dollars and then into per-capita terms using data from the Bureau of the Census. Real asset returns are constructed by subtracting inflation rates corresponding to each notion of consumption from nominal returns.

Table 1 presents the estimation results based on the CVM cost function for seven frequency bands, $[0, \omega_{max}]$, $\omega_{max} = \pi/20, \pi/10, ..., \pi$. For both consumption measures and both equity indexes, the discount factor ß decreases slowly from .92 to about .88 as more and more high frequencies are included. At the same time, estimated risk aversion increases sharply as frequencies higher than about .6 (2 1/2 year cycles) are added. This is most easily seen in Figure 1 which summarizes the estimation results graphically. If we view high estimates of risk aversion as a failing of the CCAPM, then Figure 1 suggests greater empirical relevance of the model over low frequencies.

In summary, both free parameters, ß and γ, attain more plausible estimated values when attention is restricted to low-frequencies. This provides compelling evidence that the consumption-based asset pricing model fares poorly because of high frequency noise that it is not capable of nor intended to match.

These results are consistent with Grossman, Melino and Shiller (1987) and Wheatley (1988) who argue that in the presence of aggregation and measurement error the CCAPM should be expected to fail over short horizons.

We have motivated GSE by emphasizing that under rational expectations Euler residuals should behave like white noise. GSE parameters should imply residuals that are closer to white noise than calibrated values. Figure 2 displays the spectra of two series of Euler residuals, alongside a confidence tunnel calculated under the null of white noise residuals. The first spectrum corresponds to the calibration exercise in Campbell, Lo and MacKinlay (1997) who update the approach of Mehra and Prescott (1985) with a more recent data set. The second spectrum is implied by low frequency GSE estimates -- including only frequencies lower than 2 1/2 years. Several features of Figure 2 are noteworthy. First, the calibrated parameter values imply residuals with substantially higher variance; the associated spectrum is everywhere higher than the GSE spectrum. Second, the calibrated spectrum is decidedly downward sloping over long horizons over about 2 years. In fact, the spectrum exceeds a 95% confidence tunnel in this range.

The unusually low discount factor of .92, estimated by GSE, appears to yield residuals with substantially lower persistence (and lower variance) than traditional values. These results convey interesting information about the workings of the CCAPM. It was previously known that large values of

γ imply an equity premium "puzzle". The GSE approach indicates that over long horizons the model is consistent with much lower risk aversion. The cost is a risk-free rate puzzle -- in this case generated by the low discount factor.

6. Conclusion

This paper suggests some new techniques for estimating dynamic rational expectations models which are designed or expected to match only subsets of the fluctuations in observed data. An application to the consumption-based asset pricing model illustrated the utility of excluding high frequencies. Generalized spectral estimators may also be of use in estimating business cycle models or long term growth models because of their inherent focus on subsets of frequencies. It is hoped that our approach strengthens the ability of economists to constructively assess the dimensions along which models succeed and fail to match empirical observations.

References

Amemiya, T. (1985). *Advanced Econometrics*. Oxford: Blackwell.

Andrews, D.W.K. (1987), "Consistency in nonlinear Econometric Models: A Generic Uniform Law of Large Numbers," *Econometrica*, 6, 1465-1471.

Apostol, T. M. (1974). *Mathematical Analysis*. Reading, MA: Addison-Wesley.

Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory*. New York: Holden Day.

Canova, F. (1993), "Detrending and Business Cycle Facts," CEPR Discussion Paper No. 782.

Cogley, T. (1995), "Estimating Dynamic Rational Expectations Models when the Trend Specification is Uncertain," manuscript, Federal Reserve Bank of San Francisco.

Diebold, F. X., L.E. Ohanian, J. Berkowitz (1998), "Dynamic Equilibrium Economies: A Framework for Comparing Models and Data," *Review of Economic Studies*, 65, 433-451.

Domowitz, I. and H. White (1982), "Misspecified Models with Dependent Observations," *Journal of Econometrics*, 20, 35-58.

Durlauf, S. N. (1991), "Spectral Based Testing of the Martingale Hypothesis," *Journal of Econometrics*, 50, 355-376.

Eichenbaum, M. (1996), "Some Comments on the Role of Econometrics in Economic Theory," Economic Perspectives, Federal Reserve Bank of Chicago, 22-31.

Engle, R. F. (1974), "Band Spectrum Regression," *International Economic Review*, 15, 1-11.

Engle, R. F. (1980), "Exact Maximum Likelihood Methods for Dynamic Regressions and Band Spectrum Regressions," *International Economic Review*, 21, 391-407.

