MASSACHUSETTS INSTITUTE of TECHNOLOGY
LIBRARIES

M.I.T. LIBRARIES - DEWEY

HD28
.M414
no.
3728-
94

18

FINITE SAMPLE PROPERTIES OF SOME ALTERNATIVE GMM ESTIMATORS

Lars Peter Hansen, University of Chicago
John Heaton, Massachusetts Institute of Technology
Amir Yaron, Carnegie-Mellon University

WP#3728-94-EFA September 1994

Finite Sample Properties of Some Alternative GMM Estimators *

Lars Peter Hansen[†], John Heaton[‡] and Amir Yaron[§]

September 1994

Abstract

We investigate the small sample properties of three alternative *GMM* estimators of asset pricing models. The estimators that we consider include ones in which the weighting matrix is iterated to convergence and ones in which the weighting matrix is changed with each choice of the parameters. Particular attention is devoted to assessing the performance of the asymptotic theory for making inferences based directly on the deterioration of *GMM* criterion functions.

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

DEC 01 1994

LIBRARIES

*We thank Renee Adams, Wen Fang Liu, Alexander Monge, Marc Roston and George Tauchen for helpful comments and Marc Roston for assisting with the computations. Financial support from the National Science Foundation (Hansen and Heaton) and the Sloan Foundation (Heaton and Yaron) is gratefully acknowledged. Additional computer resources were provided by the Social Sciences and Public Policy Computing Center. Part of this research was conducted while Heaton was a National Fellow at the Hoover Institution.

[†]University of Chicago, NORC and NBER

[‡]Sloan School of Management, MIT and NBER

[§]Graduate School of Industrial Administration, Carnegie-Mellon University

1 Introduction

The purpose of this paper is to investigate the small sample properties of generalized method of moments (*GMM*) estimators applied to asset pricing models. Our alternative asymptotically efficient estimators include ones in which the weighting matrix is estimated using an initial (consistent) estimator of the parameter vector, ones in which the weighting matrix is iterated to convergence and ones in which the weighting matrix is changed for every hypothetical parameter value. The last of these three approaches has not been used very much in the empirical asset pricing literature, but it has the attraction of being insensitive to how the moment conditions are scaled. In addition, we study the advantages to basing statistical inferences directly on the criterion function rather than on quadratic approximations to it.

We address the following issues:

- How does the procedure for constructing the weighting matrix affect the small sample behavior of the *GMM* criterion function?

- What are the small sample properties of confidence regions of parameter estimators constructed using the *GMM* criterion function, compared to confidence regions constructed using standard errors?

- Is the small sample over-rejection often found in studies of *GMM* estimators reduced when using an estimator in which the weighting matrix is continuously altered?

- How are the small sample biases of the *GMM* estimators effected by the choice of procedure for constructing the weighting matrix?

Since there has been an extensive body of empirical work investigating the consumption-based intertemporal capital asset pricing model using *GMM* estimation methods, we use such models as a laboratories for our Monte Carlo experiments. As in Tauchen (1986) and Kocherlakota (1990b), all of our experiments come from single consumer economies with power utility functions. Within the confines of these economies, there is still considerable flexibility in the experimental design. Some of our experimental economies are calibrated to annual time series data presuming a century of data. Other economies are calibrated to monthly post war data. In the experiments calibrated to annual data and several of the experiments calibrated to monthly data, the moment conditions are nonlinear in at least one of the parameters of interest. Moreover, some of the specifications introduce time nonseparabilities in the consumer preferences which are motivated by either local durability or habit persistence. In the other experiments calibrated to monthly data, the moment conditions are, by design, linear in the parameter of interest. Some of these setups are special cases of the classical simultaneous equations model. For other setups, the observed data is modeled as being time averaged, introducing a moving-average structure in the disturbance terms.

The paper is organized as follows. Section 2 describes the alternative estimators we study and the related econometric literature. Section 3 specifies the Monte Carlo environments we use. Section 4 gives an overview of the calculations including a description of how inferences are made based directly on the shape of the criterion functions. Section 5 then presents the results of the Monte Carlo experiments calibrated to monthly data that are linear in the parameter of interest. Section 6 presents the results calibrated to annual and monthly data in which the moment conditions are nonlinear in the parameters. Finally, our concluding remarks are in section 7.

2 Alternative Estimators and Related Literature

One of the goals of our study is to compare the finite sample properties of three alternative *GMM* estimators, each of which uses a given collection of moment conditions in an asymptotically efficient manner. Write the moment conditions as:

$$E[\varphi(X_t, \beta)] = 0 \tag{2.1}$$

where β is the k dimensional parameter vector of interest. In (2.1) the function φ has $n \geq k$ coordinates. We assume that $\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varphi(X_t,\beta)\}$ converges in distribution to a normally distributed random vector with mean zero and covariance matrix $V(\beta)$.

Let $V_T(\beta)$ denote (an infeasible) consistent estimator of this covariance matrix. This latter estimator is typically made operational by substituting a consistent estimator for β, denoted $\{b_T^1\}$. An efficient *GMM* estimator of the parameter vector β is then constructed by choosing the parameter vector b that minimizes:

$$[\frac{1}{T}\sum_{t=1}^{T}\varphi(X_t,b)]'[V_T(b_T^1)]^{-1}[\frac{1}{T}\sum_{t=1}^{T}\varphi(X_t,b)] \tag{2.2}$$

The first two *GMM* estimators that we consider differ in the way in which this is accomplished.

Two-step Estimator

The first estimator, called the *two-step estimator*, uses an identity matrix to weight the moment conditions so that b_T^1 is chosen to minimize:

$$[\frac{1}{T}\sum_{t=1}^{T}\varphi(X_t,b)]'[\frac{1}{T}\sum_{t=1}^{T}\varphi(X_t,b)]. \tag{2.3}$$

Let b_T^2 denote the estimator obtained by minimizing (2.2).

Iterative Estimator

The second estimator continues from the two-step estimator by reestimating the matrix $V(\beta)$ using $V(b_T^{j-1})$ and constructing a new estimator b_T^j. This is repeated until b_T^j converges or until j attains some large value. Let b_T^∞ denote this estimator.

Continuous-updating Estimator

Instead of taking the weighting matrix as given in each step of the *GMM* estimation, we also consider an estimator in which the covariance matrix is continuously altered as b is changed in the minimization. Formally let b_T^c be the minimizer of:

$$[\frac{1}{T}\sum_{t=1}^{T}\varphi(X_t,b)]'[V_T(b)]^{-1}[\frac{1}{T}\sum_{t=1}^{T}\varphi(X_t,b)]. \tag{2.4}$$

Allowing the weighting matrix to vary with b clearly alters the shape of the criterion function that is minimized. While the first-order conditions for this minimization problem have an extra term relative to problems with a fixed weighting matrices, this term does not distort the limiting distribution for the estimator. [See Pakes and Pollard (1989, pages 1044-1046) for a more formal discussion and provision of sufficient conditions that justify this conclusion.] An advantage of this estimator relative to the previous two is that it is invariant to how the moment conditions are scaled even when parameter dependent scale factors are introduced. A simple example of a continuous-updating estimator is a minimum chi-square estimator used for restricted multinomial models in which the efficient distance matrix is constructed from the probabilities implied by the underlying parameters and hence is parameter dependent.

The three GMM estimators have antecedents in the classical simultaneous equations literature. Consider estimating a single equation, say

$$y_t = \beta' x_t + u_t \tag{2.5}$$

where β is the parameter of interest. Let z_t denote the vector of predetermined variables at time t which by definition are orthogonal to u_t. One way to estimate β is to use two-stage least squares which is our two-step estimator under the additional restrictions that the disturbance term is conditionally homoskedastic, and serially uncorrelated. In this case the iterative estimator converges after two-steps and hence is the two-step estimator. It is well known that the two-stage least squares estimator is not invariant to normalization. In fact Hillier (1990) criticized the two-stage least squares estimator by arguing that the object that is identified is the direction $\begin{bmatrix} 1 & -\beta' \end{bmatrix}$ but not its magnitude. Hillier then showed that the conventional two-stage least squares estimator of direction is distorted by its dependence on normalization.

As an alternative, Sargan (1958) suggested an instrumental variables type estimator that minimizes

$$\frac{\frac{1}{T}\sum_{t=1}^{T}(y_t - b'x_t)z_t' \left[\frac{1}{T}\sum_{t=1}^{T} z_t z_t'\right]^{-1} \frac{1}{T}\sum_{t=1}^{T}(y_t - b'x_t)z_t}{\frac{1}{T}\sum_{t=1}^{T}(y_t - b'x_t)^2} \tag{2.6}$$

by choice of b. Under the additional restrictions imposed on the disturbance term above, this is our continuous-updating estimator. Notice that if we ignore the denominator term in (2.6) and minimize, the solution is the two-stage least squares estimator. By including the denominator term, Sargan showed that for an appropriate choice of z_t, the solution is the (limited information) quasi-maximum likelihood estimator (using a Gaussian likelihood), which as an estimator of direction is invariant to normalization.[1]

The estimation environments that we study are more complicated than the one just described. Sometimes the moment conditions are not linear in the parameters, and the disturbance terms are often conditionally heteroskedastic and/or serially correlated. As a consequence, the two-step and iterative estimators no longer coincide. However, the estimation methods remain limited information in that the moment conditions used are typically not sufficient to fully characterize the time series evolution of the endogenous variables.