Ferson, W. and G. Constantinides (1991), "Habit Formation and Durability in Aggregate Consumption: Empirical Tests," *Journal of Financial Economics*, 29, 199-240.

Grenander, U. and M. Rosenblatt (1957). *Statistical Analysis of Stationary Time Series*. New York: Chelsea.

Hall, P. and C.C. Hyde (1980). *Martingale Limit Theory and Its Application.* New York: Academic Press.

Hannan, E.J. (1963), "Regression for Time Series," in M.Rosenblatt ed., *Time Series Analysis*. New

York: John Wiley.

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moment Estimators," *Econometrica*, 50, 1029-1055.

Hansen, L. P. (1985), "A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators," *Journal of Econometrics*, 30, 203-238.

Hansen, L.P. and T.J. Sargent (1982), "Instrumental Variables Procedures for Estimating Linear Rational Expectations Models," *Journal of Monetary Economics*, 9, 263-296.

Hansen, L.P. and T.J. Sargent (1993), "Seasonality and Approximation Errors in Rational Expectations Models," *Journal of Econometrics*, 55, 21-55.

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

Hayashi, F. and C. Sims (1983), "Nearly Efficient Estimation of Time Series Models with Predetermined, but not Exogenous, Instruments," *Econometrica*, 51, 783-798.

King, R.G. and M.W. Watson (1995), "Money, Prices, Interest Rates and the Business Cycle," Federal Reserve Bank of Chicago Working Paper No. 95-10.

Serfling, R.J. (1968). "Contributions to Central Limit Theory for Dependent Variables," *Annals of Mathematical Statistics*, 39, 1158-1175.

Sims, C.A. (1972), "Approximate Prior Restrictions in Distributed Lag Estimation," *Journal of the American Statistical Association*, 67, 169-175.

Sims, C.A. (1993), Rational Expectations Modeling with Seasonally Adjusted Data," *Journal of Econometrics*, 55, 9-19.

Watson, M.W. (1993), "Measures of Fit for Calibrated Models," *Journal of Political Economy*, 101, 1011-1041.

Wheatley, S. (1988), "Some Tests of the Consumption-Based Asset Pricing Model," *Journal of Monetary Economics*, 22, 193-215.

Technical Appendix

Lemma 1: (Boundedness) The minimand of whitening estimators is uniformly bounded.

Proof

Whitening estimators minimize a convex function of

$$\hat{\mathrm{U}}(\mathrm{s}) = \int_0^{p\,s} \left(\frac{\hat{\mathrm{f}}}{\hat{s}^{\,2}} - \frac{1}{p} \right) \mathrm{d}w$$

$$= \int_0^{p\,s} \left(\frac{\hat{\mathrm{f}}}{\hat{s}^{\,2}} \right) \mathrm{d}w - \int_0^{p\,s} \left(\frac{1}{p} \right) \mathrm{d}w$$

$$= \frac{1}{\hat{s}^{\,2}} \int_0^{p\,s} \hat{\mathrm{f}}\ \mathrm{d}w - \int_0^{p\,s} \left(\frac{1}{p} \right) \mathrm{d}w = \hat{\mathrm{F}}(p\ \mathrm{s}) - \mathrm{s}.$$

The sample spectral distribution, $\hat{\mathrm{F}}(w)$, is maximized at $\omega=\pi$ where $\hat{\mathrm{F}}(w)=1$. Since $0 \le s \le 1$, we conclude that $|\hat{\mathrm{U}}(\mathrm{s})| \le 1$. The convex cost functions, C(·), considered in the present paper are bounded on compact sets implying that $\mathrm{C}(\hat{\mathrm{U}}(\mathrm{s}))$ is itself bounded.[6] ■

Lemma 1 has several important implications. First, it follows that $\mathrm{EC}(\hat{\mathrm{U}}_{\mathrm{T}}(\mathrm{s_i})) = \mathrm{m}(\theta)_{\mathrm{iT}} < \infty$ and $\mathrm{var}(\mathrm{C}(\hat{\mathrm{U}}_{\mathrm{T}}(\mathrm{s_i}))) = \sigma^2(\theta)_{\mathrm{iT}} < \infty$, for any $\theta \in \Theta$ since bounded functions are integrable (we have already assumed a finite measure). The lemma will also be sufficient to establish a Lipschitz condition on $\mathrm{C}(\hat{\mathrm{U}}_{\mathrm{T}}(\mathrm{s_i}))$ needed to prove a uniform LLN.