The second goal of our analysis is to compare the reliability of confidence regions computed using quadratic approximations to criterion functions to ones based directly on the deterioration of the original criterion functions. The former approach is more commonly used in the empirical asset pricing literature partially because it is easier to implement. The latter approach exploits the chi-square feature of the appropriately scaled criterion functions. From the vantage point of hypotheses testing, the plausibility of an observed deterioration of the criterion function caused by imposing parameter restrictions can be assessed by using the appropriate chi-square distribution. Using this same insight, confidence regions can be

[1] See Imbens (1992) for an alternative GMM estimator that is invariant to normalization. Imbens' estimator is constructed to coincide with the maximum likelihood estimator when the data is multinomial

computed by using the appropriate chi-square distribution to prespecify some increment in the criterion function and inferring the set of parameter values that imply no more than that increment. Such confidence regions can have unusual shapes and, in fact, may not even be connected. One of the key questions of this investigation is whether or not they lead to more reliable statistical inferences.

One of our reasons for studying the performance of criterion-function-based inference comes from the work of Magdalinos (1994). Within the confines of the classical simultaneous equations paradigm, Magdalinos studied the performance of alternative tests of instrument admissibility. As a result of his analysis, Magdalinos recommended altering the weighting matrix to embody the restrictions as is done in the continuous-updating method. In addition, he found that test statistics are better behaved using the limited information maximum likelihood estimator than the two-stage least squares estimator. Recall that the former estimator coincides with our continuous-updating estimator and the later to our two-step and iterated estimators in the classical simultaneous equations estimation environment considered by Magdalinos. Another reason is Nelson and Startz's (1990) criticism of the use of instrumental variables methods for studying consumption-based asset pricing models. These authors were concerned about the behavior of instrumental variables estimators when the instruments are poorly correlated with the endogenous variables. Their arguments are based on analogies to results derived formally for t-statistics and over-identifying restrictions tests in the classical simultaneous equations setting. The question of interest to us is the extent to which criterion-function based inference and continuous updating can help overcome the concerns of Nelson and Startz.

The finite sample properties of the two-step and iterative *GMM* estimators in an asset-pricing setting have been studied previously by Tauchen (1986), Kocherlakota (1990b), and Ferson and Foerster (1991). These investigators did not study the properties of the continuous-updating estimator, nor did they study the behavior of criterion function-based confidence regions. Further Tauchen (1986) and Kocherlakota (1990b) considered only the case of time separable preferences for the representative consumer.

3 Monte Carlo Environment

We consider several Monte Carlo environments to assess the finite sample properties of the estimators described in section 2. The data generating mechanisms are constructed to be consistent with a representative agent consumption-based asset pricing model (*CCAPM*). Estimators of the parameters of the representative agent's utility function are considered along with tests of the overidentifying conditions implied by the model. We use the *CCAPM* as the basis of our experiments because *GMM* has been used extensively in studying this model.[2] Further, the *CCAPM* forms the basis for two studies of the finite sample properties of *GMM* conducted by Tauchen (1986) and Kocherlakota (1990b).

Preferences and Euler Equations

In the model the representative consumer is assumed to have preferences over consumption given by:

$$U_0 = E\left[\sum_{t=0}^{\infty}\delta^t\frac{(c_t+\theta c_{t-1})^{1-\gamma}-1}{1-\gamma}\right],\gamma>0 \tag{3.1}$$

where c_t is consumption at date t. The parameter θ captures some time nonseparability in preferences.[3] If $\theta > 0$, consumption is durable or substitutable over time. If $\theta = 0$, the preferences of the consumer are time additive. If $\theta < 0$, consumption is complementary over time and the preferences of the representative consumer exhibit habit persistence.

We consider estimators of the parameters δ, γ and θ as well as tests of the model based on implications of the Euler equations. Let $mus_t \equiv (c_t+\theta c_{t-1})^{-\gamma}$, which can be interpreted as the indirect marginal utility for consumption "services" as measured by $s_t = c_t + \theta c_{t-1}$.

[2]See, for example, Dunn and Singleton (1986), Eichenbaum and Hansen (1990), Epstein and Zin (1991), Ferson and Constantinides (1991), and Hansen and Singleton (1982).

[3]For models with time nonseparability in preferences see, for example, Abel (1990), Constantinides (1990), Detemple and Zapatero (1991), Dunn and Singleton (1986), Eichenbaum and Hansen (1990), Gallant and Tauchen (1989), Heaton (1993, 1994), Novales (1990), Ryder and Heal (1973) and Sundaresan (1989).

Similarly, let $muc_t \equiv (c_t + \theta c_{t-1})^{-\gamma} + \theta\delta E[(c_{t+1} + \theta c_t)^{-\gamma} \mid \mathcal{F}_t]$ where $\mathcal{F}_t$ gives the information set at time t. The Euler equation for a representative agent's portfolio allocation decision is given by:

$$muc_t = E(\delta muc_{t+1} R_{t+1} \mid \mathcal{F}_t) \tag{3.2}$$

where R_{t+1} is a gross return on an asset from t to $t+1$. Since aggregate consumption is growing over time we divided (3.2) by mus_t to induce stationarity. The (normalized) Euler equation that we consider is then given by:

$$\frac{muc_t}{mus_t} = E\left(\delta \frac{muc_{t+1}}{mus_t} R_{t+1} \mid \mathcal{F}_t\right). \tag{3.3}$$

Removing conditional expectations from (3.3) results in the Euler equation error:

$$\phi_{t+2}(\delta, \gamma, \theta) = \frac{muc_t^*}{mus_t} - \delta \frac{muc_{t+1}^*}{mus_t} R_{t+1} \tag{3.4}$$

where $muc_t^* \equiv (c_t + \theta c_{t-1})^{-\gamma} + \delta\theta(c_{t+1} + \theta c_t)^{-\gamma}$. Notice that $E(\phi_{t+2} \mid \mathcal{F}_t) = 0$. Further notice that when θ is not zero, ϕ_{t+2} has an MA(1) structure. By choosing instruments, z_t, in $\mathcal{F}_t$, unconditional moment conditions are given by:

$$E[\varphi_{t+2}(\delta, \gamma, \theta)] = E[z_t \phi_{t+2}(\delta, \gamma, \theta)] = 0. \tag{3.5}$$

Finally, notice that ϕ_{t+2} can be expressed in terms of consumption ratios and returns, which we take to be stationary processes.

One unpleasant feature of these moment conditions is that they can always be made to be satisfied in a degenerate fashion. Suppose that $\gamma = 0$ and $\delta\theta = -1$, then clearly $muc_t^* = 0$ and the moment conditions are trivially satisfied. Without imposing additional constraints on the parameter vectors, this degeneracy in the moment conditions creates problems for the two-step and iterative estimators. Of course when time-separability is imposed ($\theta = 0$), these problematic parameter values cannot be reached. When θ is permitted to be different from zero. Eichenbaum and Hansen (1990) were led to divide the moment conditions by $1 + \delta\theta$ in order that the two-step and iterative estimators not be driven to the degenerate values. We will do likewise. An attractive property of the continuous-updating estimator is its insensitivity to parameter dependent scale factors and hence to the moment transformation used by Eichenbaum and Hansen (1990). Moreover, the criterion of the continuous-updating estimator does not necessarily tend to zero in the vicinity $\gamma = 0$ and $\delta\theta = -1$. Although the estimated mean for $\{\varphi_{t+2}\}$ becomes small in the neighborhood of these parameter values so does the estimated asymptotic covariance matrix, and the criterion function for the continuous-updated estimator plays off this tension.

We build several Monte Carlo environments to simulate returns and consumption growth that are consistent with (3.3). These are used to assess the finite sample properties of the estimators of section 2 based upon the moment conditions (3.5).

3.1 Log-Normal Model

Time Additive Model, No Time Averaging

In our first Monte Carlo environment we model consumption growth and returns directly by assuming that they are jointly log-normally distributed as in Hansen and Singleton (1983). Let $Y(t) = [\log(c_t/c_{t-1}) \;\; \log R_t^e \;\; \log R_t^f]'$ where c_t is aggregate consumption at time t, R_t^e is the gross return on a stock index at time t and R_t^f is the gross return on a bond at time t . We assume that:

$$Y_t = \mu + \mathrm{B}(L)\epsilon_t \tag{3.6}$$

where ϵ_t is a normally distributed three dimensional random vector that is independent over time, has zero mean and covariance matrix I, and where μ is the mean of Y_t. Further B(L) is a matrix of polynomials in the lag operator. To use this assumption about the dynamics of consumption and returns along with the Euler equation (3.2) we assume that the preferences of the representative agent are time additive. In this case the Euler equation error for each return can be written as:

$$-\gamma \log(c_{t+1}/c_t) + \log R_{t+1} - \kappa = \eta_{t+1} \tag{3.7}$$

where $E(\eta_{t+1} \mid \mathcal{F}_t) = 0$ and κ is a constant [see, for example, Hansen and Singleton (1983)].