---

[6] The boundedness of the cost function is implied by the assumption that C(@) is continuous and the fact that {x: |x|# 1} is compact (see Apostol 1974, Theorem 4.27).

Lemma 2: (Chebyshev LLN) The stochastic sum $T^{-1}\sum_{i=1}^{T}\hat{U}_T(i/T)$ *admits a weak law of large numbers, pointwise in q,*

(18) $$\mathrm{pr}_T\left[T^{-1}\sum_{i=1}^{T}\left(C(\hat{U}_T(i/T)) - m_{iT}\right)\geq e\right]\to 0, \text{ for all } e > 0.$$

Proof

We will use Chebyshev's inequality to derive a standard weak LLN for triangular arrays. Write the sum as $T^{-1}\Sigma_i C(\hat{U}_T(s_{2i/T}))=T^{-1}Q_T(\theta)$ for notational convenience.

Then $E\,T^{-1}Q_T(\theta) = T^{-1}\Sigma_i m_{iT} = m_T$. Now $\mathrm{var}(T^{-1}Q_T(\theta)) = \mathrm{var}(T^{-1}\Sigma_i C(\hat{U}_T(s_i))) =$

$$= T^{-2}\Sigma_i s^2_{iT} + T^{-2}\Sigma_{i\neq j}\mathrm{cov}\left(C(\hat{U}_T(s_i)), C(\hat{U}_T(s_j))\right)$$

The first term is $O(T^{-1})$ since the s^2_{iT} are all finite. Now consider the second term. For fixed i,j, $\mathrm{cov}((\hat{F}_T(w_i), \hat{F}_T(w_j)) = O(T^{-1})$ with $\hat{F}_T(w_i)$ asymptotically Normal (e.g., Brillinger (1975), p.168). It follows from a result in Hall and Hyde (1980) that $\mathrm{cov}\left(C(\hat{U}_T(s_i)), C(\hat{U}_T(s_j))\right)$ is at most $O(T^{-1})$.[7]

The sum $\Sigma_{i\neq j}$ contains at most T(T-1) terms, so

$$\left| T^{-2}\Sigma_{i\neq j}\mathrm{cov}\left(C(\hat{U}_T(s_i)), C(\hat{U}_T(s_j))\right)\right| \leq T^{-2}\,T(T-1)\,R_T$$

where $R_T = O(T^{-1})$. Therefore $\left| T^{-2}\Sigma_{i\neq j}\mathrm{cov}\left(C(\hat{U}_T(s_i)), C(\hat{U}_T(s_j))\right)\right| = O(T^{-1})$ and

$$\mathrm{pr}_T\left([T^{-1}Q_T(\theta) - m_T]\geq e\right)\to 0 \text{ for fixed } e \blacksquare$$

Proof of Theorem 1

Boundedness of the minimand (established in Lemma 1) is stronger than assumption A4 in Andrews (1987) and together with our assumptions B1, B3 and B9, we can invoke Andrews' main Theorem. We have

---

[7]Each $C(\hat{U}_T(s_i))$ is measurable with respect to the sigma algebra $\ddot{o}_{iT} = s(\hat{F}_T(w_i))$. It follows from Theorem A.5 in Hall and Hyde that $\mathrm{cov}\left(C(\hat{U}_T(s_i)), C(\hat{U}_T(s_j))\right)$ are of the same magnitude as $\mathrm{cov}(\hat{F}_T(w_i), \hat{F}_T(w_j))$.

(19) $$\sup_{q \in \Theta} | T^{-1} \sum_{i=1}^{T} C(\hat{U}_T(i/T)) - m_T | \to 0$$

almost surely.

The consistency of $\hat{q}_T = \arg\min_{q} T^{-1} \sum C\left(\hat{U}(s)\right)$ follows directly by, for example, Theorem 2.2 in Domowitz and White (1982) or Theorem 4.1.1 in Amemiya (1985).

Proof of Theorem 2

By assumption $\frac{\partial^2}{\partial q^2} Q_T(q)$ exists and is continuous. From the proof of Theorem 1

$$T^{-1} Q_T(q) \xrightarrow{p} Q(q), \text{ uniformly in } q.$$

By assumption C1 and the continuous mapping theorem, it follows that

(20) $$T^{-1} \frac{\partial^2}{\partial q^2} Q_T(q)\Big|_{q_T} \xrightarrow{p} \frac{\partial^2}{\partial q^2} Q(q_0).$$

Now from lemma 1 and assumption C1, $\frac{\partial}{\partial q} C(U_{g(q)}(s))|_{q_0}$ is bounded uniformly in s. This implies that Lindeberg's condition is satisfied for $\frac{\partial}{\partial q} C(U_{g(q)}(s))|_{q_0}$. Given assumption C2 we can apply Liapunov's CLT adapted slightly to accommodate the dependence (see, for example, Billingsley (1986)),