The relation (3.7) implies a set of restrictions on the law of motion (3.6). To impose these restrictions we consider several finite order parameterizations of B(L) and use the methods described by Hansen and Sargent (1991). These parameterizations are of the form:

$$\mathrm{B(L)} = \frac{\mathrm{C(L)}}{\alpha(\mathrm{L})} \tag{3.8}$$

where C(L) is a 3 by 3 matrix of polynomials in the lag operator and α(L) is a 1 by 1 polynomial in the lag operator. We restrict the polynomial α(L) to be second order and considered several different orders for C(L).

To estimate the constrained law of motion, we used monthly data from 1959,2 to 1992,12. Aggregate consumption is seasonally adjusted real aggregate consumption of nondurables plus services for the U.S. taken from CITIBASE. These data were converted to a per capita measure by dividing by total U.S. population for each month, obtained from CITIBASE. The equity return is the value weighted return from CRSP, and the bond return is the Fama-Bliss risk free return from CRSP. Each of these return series was converted into a real return using the implicit price deflator for nondurables and services from CITIBASE.

For simplicity we removed the sample mean from the vector Y_t so that the constants in (3.6) and (3.7) did not have to be estimated. As a result, the only preference parameter to be estimated is γ. The results of estimating the law of motion (3.6) using exact Maximum Likelihood for different orders of the polynomial C(L) are given in table 3.1. The column labeled "Unrestricted Log-Likelihood" reports the log-likelihood in the case of unrestricted estimation of the polynomials in the lag operator. The columns labeled "No Time Avg."

report the log-likelihood and the estimated value of γ under the restrictions implied by (3.7).[4] Notice that there is substantial improvement in the log-likelihood in moving from a first to a second order polynomial for C(L) in the unrestricted case. This indicates that more than a first order polynomial is needed for C(L) . There is little improvement in going to a third order polynomial in the unrestricted case.

For both the second and third order polynomial cases there is great deterioration in the log-likelihood when the model restrictions are imposed. This is consistent with the results reported by Hansen and Singleton (1983). Also there is little improvement in the log-likelihood in moving from a second order to a third order C(L). For this reason we used the point estimates from the restricted model with a second order C(L) to conduct our Monte Carlo experiments for the log-normal model with no time averaging.

In assessing the finite sample properties of *GMM* estimators in this case we constructed 500 Monte Carlo samples each with a sample size of 400. A sample size of 400 approximates the size of available monthly consumption data. We constructed moment conditions based upon the Euler equation errors for both the bond and the stock returns simultaneously. For each Euler equation error we used one period lagged (log) bond and (log) stock returns and one period lagged (log) consumption growth as instruments. We did not include constants as instruments since the data is simulated under the assumption that it has a zero mean.

Time Additive Model with Time Averaging

As a further data generating mechanism, we also consider an example in which the decision interval of the representative agent is much smaller than the interval of the data. Suppose that there are n decision periods within each observation period. For example if the representative agent's decision interval is a week and data is observed monthly, then n

[4]In searching for the maximized log-likelihood, larger values of the log-likelihood were found for values of γ larger than 50. The results reported in table 3.1 for the constrained models correspond to local maxima. We used the local maximizers for our simulations because they result in plausible values for γ.

would be approximately 4. The representative agent's utility function at time t is given by:

$$U_t = E\left[\sum_{h=0}^{\infty}(\delta_n)^h \frac{\left(c_{t+\frac{h}{n}}\right)^{1-\gamma} - 1}{1-\gamma} \mid \mathcal{F}_t\right]. \tag{3.9}$$

If we maintain the assumption that consumption and returns are jointly lognormally distributed, the Euler equation error η_{t+1} is again given by (3.7). Moreover, this error can be decomposed as:

$$\eta_{t+1} = \sum_{h=1}^{n} \zeta_{t+\frac{h}{n}} \tag{3.10}$$

where

$$\zeta_{t+\frac{h}{n}} \equiv E(\eta_{t+1} \mid \mathcal{F}_{t+\frac{h}{n}}) - E(\eta_{t+1} \mid \mathcal{F}_{t+\frac{h-1}{n}}) \tag{3.11}$$

In this environment, we presume that observed consumption does not correspond to the actual point-in-time consumption the representative agent, but instead is an average of actual consumption over one unit of time. Specifically, suppose that observed consumption, c_t^a, is a geometric average of actual consumption:

$$c_t^a = \left(\prod_{h=1}^{n} c_{t-1+\frac{h}{n}}\right)^{1/n} . \tag{3.12}$$

and similarly for the observed return, R_t^a. Averaging (3.7) over time implies that:

$$-\gamma \log(c_{t+1}^a/c_t^a) + \log R_{t+1}^a = \kappa + \frac{1}{n}\sum_{\tau=1}^{n}\sum_{h=1}^{n} \zeta_{t-1+\frac{\tau}{n}+\frac{h}{n}}. \tag{3.13}$$

Notice that in this case the Euler equation error:

$$\eta_{t+1}^a \equiv \frac{1}{n}\sum_{\tau=1}^{n}\sum_{h=1}^{n} \zeta_{t-1+\frac{\tau}{n}+\frac{h}{n}} \tag{3.14}$$

is predictable at time t. However $E(\eta_{t+1}^a \mid \mathcal{F}_{t-1}) = 0$ so that instruments can be chosen from the information set at time $t-1$. An instrumental variables estimator of this model must account for the MA(1) structure of the moment condition.

We estimated the log-linear law of motion (3.8) for consumption and asset returns[5] under the restriction implied by (3.13). Notice that this model imposes a weaker set of restrictions than does the model that takes no account of time averaging.[6] We consider the case of a third order polynomial for C(L) and a second order polynomial for α(L). The results of this estimation are reported in table 3.1 in the columns labeled "Time Averaged." Notice that the log-likelihood function improves somewhat compared to the case where time averaging is ignored. However the model is still substantially at odds with the data. The estimated value of γ is slightly larger as well.

We used this model to create 500 Monte Carlo draws each with a sample size of 400. As in the case of no time averaging, we studied estimators based upon the Euler equations for both returns. In this case the instruments were (log) stock returns, (log) bond returns and (log) consumption growth all lagged 2 periods.

3.2 Discrete-State Models

In our second set of Monte Carlo environments we follow Tauchen (1986) and Kocherlakota (1990b) and consider a Markov chain model for aggregate consumption and dividend growth. Aggregate consumption is assumed to represent the endowment of the representative consumer and dividends represent the cash flow from holding stock. We form one-period stock returns and the returns to holding a one-period (real) discount bond. Each of these returns

[5]The actual consumption data is an arithmetic average of consumption expenditures over a period. In fitting the model to the data we are assuming that geometric averages and arithmetic averages are approximately the same. This assumption was made by Grossman, Melino and Shiller (1987), Hall (1988) and Hansen and Singleton (1993). Notice also that the returns in (3.13) are time averaged. For simplicity, in estimating the law of motion (3.8) we use the monthly CRSP series directly. Hansen and Singleton (1993) constructed time averaged returns from daily data in their analysis of this model. For the consideration of the finite sample properties of GMM estimators, this is unlikely to be an important issue.

[6]There is a additional restriction on the first-order autocorrelation of the Euler equation error for the stock return. In the limit case of continuous decision making, the first-order autocorrelation of the error should be 0.25 as discussed by Grossman, Melino and Shiller (1987) and Hall (1988). We do not impose this restriction in our estimation.

can be represented as functions of the state of the Markov chain. Construction of these returns in the case of time additive utility is described by Kocherlakota (1990b) and Tauchen (1986).

Annual Model

To calibrate the first Markov chain we used the method described by Tauchen and Hussey (1991) to approximate a first-order VAR for consumption and dividend growth. The parameters of the VAR are taken from Kocherlakota (1990b) and are given by:

$$\begin{bmatrix} \ln \xi_t \\ \ln \lambda_t \end{bmatrix} = \begin{bmatrix} 0.004 \\ 0.021 \end{bmatrix} + \begin{bmatrix} 0.117 & 0.414 \\ 0.017 & 0.161 \end{bmatrix} \begin{bmatrix} \ln \xi_{t-1} \\ \ln \lambda_{t-1} \end{bmatrix} + \epsilon_t \tag{3.15}$$

where ξ_t is the gross growth rate of real annual dividends on the S&P500, where λ_t is the gross growth rate of U.S. per capita real annual consumption, and where

$$E(\epsilon_t \epsilon_t') = \begin{bmatrix} 0.01400 & 0.00177 \\ 0.00177 & 0.00120 \end{bmatrix}. \tag{3.16}$$

Further ϵ_t is assumed to be normally distributed and uncorrelated over time. The Markov chain for $[\xi_t \ \lambda_t]'$ is chosen to have 16 states.

In simulating data from this model we chose several values of the preference parameters of the representative consumer. These are presented in table 3.2. Preference setting TS1 was used by Tauchen (1986) and preference setting TS2 was used by Kocherlakota (1990b). As shown by Kocherlakota (1990b), these latter parameters, along with the Markov chain model of endowments, imply first and second moments for asset returns that mimic their sample counterparts. The large value of δ in TS2 is not inconsistent with the existence of an equilibrium in the model because of the large value of γ [Kocherlakota (1990a)].

We restrict our attention to "moderate" values of δ and γ in our examination of time nonseparable preferences, and we consider a range of values of θ. Parameter setting TNS1 introduces a modest degree of durability by letting $\theta = 1/3$ and TNS2 introduces habit persistence with $\theta = -1/3$. TNS3 results in a more extreme amount of habit persistence by setting $\theta = -2/3$. The asymmetry (in magnitude) across the specifications of θ is guided in part by the *a priori* notion that there should only be limited amount of durability in the goods classified as "nondurable" in NIPA and by Ferson's and Constantinides's (1991) empirical evidence for a substantial degree of habit persistence.

Since there is great flexibility in the construction of moment conditions, we chose several sets of moment conditions that differ in the number of returns and instruments that are used. The alternative sets are listed in table 3.3. The Euler equation for each listed return was multiplied by the listed instruments to construct the moment conditions. R_t^e denotes the return to holding the stock and R_t^f to holding the bond. We used the dividend-price ratio, d_t/P_t^e, as an instrument in moment set M4 because of its ability to predict equity returns. GMM estimates and test statistics were computed for 500 replications of a sample of size 100. This sample size corresponds approximately to the length of most annual data sets.

Monthly Model

We repeated some of the experiments using a law of motion calibrated to postwar monthly data. As in the construction of the Markov chain for the annual model, we started with a first-order VAR for consumption and dividend growth. The consumption data used in estimating the VAR was aggregate U.S. expenditures on nondurables and services described in subsection 3.1. We constructed dividends implied by the monthly CRSP value-weighted NYSE portfolio return. These dividends were converted to real dividends using the implicit price deflator for monthly nondurables and services taken from CITIBASE. This dividend series is highly seasonal because of the regular dividend payout policies of most companies. To avoid modeling this seasonality we let $\xi_t \equiv \log(d_t/d_{t-12})/12$ and we assumed that ξ_t represents the one-period dividend growth of the model. The series $\{\log(d_t/d_{t-12})\}$ appears to be stationary.

The parameters of the VAR estimated using this data are given by:

$$\begin{bmatrix} \ln \xi_t \\ \ln \lambda_t \end{bmatrix} = \begin{bmatrix} 0.0012 \\ 0.0019 \end{bmatrix} + \begin{bmatrix} -0.1768 & 0.1941 \\ 0.0267 & -0.2150 \end{bmatrix} \begin{bmatrix} \ln \xi_{t-1} \\ \ln \lambda_{t-1} \end{bmatrix} + \epsilon_t \tag{3.17}$$

where

$$E(\epsilon_t \epsilon_t') = \begin{bmatrix} 0.1438 & 0.0001 \\ 0.0001 & 0.0145 \end{bmatrix} \times 10^{-3}. \tag{3.18}$$

As in the case of the annual Markov chain model, we approximated the VAR of (3.17) and (3.18) with a 16 state Markov chain using the methods of Tauchen and Hussey (1991). The Monte Carlo data consisted of 500 replications of a sample size of 400. We focused exclusively on the more "moderate" preference configuration (TS1) adjusting δ for the shorter sampling interval. (More precisely, we used the twelfth root of .97 in place of .97 for δ.) In addition, we generated Monte Carlo data using nonseparable specification TNS1 and TNS2, again with δ adjusted appropriately.

4 Overview of Descriptive Statistics for Monte Carlo Results

In describing the results of the various Monte Carlo experiments in sections 5 and 6, we focus most of our discussion on the following calculations:

1. We found the minimum value of the criterion function. Call this J_T. The limiting distribution of TJ_T is chi-square with degrees of freedom equal to the number of moment conditions minus the number of parameters estimated. We used this limiting distribution to test the overidentifying moment conditions.

2. We evaluated the criterion function at the true parameter vector. Call this J_T^{tr}. The limiting distribution of TJ_T^{tr} is chi-square with degrees of freedom equal to the number of moment conditions. With this limiting distribution, we characterize the family of

parameter vectors that look plausible from the standpoint of the moment conditions. It is of interest to see how well the limiting distribution performs in making this assessment.

3. We found the minimum value of the criterion function when γ is constrained to be its true value. Call this J_T^γ. Since γ is the only parameter estimated in the log-normal model, in this case J_T^γ coincides with J_T^{tr}. The limiting distribution of $T(J_T^\gamma - J_T)$ is chi-square with one degree of freedom. This limiting distribution allows us to construct a confidence region for γ based on the increments of the criterion function from its unconstrained minimum. By evaluating $T(J_T^\gamma - J_T)$ we determined whether the true value of γ is in the resulting interval for alternative confidence levels.

4. We constructed the more standard confidence intervals for γ based on a quadratic approximation to the criterion function. Let γ_T be an estimator of γ, and σ_T^γ be the estimated asymptotic standard error[7] of the estimator γ_T. Formally, we study $T(\gamma_T - \gamma)^2/(\sigma_T^\gamma)^2$ which has an asymptotic chi-square distribution with one degree of freedom. Notice that this object is just the Wald statistic for the hypothesis that the true value of the parameter is γ.

Our Monte Carlo calculations are greatly simplified by our knowledge of the true parameter vector. In empirical work, the corresponding computations would be more complicated. For instance, to construct a confidence interval for γ based on the original criterion function, a researcher would have to characterize numerically the hypothetical values of this parameter that are consistent with a prespecified deterioration in the criterion while concentrating out all of the other parameters. When there are very few remaining components in the parameter vector (in our examples zero, one or two), this concentration is tractable. However, this

[7]The standard errors for the continuous-updating estimator were based only upon first derivatives of the sample moment conditions with respect to the parameters. Standard errors constructed in this way are analogous to the standard errors typically constructed in *GMM* estimation [see Hansen (1982)]. An alternative way to proceed would be to base the standard errors on derivatives of the criterion function with respect to the parameters.

approach may become very difficult when the parameter vector is large.

In reporting our Monte Carlo results we use one of the graphical methods advocated by Davidson and McKinnon (1994). For each Monte Carlo setup we computed the empirical distributions of the statistics and compare them to the corresponding chi-square distributions. The results are plotted on a set of figures constructed as follows. For each probability value (depicted on the x-axis), we computed the corresponding chi-square critical value and the fraction of the actual computed statistics that are above that value (depicted on the y-axis). Thus the 45 degree line (depicted as ...) is the appropriate reference for assessing the quality of the limiting distribution. Following Davidson and McKinnon (1994), these plots are referred to as "p-value" plots, and we present the results for the interval [0 0.5] since this bounds the region of probability values used in most applications. Although we use probability values as our basis of comparison, confidence intervals at alternative significance levels can be assessed by simply subtracting the probability values from one.

The figures are organized as follows. For each Monte Carlo setup we first consider a single figure with four graphs titled as follows: "Minimized", "True", "Constrained - Minimized" and "Wald" corresponding to the statistics TJ_T, TJ_T^{tr}, $T(J_T^{\gamma} - J_T)$, and $T(\gamma_T - \gamma)^2/(\sigma_T^{\gamma})^2$, respectively. To provide a formal statistical measure of the distance between the empirical distributions and their theoretical counterparts, on each figure a band about the 45 degree line is plotted using dotted lines. This band is a 90 percent confidence region based on the Kolmogorov-Smirnov Test. This states that the probability that the maximal difference between the empirical distribution and the theoretical one will lie within those lines is 90 percent. Maximal differences within these bands are not statistically significant at the 10% significance level.[8]

In each graph, the dashed line gives the Monte Carlo results for the two-step estimator, the dot-dash line for the iterated estimator and the solid line for the continuous-updating estimator. For the minimized criterion function results there is a necessary ordering between

[8]The Kolmogorov-Smirnov confidence region is based on calculating the supremum between the empirical distribution and the 45 degree line over the region [0,1].

the continuous-updating and iterative estimator. When the iterative estimator converges, the value of the criterion function can also be obtained by the continuous-updating estimator. Since the continuous-updating estimator minimizes its criterion, this minimized value must be smaller than the criterion for the iterative estimator. As a result, the plot for the minimized criterion of the continuous-updating estimator must lie below the plot for the iterative estimator unless the iterative estimator fails to converge. There is no natural ordering between the results for the two-step estimator and the continuous-updating estimator or between the two-step estimator and the iterative estimator.

To complement our p-value plots, we also provide some results summarizing the performance of the implied parameter estimators. The finite sample properties of the point estimates are of interest in their own right and in some cases provide additional insights into the behavior of the p-value plots.

5 Monte Carlo Results, Log-Normal Model

For the case of time-averaged data, $\{\varphi_t\}$ has an MA(1) structure as we discussed in section 3.1. To account for this, the estimator of $V_T(b)$ was computed by first considering an estimator of the form:

$$V_T(b) = \frac{1}{T}\left\{\sum_{t=1}^{T}\varphi_T(X_t,b)\varphi_T(X_t,b)' + \sum_{t=2}^{T}\left[\varphi_T(X_{t-1},b)\varphi_T(X_t,b)' + \varphi_T(X_t,b)\varphi_T(X_{t-1},b)'\right]\right\} \tag{5.1}$$

where $\varphi_T(X_t,b) = \varphi(X_t,b) - \frac{1}{T}\sum_{t=1}^{T}\varphi(X_t,b)$. When this estimator was not positive definite, we used an estimator proposed by Durbin (1960) [see also Eichenbaum, Hansen and Singleton (1988)].[9] Durbin's estimator is obtained by first approximating the MA(1) model with a finite order autoregression. The residuals from this autoregression are used to approximate the

[9] In the Monte Carlo experiments with time averaging, the use of this covariance matrix estimator was not necessary at any of the converged parameter estimates.

innovations. Then the parameters of the MA(1) model are estimated by running a regression of the original time series onto a one-period lag of the "approximate" innovations. Finally, an estimate of $V_T(b)$ is formed using the estimated moving-average coefficients and sample covariance matrix for the residuals. This procedure has the advantage that the finite-order moving average structure of $\{\varphi_t\}$ is imposed and the estimator is positive semidefinite by construction. However it does rely on the choice of a finite order autoregression to use in the approximation. In implementing the estimator we ran a 12th order autoregression in the initial stage.