(21) $$\mathrm{T}^{-1/2}\frac{\partial}{\partial q}\mathrm{Q}_{\mathrm{T}}(q)\Big|_{q_0} / \operatorname{var}\left(\mathrm{T}^{-1/2}\frac{\partial}{\partial q}\mathrm{Q}_{\mathrm{T}}(q)\Big|_{q_0}\right) \rightarrow \mathrm{N}(0,1).$$

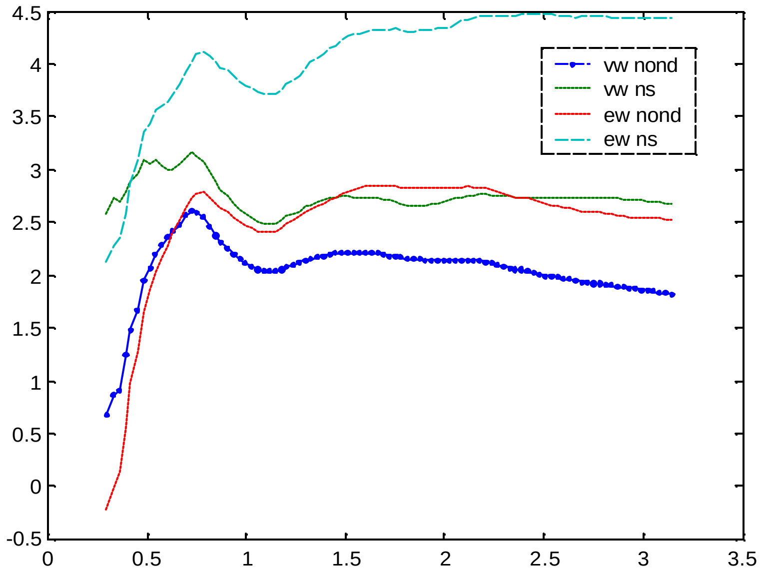
The remainder of the proof follows from standard arguments such as Theorem 4.1.3 in Amemiya (1985).

Table 1. CCAPM Generalized Spectral Estimation Results

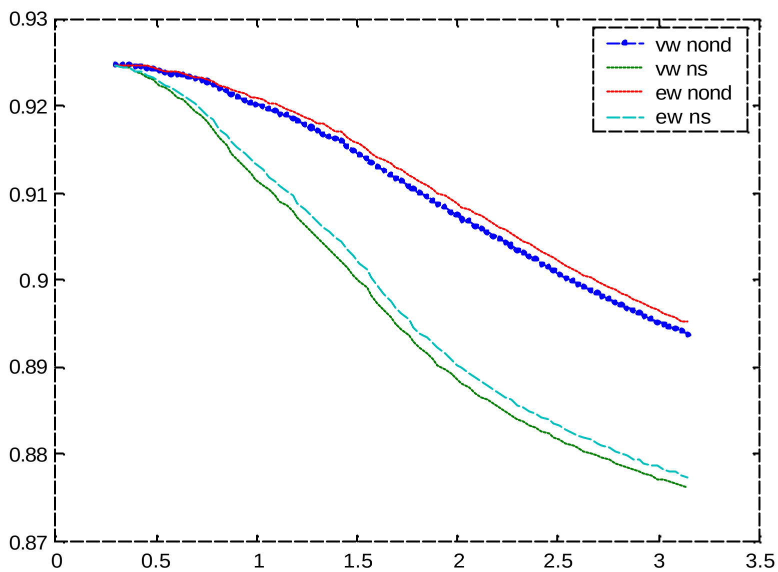
Maximum Frequency	π/20	π/10	π/5	π/4	π/3	π/2	π
Nondurables							
Value-Weighted							
β	0.9245	0.9237	0.9217	0.9202	0.9179	0.9113	0.8938
	0.0861	0.0737	0.0602	0.0550	0.0487	0.0390	0.0192
γ	1.4763	2.2923	2.3812	2.1065	2.0992	2.1868	1.8238
	0.2673	0.2294	0.1892	0.1741	0.1564	0.1333	0.1187
Equal-Weighted							
β	0.9237	0.9216	0.9154	0.9116	0.9063	0.8942	0.8762
	0.0859	0.0730	0.0310	0.0515	0.0413	0.0257	0.0093
γ	2.8803	3.0391	2.8863	2.5748	2.5857	2.7062	2.6848
	0.2673	0.2302	0.1890	0.1800	0.1652	2.4690	1.0691
Nondurables Plus Services							
Value-Weighted							
β	0.9246	0.9239	0.9221	0.9208	0.9187	0.9125	0.8952
	0.0862	0.0738	0.0603	0.0552	0.0489	0.0395	0.0197
γ	0.9701	2.1740	2.6830	2.4681	2.5206	2.8395	2.5236
	0.2673	0.2294	0.1891	0.1742	0.1571	0.1421	2.0423
Equal-Weighted							
β	0.9238	0.9220	0.9166	0.9132	0.9082	0.8962	0.8774
	0.0859	0.0731	0.0583	0.0520	0.0436	0.0267	0.0098
γ	2.8786	3.5988	4.0187	3.7932	3.8501	4.3270	4.4349
	0.2673	0.2300	0.2777	0.2156	5.1702	3.5234	1.5801