Criterion functions

Figures 5.1 and 5.2 report the properties of the criterion functions for the two Monte Carlo experiments. Figure 5.1 is for the case of no time averaging of the data and figure 5.2 is for the case of time averaging. The lower right plots in the figures report the results using the Wald (approximate quadratic) criteria. The results for the Wald and "Constrained-Minimized" criteria are identical for the Iterative and Two-Step estimators. This occurs because the model is linear in the parameters and the weighting matrix is fixed in constructing $(J_T^{tr} - J_T)$. For the continuous-updating estimator the results for the Wald and the "Constrained-Minimized" criteria are different due to the dependence of the weighting matrix on the hypothetical parameter values.

Notice that the small sample distributions of the minimized criterion functions for the iterative and two-step estimators are greatly distorted. The small sample size of tests of the overidentifying restrictions based upon the minimized criterion values are too large leading to over rejections of the model when using these estimators. The minimized criterion function for the continuous-updating estimator is much better behaved and the small sample distribution is very close to being χ^2 for both Monte Carlo experiments. Tests of the overidentifying restrictions of the models using the minimized value of the criterion function of the continuous-updating estimator have the correct size for the model without time averaging and similarly for the model with time-averaging for probability values less than about 0.1. Even for probability values greater than 0.1, the distribution of the minimized criterion for the continuous-updating estimator is not greatly distorted.

The finite sample coverage probabilities of the two ways of constructing confidence regions for γ are depicted under the headings "Wald" and "Constrained-Minimized". Recall that these two methods coincide when they are based on the two-step and iterative estimators, but differ when the continuous-updating estimator is used. The small sample coverage probabilities are greatly distorted for the intervals constructed with the iterative and two-step estimators. In particular, they do not contain the true parameter value as often as is to be expected from the limiting distribution. In the case of the continuous-updating estimator, coverage rates again are too small for confidence intervals built from the Wald criteria, but the distortion is substantially smaller than with the other two estimators. Finally, the coverage rates of the confidence regions implied by the "True-Minimized" criteria for the continuous-updating estimator accord well with the asymptotic distribution and are clearly better than the coverage rates for the other three criteria.

These Monte Carlo results for the log-normal model support the following remedy for the concerns raised by Nelson and Startz (1990). From the standpoint of hypothesis testing and confidence interval construction, use of the continuous-updating criterion is much more reliable than the other methods we study. The tests of the overidentifying restrictions based on the continuous-updating estimator do not reject too often and in fact are quite well approximated by the limiting distribution. While confidence intervals based on the Wald criteria can be badly distorted, particularly for the two-step and iterative estimators, confidence regions constructed from the continuous-updating criteria have coverage probabilities that are close to the ones implied by the asymptotic theory.

Parameter Estimates

Table 5.1 reports summaries of measures of central tendency for the three estimators of

γ along with 10% and 90% quantiles. The medians for the two-step and iterative estimators are considerably lower than the true value of γ while the median bias for the continuously-updated estimator is much smaller. However, the distribution for the continuous-updating estimator is also more disperse as evidenced by the larger increment between the 10% and 90% quantiles. Moreover, the Monte Carlo sample means for the continuous-updating estimator is much more severely distorted than they are for the other two estimators. The enormous sample means for the continuous-updating estimator occur because in the case of no time averaging and of time averaging there were 23 and 31 samples, respectively, in which the estimates are, in absolute value, larger than 100. When these are removed from the Monte Carlo samples, the sample means of the continuous-updating estimator are closer to the true values than are the means for the other two estimators.

Recall that the analog to the two-step and iterative estimator in the classical simultaneous equations model is two-stage least squares and that the analog to the continuous-updating estimator is limited information (quasi) maximum likelihood. It is known from the literature that there are settings in which two-stage least squares has finite moments but limited information maximum likelihood does not [*e.g.*, see Sawa (1969) and Mariano and Sawa (1972)]. In light of these theoretical results and our Monte Carlo findings, the continuous-updating estimator is not an attractive alternative to the other estimators we consider if our bases of comparison are the (untruncated) moment properties [or even relative squared errors as in Zellner (1978)]. On the other hand, Anderson, Kunitomo and Sawa (1982) advocated use the limited information estimator over the two-stage least squares estimator because, among other things, the median bias of the former estimator is smaller. We also find less distortion in the medians for the continuous-updating estimator in our experiments.[10]

As we noted previously, one attractive attribute of the continuous-updating estimator is its invariance to *ad hoc* (parameter dependent) transformations of the moment conditions.

[10] Even though our model is linear in variables and parameters, the continuous-updating estimator does not coincide with limited information maximum likelihood in our setting. Among other things, we use a heteroskedasticity consistent estimator of $V_T(b)$.

For instance, when the object of interest is a "direction" in the sense of Hillier (1990), the continuous-updating estimator is invariant to normalization. Hillier's defense of limited information maximum likelihood over conventional two-stage least squares is that the former is a better estimator of direction.[11] To see whether such a conclusion might well extends to comparisons between the continuously-updating estimator and the other two estimators we consider, we report smoothed distributions of the estimated "direction" in figures 5.3 and 5.4. We measure direction by the angle (as measured in radians) between the horizontal axis and the point $(1, \gamma)$. Since there is still a sign normalization that must be imposed for identification, we restrict attention to the interval $[-\pi/2, \pi/2]$. In smoothing the histogram, we used Gaussian kernel with a bandwidth of 0.1. The value of the density estimate is plotted at each of the sample points using a circle. The shape of the smoothed distribution along with the mass of the plotted circles provides evidence about the small sample distribution of the parameter estimators. Notice that the primary modes of the continuous-updating angle estimator are very close to the true parameter values, while the modes of the other two angle estimators are distorted. Moreover, the density estimates for the modal angle are larger for the continuous-updating method. However, the Monte Carlo distributions for the continuous-updating angle estimator also have secondary modes near $-\pi/2$, corresponding to large in magnitude estimates of γ with the wrong sign.

Continuous-Updating Criterion Function

The criterion function for the continuous-updating estimator can sometimes lead to extreme outliers for the minimizing value of γ. This occurs in the two Monte Carlo experiments for some of the trials as we discussed above. To see why this can occur suppose for simplicity that there is a single return under consideration, no time averaging and several instruments.

[11] Hillier (1990) considered an alternative *GMM* estimator of direction in which the absolute value of the parameter vector is normalized to one instead of one of the elements of that vector. He found that this alternative estimator has finite sample properties that compare more favorably to those of the limited information maximum likelihood estimator.

The moment conditions are constructed using:

$$\phi(X_t, g) \equiv [\log R_{t+1} - g \log(c_{t+1}/c_t)]z_t \tag{5.2}$$

where z_t is a vector of instruments and $E[\phi(X_t, \gamma)] = 0$ [see (3.7)]. Since the moment conditions are linear in g, the criteria for the iterated and two step estimators are quadratic in g. In contrast, the criterion for the continuous-updating estimator converges as g gets large. To see this, observe that for a large value of g the sample average of $\phi(X_t, g)$ is approximately g times the sample mean of $\log(c_{t+1}/c_t)z_t$ and the sample covariance is approximately g^2 times the sample covariance of $\log(c_{t+1}/c_t)z_t$. Therefore for large g, the criterion function is approximately a quadratic form that is $(\frac{1}{T})$ times the chi-square test statistic for the null hypothesis that

$$E[\log(c_{t+1}/c_t)z_t] = 0. \tag{5.3}$$

As a result it is possible for the minimized criterion for the continuous-updating estimator to occur for a very large value of g.

To further illustrate this potential problem, the upper plot in figure 5.5 is of the criterion function for the continuous-updating estimator for a Monte Carlo draw in which the value of g that minimizes the criterion function is 673750.4. The lower plot in figure 5.5 gives the average value of the criterion function over the Monte Carlo experiments. These results are for the case of no time averaging. Notice that in the upper plot, the criterion function approaches its lowest value as g becomes large in absolute value. Even in the lower plot the criterion function asymptotes to a local minimum for large negative values of g. The numerical search used to implement the estimator could be complicated by the flat sections of the criterion function and the search routine could end up spuriously searching in the direction of very large values of g. When the parameter vector is of low dimension, this can easily be assessed by gridding the parameter vector and evaluating the criterion function at the grid points. This is what we did to make sure that large estimates of g were not due to numerical problems. However when the parameter vector is of large dimension, implementing the continuous-updating estimator may sometimes be difficult.

6 Monte Carlo Results, Markov Chain Models

6.1 Time Separable Preferences, Annual Data

First we consider the results for time separable preferences where ($\theta = 0$) using the Markov Chain model calibrated to annual data. For these runs $V_T(b)$ is computed as a simple covariance matrix estimator. The resulting p-value plots are depicted in figures 6.1 through 6.6 for different combinations of the preference settings and moment sets given in tables 3.2 and 3.3. Features of the Monte Carlo distributions for the point estimates of γ are reported in table 6.1.