Notes: Alternative estimates of CCAPM parameters. Estimates based on two sets of NYSE index returns and seasonally-adjusted per-capita expenditure on two measures of consumption. Data are quarterly and span from January 1947 to December 1998. Standard errors are displayed beneath the point estimates. The estimator is the CVM version of GSE for low frequency bands, $\omega \in [0, \omega_{max}]$, indexed by ω_{max} .

Figure 1. CCAPM Parameter Estimates:
Expanding Frequency Bands



Estimates of γ

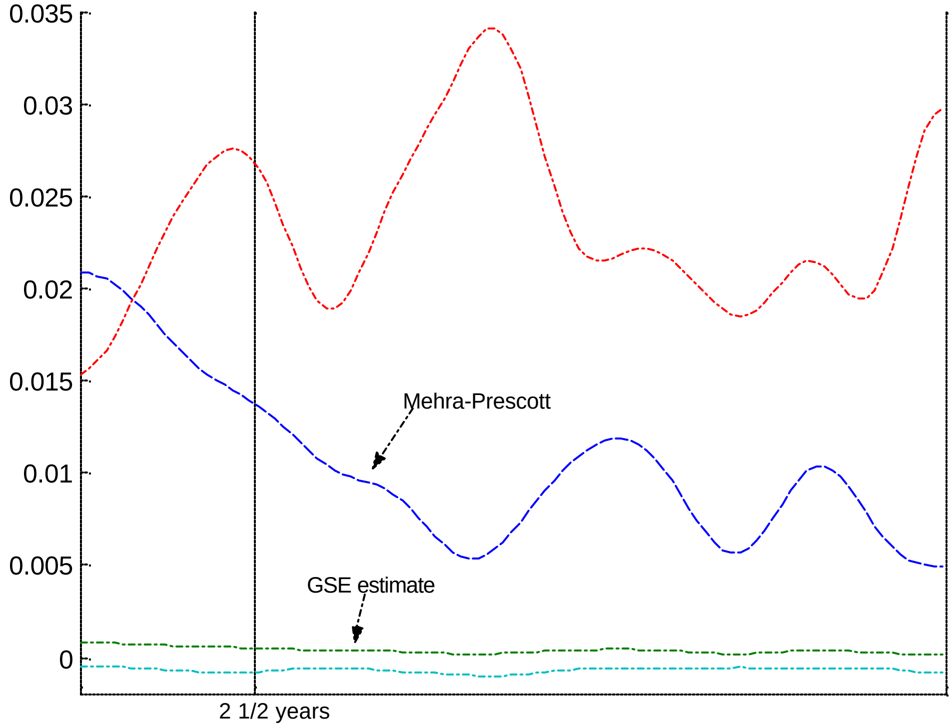


Estimates of β

Notes: GSE estimates of the two parameter CCAPM over 90 frequency bands, $[0, \omega_{max}]$, for $\omega_{max} = 10\pi/100, 11\pi/100, \ldots,$

π using the CVM cost function. When $\omega_{max} = \pi$, all frequencies are included in estimation. Plots are based on two sets of NYSE index returns and seasonally-adjusted per-capita expenditure on two measures of consumption.

Figure 2
Model Spectra and Confidence Tunnel
CRSP Equal-Weighted Returns



Notes: Two model-implied spectra shown with a 95% confidence tunnel calculated under the null of white noise. The spectrum labeled "Mehra-Prescott" corresponds to the Euler residual of the CCAPM with β=.99 and γ=19. The spectrum marked "GSE" is implied by parameters estimated over frequency band $[0,\pi/5]$ with the CVM cost function. The estimated values are β=.92 and γ=3.7.