Results for $\gamma = 1.3$ and $\delta = .97$

We start by discussing the results obtained using the more "moderate" values of the preference parameters TS1 which were used by Tauchen (1986). Figure 6.1 includes a replication of findings in Tauchen (1986) using moment conditions M1. It is an example in which it is known from Tauchen that the over-identifying restrictions test "under rejects." This phenomenon can be seen in the upper left plot in figure 6.1 by noting that the dot-dash line is below the 45 degree line. Since the minimized value of the continuous-updating estimator is always less than or equal to that of the iterative estimator (when the iterative estimator converges), we expect the under rejection to be more pronounced when the over-identifying restrictions tests is based on the continuous-updating estimator. Indeed, the under rejection is more substantial for both the two-step and continuous-updating estimators. Interestingly, the underlying central limit approximation for the continuous-updating estimator looks quite good as depicted by the upper right plot in figure 6.1. The criterion-function based confidence sets are evaluated in the lower left plot in figure 6.1. The confidence sets based on the continuous-updating estimator do not contain γ as often as predicted by the asymptotic theory as might be anticipated from the downward distortion of the minimized criterion functions. In contrast, the limit theory works well for the confidence intervals based on the Wald statistic when the continuous-updating estimator is used.

Consider next the set of eight moment conditions M2. Given the large number of moment conditions relative to sample size, it is not surprising that the underlying central limit approximations are much less accurate (see the upper right plot in figure 6.2). While the asymptotic approximations are better with the continuous-updating estimator, at least for smaller probability values, the criteria evaluated at the true parameters still are too large when compared to the magnitudes predicted by the asymptotic theory. On the other hand, the minimized criteria functions are much better behaved, especially for the continuous-updating estimator with probability values less than .15 (see the upper left plot in figure 6.2). Confidence intervals based on criteria function behavior performed poorly in this setting, although they performed better for the continuous-updating estimator than for the other two estimators. Moreover, the criterion-function based confidence intervals for the continuous-updating estimator proved to be more reliable that the confidence intervals based on the Wald statistic.

Figures 6.3 and 6.4 report plots for the sets of four moment conditions M3 and M4. The reduction in moment conditions (relative to M2) leads to an improvement in the underlying central limit approximations depicted in the upper right portions of the figures. This is especially true for the continuous-updating estimator used in conjunction with moment conditions M4. The minimized criterion function values used to test the over-identifying restrictions behave as predicted by the asymptotic theory for all three estimators when moment conditions M3 are used. In contrast there is substantial under rejection for all three estimation methods when moment conditions M4 are used. Interestingly, the dot-dashed line is below the solid line in the upper left plot in figure 6.4. This means that the minimized values of the criteria for the continuous-updating estimator are not always below those of the iterative estimator. The reason for this apparent anomaly is that the iterations on the weighting matrix often did not converge for this experiment. The criterion-function-based confidence intervals work quite well for the continuous-updating estimator when moment conditions M3 are used and both methods for constructing confidence intervals work well for the two-step estimator when moment conditions M4 are used.

In summary, with the possible exception of the over-identifying restriction test using moment conditions M1, the asymptotic approximations for inferences based on the iterative estimator perform worse than the corresponding approximations for the other two estimators. In comparing the continuous-updating estimator to the two-step estimator, we found the following. First the asymptotic distribution for the over-identifying restrictions tests is more reliable when based on the continuous-updating estimator. Second, for both estimators confidence intervals constructed based on the original criterion functions, are often distorted, but with the exception of M1 they are no less reliable and often more reliable than confidence intervals constructed via quadratic approximations.

Of course, assessing the reliability of the asymptotic theory as applied to the different parameter estimators is a different question than assessing the performance of the parameter estimators themselves. In regards to this latter question, the results in the first portion of table 6.1 show that the continuous-updating estimator tends to have either comparable or considerably less bias in the medians than the other two estimators. On the other hand, the dispersion in the estimators as measured by the width between the .10 and .90 quantiles is always less and often much less for the iterative estimator than for the continuous-updating estimator.

Results for $\gamma = 13.7$ *and* $\delta = 1.139$

Next we consider results using the preference specification considered by Kocherlakota (1990b) (TS2 in table 3.1). We only consider the performance of *GMM* estimators obtained using moment conditions M3 and M4. Our results are displayed in figures 6.5 and 6.6 and table 6.1. With this change in parameter configuration, the results for the M3 moment conditions are similar to those in figure 6.3. However under the M4 moment conditions, there is no longer evidence of under rejection of the over-identifying restrictions. As before, the weighting matrices for the iterative method often failed to converge for the M4 runs. In contrast to our earlier findings, the limit theory no longer provides a good guide for the coverage probabilities for the criterion-function-based confidence sets when the two-step estimator is used in conjunction with moment conditions M4 (compare figures 6.4 and 6.6). In regards to the parameter estimates, the continuous-updating estimator again has less median bias than the other two estimators but more dispersion (as measured by the distance between quantiles.)

6.2 Time Separable Preferences, Monthly Data

To explore the extent to which the limiting distribution provides a better guide for inference for larger sample sizes (with less extreme data points), we redid some our calculations using simulations calibrated to monthly data as described in subsection 3.2. We focused exclusively on the more "moderate" preference configuration adjusting δ accordingly. In this case we only looked at estimators constructed using moment conditions M2 and M3. We are particularly interested in moment set M2 because of its common use in practice when analyzing post war data. Our results are reported in figures 6.7 and 6.8 and table 6.2. Notice that all of the asymptotic approximations are consistently reliable for the continuous-updating estimation method. In sharp contrast, large sample inferences for the two-step estimator are of particularly poor quality with the exception of the over-identifying restrictions test using M3. Also, of note is that the iterative estimates and the continuous-updating estimates are very close to one another when M3 is used. This is reflected in quantiles reported in table 6.2 as well as in the "Minimized" and "Constrained - Minimized" graphs. Presumably, the reason for this is that the weighting matrix tends to be a relatively "flat" function of the parameters.

In regards to the parameter estimates of γ, both the continuous-updating estimator and the iterative estimators have distributions that are much more concentrated around the true parameter value than the distributions for the two-step estimator (again see table 6.2).[12] These particular Monte Carlo experiments are ones in which the unconditional moment restrictions provide much more identifying information about the power parameter γ than any of the other experiments we report on. Not only are estimates more accurate than the estimates obtained from the Monte Carlo experiments calibrated to annual data, but also the estimates from the log-normal Monte Carlo experiments reported in section 5 which have the same sample size.[13]

6.3 Time Nonseparable Preferences, Annual Data

As we discussed in subsection 6.1, the continuous-updating estimator generally provides more reliable inference in the case of time separable preferences when data are generated from the annual Markov Chain model. However even for that estimator, it is only when moment conditions M3 are used that the distributions of the criteria TJ_T ("Minimized"), TJ_T^{tr} ("True") and $T(J_T^{\gamma} - J_T)$ ("Constrained-Minimized") accord well with the corresponding chi-square distributions. For these reasons we consider results with time nonseparable preferences using only moment conditions M3 and only for the continuous-updating estimator. To construct our Monte Carlo data sets we used the three time nonseparable settings of the parameters listing in table 3.2 as TNS1, TNS2 and TNS3 ($\theta = 1/3, \theta = -1/3, \theta = -2/3$). We also used data generated with $\theta = 0$, but still estimated the parameter θ. Unlike the case of time separable preferences when the restriction $\theta = 0$ is imposed in the estimation, the estimator

[12] The performance of the two-step estimator could potentially be improved by using a different weighting matrix in the first-step. For example, the residuals from nonlinear two-stage least squares applied to each Euler equation could be used to estimate the asymptotic covariance matrix of the moment conditions. Another possibility is to use the covariance matrix of the prices of the "synthetic" securities implicit in the use of instrumental variables. See Hansen and Jagannathan (1993) for a discussion of this weighting matrix.

[13] Presumably, the main reason for this disparity is that the Markov chain models are calibrated to dividend behavior rather than return behavior. As is well known from the empirical asset pricing literature, the dividend calibrations imply returns that are less volatile than historical time series because of some fundamental model misspecification.

$V_T(b)$ of the asymptotic covariance matrix accommodates an MA(1) structure in the Euler equation errors. As in section 5 we used the $V_T(b)$ estimator given in (5.1) except when it was not positive definite in which case we shifted to Durbin's (1960) estimator with a 4th order autoregression.

Figure 6.9 presents the p-value plots for the criterion functions for the different settings for θ.[14] In contrast to the time separable case (the solid line in figure 6.3), the distribution of the minimized criteria imply small sample over-rejection of the moment conditions for each of the settings for θ. Further, even when evaluated at the true parameters, the criterion functions are not distributed as a chi-square. This occurs for all four settings of θ including $\theta = 0$ (time separable preferences). Evidently the estimator of the asymptotic covariance matrix of the moment conditions, which assumes an MA(1) structure for the errors, causes small sample distortion of the *GMM* criterion function. Notice that the distributions of the Wald statistics are very far from being chi-square . Consistent with the results reported in section 5 and subsections 6.1 and 6.2, confidence intervals for γ constructed using the criterion function perform much better than those based on the Wald statistic.

Table 6.3 reports statistics summarizing the properties of the estimators of γ and θ. The estimator of γ does not in general perform as well as in the time separable preference case (see table 6.1 for the comparison). The median for γ_T is substantially below the true value of γ in the case of $\theta = 0$ and $\theta = -1/3$ and above for $\theta = 1/3$. Further, the dispersion of the estimators of γ is considerably larger than when time separability is correctly imposed, at least in part due to having to estimate an additional parameter and to accommodate MA(1) terms in the estimator of the asymptotic covariance matrix $V(b)$. Regarding the estimators of θ, there is substantial dispersion for each case as evidenced by the 10% and 90% quantiles. Notice further that there is some median bias in the estimators of θ for the cases of $\theta = 1/3$, 0 and -1/3 (panels A, B and C of table 6.3). In summary, the annual data

[14] The Monte Carlo experiments used to construct these results are independent across the different values of θ. The p-value plots of Figures 5.1, 5.2 and 6.1-6.8 considered results for fixed preference parameters and the three different estimators. The same Monte Carlo data was used for the experiments for each estimator in these plots.

do not permit simultaneous estimation of θ and γ with any reasonable precision, at least for moment conditions M3.

6.4 Time Nonseparable Preferences, Monthly Data

Using the Monthly Markov chain model we also examined the time nonseparable model for $\theta = 1/3$, 0 and -1/3 and for moment conditions M3. We further considered moment conditions M2 since these conditions are often used in practice and since the continuous-updating estimators demonstrated reasonable small sample properties under time-separable preferences. We did not consider the case of $\theta = -2/3$ with the monthly model because the Markov chain model implies that the covariance matrix of the Euler equation errors is close to being singular at this value of θ. Once again we used the estimator of $V(b)$ given by (5.1). When Durbin's (1960) estimator was necessary we used a 12th order autoregression.

Figure 6.10 reports the results for criterion functions using moment conditions M2. Under M2, the minimized criterion performs reasonably well for all three settings of θ. Tests of the overidentifying restrictions of the model have the correct small sample size in this case. Notice further that the criterion $T(J_T^{\gamma} - J_T)$ ("Constrained-Minimized") is close to being chi-square distributed, but that the distribution of the Wald statistic is very far from chi-square. Hence, we continue to find that inferences based directly on criterion functions are much more reliable than those based on quadratic approximations to the criterion functions. Finally, we report summary statistics of the central tendency of the estimators of γ and θ in table 6.4, panels A, B and C. Lack of prior knowledge of the parameter θ again causes the estimator of γ to be much less precise as measured by the distance between the 90% and 10% quantiles. (For instance, compare the first column of table 6.2 to panel B of table 6.4.)

Figure 6.11 presents the p-value plots for moment conditions M3. In this case the distribution of the minimized criterion functions and the "Constrained-Minimized" criteria are not chi-square. The model's over identifying conditions are rejected too often for all of the parameter settings and confidence intervals for the parameter γ have the wrong coverage probabilities. With moment conditions M3, allowing for time nonseparability results in a substantial number of very large estimates of γ as reflected by the size of the 90% quantiles (see table 6.4, panels C, D and E). It appears that the addition of returns as instruments (the difference between moment conditions M2 and M3) improves the performance of the estimator and the quality of the central limit approximations.

7 Concluding Remarks

In this paper we examined the finite sample properties of three alternative *GMM* estimators that differ in the way in which the moment conditions are weighted. Particular attention was paid to both the performance of tests of over-identifying restrictions and to comparing alternative ways of constructing confidence sets. In documenting finite sample properties, we used several different specifications of the consumption-based CAPM. The experiments differed substantially in the amount sample information there is about the parameters of interest. While the experiments do not uniformly support the conclusion that one estimator dominates the others, some interesting patterns emerged.

- Continuous-updating in conjunction with criterion-function based inference often performed better than other methods for annual data, however the large sample approximations are still not very reliable.

- In monthly data the central limit approximations for the continuous-updating estimation method applied in conjunction with the criterion function-based method of inference performed well in most of our experiments, including ones in which the parameters are estimated very accurately and ones in which there is a substantial amount of dispersion in the estimates.

- Confidence intervals constructed using quadratic approximations to the criterion function performed very poorly in many of our experiments.

- The continuous-updating estimator typically had less median bias than the other estimators, but the Monte Carlo sample distributions for this estimator sometimes had much fatter tails.

- The tests for over identifying restrictions are, by construction, more conservative when the weighting matrix is continuously updated, and in many cases this led to a more reliable test statistic. (However, we made no attempts at comparing power even for size corrected tests.)

Our Monte Carlo experiments for monthly data were sufficiently successful to convince us to reexamine some of the empirical evidence for the consumption-based CAPM. In most tests of the consumption-based CAPM, the model's over identifying conditions are rejected [see, for example, Hansen and Singleton (1982)]. Since the two-step or iterative estimator is typically used in practice, one potential explanation for these rejections could be the tendency of these estimators to result in over rejection of the model in small samples. To assess this possibility we estimated the time separable and time nonseparable models using the continuous-updating estimator. We used the consumption and return data described in subsection 3.1 along with moment conditions M2 given in table 3.3.

Estimation of the time separable model resulted in point estimates of δ and γ of 0.25 and 720.65 respectively. This is an example where the tail behavior of the criterion results in large estimated value of γ. The minimized *GMM* criterion was 5.94 with an implied p-value of 0.43 hence it appears that the continuous-updating estimator implies that the model is not at odds with the data. However the estimate parameters are very far from being economically plausible. As we found in several of our Monte Carlo experiments, with the continuous-updating estimator, extreme point estimates of the parameters are possible. However in those cases there typically was little deterioration in the criterion function when evaluated near the true parameter values so in practice it is important to evaluate the criterion function at plausible values of the parameters. In this case we restricted γ to the range [0 20] and estimated δ for each hypothetical value of γ. The resulting minimized criterion as a function of γ is plotted in the top panel of figure 7.1. Notice that for this range of γ the minimized criterion function is well above 30 where the implied p-value is essentially zero. As a result, once a plausible set of parameters is considered, the model is still rejected when using the continuous-updating estimator.

In estimation of the time nonseparable model the point estimates of the parameters were also quite implausible with estimates of δ, γ and θ of 1.20, 267.96 and 0.32 respectively. The bottom panel of figure 7.1 presents the criterion function for the continuous-updating estimator with γ restricted to the range [0 20]. At $\gamma = 20$ the criterion reaches a minimum of 13.55 with an implied p-value of .035. As a result, even at this extreme value for γ the model is still substantially at odds with the data.

In summary, although the continuous-updating estimator does not save the consumption-based CAPM, the experiments that we have presented provide evidence that it should be of use in many *GMM* estimation environments.

References

Abel, A. (1990): "Asset Prices Under Habit Formation and Catching Up with the Joneses," *A.E.R. Papers and Proceedings*: 38-42.

Anderson, T. W., N. Kunitomo, and T. Sawa (1982): "Evaluation of the Distribution Function of the Limited Information Maximum Likelihood Estimator," *Econometrica*, 50. 1009-1028.

Constantinides, G. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98: 519-543.

Davidson, R. and J. G. McKinnon (1994): "Graphical Methods for Investigating the Size and Power of Hypothesis Tests," Queens University Discussion Paper 903.

Detemple, J. and F. Zapatero (1991): "Asset Prices in an Exchange Economy with Habit Formation," *Econometrica*, 59, 1633-1657.

Dunn, K., and K. Singleton (1986): "Modeling the Term Structure of Interest Rates Under Nonseparable Utility and Durability of Goods," *Journal of Financial Economics*, 17: 27-55.

Durbin J. (1960): "The Fitting of Time Series Models," *Review of the International Statistical Institute*, 28, 233-244.

Eichenbaum, M., and L. Hansen (1990): "Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data," *Journal of Business and Economic Statistics*, 8, 53-69.

Eichenbaum, M., L. Hansen and K. Singleton (1988): "A Time Series Analysis of a Representative Agent Model of Consumption and Leisure Choice under Uncertainty," *Quarterly Journal of Economics*, 51-78.

Epstein, L., and S. Zin (1991): "Substitution, Risk Aversion, and Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 199, 263-286.

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

Ferson, W. and S. Foerster (1991): "Finite Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models," manuscript.

Gallant, A. and G. Tauchen (1989): "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes," *Econometrica*, 57, 1091-1120.

Grossman, S. J., A. Melino, and R. Shiller (1987): "Estimating the Continuous Time Consumption Based Asset Pricing Model," *Journal of Business and Economic Statistics*, 5: 315-327.

Hall, R. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96: 339-357.

Hansen, L. (1992): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50: 1029-1054.

Hansen, L. and R. Jagannathan (1993): "Assessing Specification Errors in Stochastic Discount Factor Models," manuscript.

Hansen, L. and T. Sargent (1991): "Exact Linear Rational Expectations Models," in *Rational Expectations Econometrics*, ed. by L. Hansen and T. Sargent, Boulder: Westview Press.

Hansen, L. and K. Singleton (1982): "Generalized Instrumental Variable Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50: 1269-1286.

Hansen, L., and K. Singleton (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91: 249-265.

Hansen, L. and K. Singleton (1993): "Efficient Estimation of Linear Asset Pricing Models with Moving-Average Errors." Manuscript, University of Chicago.

Heaton, J. (1993): "The Interaction Between Time-Nonseparable Preferences and Time Aggregation," *Econometrica*, 61, 353-385.

Heaton, J. (1994): "An Empirical Investigation of Asset Pricing with Time Nonseparable Preferences," forthcoming in *Econometrica.*

Hillier, G. (1990): "On the Normalization of Structural Equations: Properties of Direction Estimators," *Econometrica*, 58, 1181-1194.

Imbens W. G. (1992): "A New Approach to Generalized Method of Moments Estimators," Manuscript.

Kocherlakota, N.R. (1990a): "A Note on the 'Discount' Factor in Growth Economies," *Journal of Monetary Economics*, 25, 43-48.

Kocherlakota, N. (1990b): "On Tests of Representative Consumer Asset Pricing Models," *Journal of Monetary Economics*, 26, 285-304.

Magdalinos A. M. (1994): "Testing Instrument Admissibility: Some Refined Asymptotic Results," *Econometrica*, 62: 373-405.

Mariano, R.S., and T. Sawa (1972): "The Exact Finite-Sample Distribution of the Limited-Information Maximum Likelihood Estimator in the Case of Two Included Endogenous Variables," *Journal of the American Statistical Association*, 67, 159-163.

Nelson, C. R., and R. Startz (1990): "The Distribution of the Instrumental Variables Estimator and Its t-Ratio When the Instrument is a Poor One," *Journal of Business*, 63: 125-164.

Novales, A. (1990): "Solving Nonlinear Rational Expectations Models: A Stochastic Equilibrium Model of Interest Rates," *Econometrica*, 58, 93-111.

Pakes, A. and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57: 1027-1057.

Ryder, H.E., Jr., and G.M. Heal (1973): "Optimal Growth with Intertemporally Dependent Preferences," *Review of Economic Studies*, 40: 1-31.

Sargan, J. (1958): "The Estimation of Economic Relationships using Instrumental Variables," *Econometrica*, 26: 393-415.

Sawa, T. (1969): "The Exact Sampling Distribution of Ordinary Least Squares and Two-Stage Least Squares Estimator," *Journal of American Statistical Association*, 64, 923-937.

Sundaresan, S. (1989): "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies*, 2, 73-89.

Tauchen, G. (1986): "Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data," *Journal of Business and Economic Statistics*, 4: 397-425.

Tauchen, G. and R. Hussey (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59: 371-396.

Zellner, A. (1978): "Estimation of Functions of Population Means and Regression Coefficients Including Structural Coefficients; A Minimum Expected Loss (MELO) Approach," *Journal of Econometrics*, 8, 127-158.

Table 3.1: MLE of Monthly Law of Motion

Order of C(L)	Unrestricted Log-Likelihood	No Time Avg.		Time Averaged	
		Log-Likelihood	γ	Log-Likelihood	γ
1	3220.6	-	-	-	-
2	3237.3	3214.9	4.55	-	-
3	3238.8	3218.1	4.21	3230.3	4.99

Table 3.2: Preference Settings for Discrete States Space Model

$$U = E\left[\sum_{t=0}^{\infty} \delta^t \frac{(c_t + \theta c_{t-1})^{1-\gamma} - 1}{1-\gamma}\right]$$

Preference Case	δ	γ	θ
TS1	0.97	1.3	0
TS2	1.139	13.7	0
TNS1	0.97	1.3	1/3
TNS2	0.97	1.3	-1/3
TNS3	0.97	1.3	-2/3

Table 3.3: Moment Conditions

Moment Set	Returns	Instruments, z_t	# of Moment Conditions
M1	R^e_{t+1}	R^e_t, λ_t, 1	3
M2	R^e_{t+1}, R^f_{t+1}	R^e_t, R^f_t, c_t/c_{t-1}, 1	8
M3	R^e_{t+1}, R^f_{t+1}	λ_t, 1	4
M4[a]	R^e_{t+1}, R^f_{t+1}	d_t/P^e_t, R^f_t, 1	4

[a] Lagged bond return used as an instrument only for the bond return. Dividend-price ratio used as an instrument only for the stock return.

Table 5.1: Properties of Estimators of γ, Log-Normal Model

	Continuous-Updating	Iterative	Two-Step
	A. No Time Averaging, True $\gamma = 4.55$		
Median	3.72	1.64	1.73
Mean	7171.17	1.81	2.06
Truncated Mean[a]	4.47	1.81	2.06
10% Quantile	-3.67	-0.23	-0.20
90% Quantile	18.75	4.00	4.33
	B. Time Averaging, True $\gamma = 4.99$		
Median	3.48	1.75	1.74
Mean	-518.14	1.88	2.04
Truncated Mean[a]	3.18	1.88	2.04
10% Quantile	-10.48	-0.29	-0.66
90% Quantile	11.52	4.22	4.85

[a]Estimates with absolute values greater than 100 were excluded from the computation of the truncated means.

Table 6.1: Properties of Estimators of γ, Markov Chain Model, Annual Data

	Continuous-Updating	Iterative	Two-Step
	A. TS1-M1, True $\gamma = 1.3, \delta = .97$		
Mean	0.98	1.58	1.53
Truncated Mean[a]	1.25	1.58	1.53
Median	1.50	1.54	1.58
10% Quantile	−5.81	−1.91	−1.83
90% Quantile	7.80	4.61	4.59
	B. TS1-M2, True $\gamma = 1.3, \delta = .97$		
Mean	1.90	0.85	1.08
Truncated Mean[a]	1.90	0.85	1.08
Median	1.15	0.70	0.78
10% Quantile	0.61	0.13	−1.25
90% Quantile	5.76	1.60	4.09
	C. TS1-M3, True $\gamma = 1.3, \delta = .97$		
Mean	1.36	1.35	1.35
Truncated Mean[a]	1.53	1.35	1.35
Median	1.17	1.10	1.25
10% Quantile	0.61	−1.35	−2.71
90% Quantile	4.10	3.51	5.59
	D. TS1-M4, True $\gamma = 1.3, \delta = .97$		
Mean	2.84	0.78	0.90
Truncated Mean[a]	2.13	0.78	0.90
Median	1.03	0.35	0.59
10% Quantile	−0.15	−1.01	−2.73
90% Quantile	7.12	4.21	5.50
	E. TS2-M3, True $\gamma = 13.7, \delta = 1.139$		
Mean	6.29	13.68	14.10
Truncated Mean[a]	14.80	13.68	14.10
Median	14.14	13.10	13.43
10% Quantile	8.96	9.01	9.85
90% Quantile	22.68	19.40	19.54
	F. TS2-M4, True $\gamma = 13.7, \delta = 1.139$		
Mean	16.94	10.81	12.51
Truncated Mean[a]	14.78	10.81	12.50
Median	14.34	10.98	12.06
10% Quantile	4.47	2.81	6.78
90% Quantile	23.68	18.42	18.97

[a]Estimates with absolute values greater than 100 were excluded from the computation of the truncated means.

Table 6.2: Properties of Estimators of γ, Markov Chain Model, Monthly Data

True $\gamma = 1.3$, $\delta = 0.97^{1/12}$

	Continuous-Updating	Iterative	Two-Step
		A. TS1-M2	
Mean	1.37	1.25	1.57
Truncated Mean[a]	1.37	1.25	1.57
Median	1.28	1.18	1.17
10% Quantile	1.02	0.96	−2.12
90% Quantile	1.85	1.61	4.63
		B. TS1-M3	
Mean	1.38	1.38	1.84
Truncated Mean[a]	1.38	1.38	1.84
Median	1.27	1.27	1.27
10% Quantile	1.00	1.00	−2.95
90% Quantile	1.89	1.89	6.50

[a]Estimates with absolute values greater than 100 were excluded from the computation of the truncated means.

Table 6.3: Properties of Estimators of γ and θ, Markov Chain Model, Annual Data

Continuous-Updating Estimator, True $\gamma = 1.3, \delta = .97$

	γ	θ
		A. TNS1-M3, True $\theta = 1/3$
Mean	4.56	5.99×10^5
Truncated Mean[a]	4.56	0.83
Median	1.71	0.52
10% Quantile	0.64	0.10
90% Quantile	10.83	1.25
		B. TS1-M3, True $\theta = 0$
Mean	5.88	-0.34
Truncated Mean[a]	3.67	-0.34
Median	0.73	-0.12
10% Quantile	0.00	-0.92
90% Quantile	10.92	0.18
		C. TNS2-M3, True $\theta = -1/3$
Mean	3.05	-0.46
Truncated Mean[a]	2.69	-0.46
Median	0.97	-0.42
10% Quantile	0.01	-0.93
90% Quantile	8.28	-0.01
		D. TNS3-M3, True $\theta = -2/3$
Mean	0.41	0.07
Truncated Mean[a]	2.16	-0.64
Median	1.22	-0.69
10% Quantile	0.04	-0.94
90% Quantile	5.79	-0.40

[a]Estimates with absolute values greater than 100 were excluded from the computation of the truncated means.

Table 6.4: Properties of Estimators of γ and θ, Markov Chain Model, Monthly Data

Continuous-Updating Estimator, True $\gamma = 1.3, \delta = .97^{1/12}$

	γ	θ
	A. TNS1-M2, True $\theta = 1/3$	
Mean	13.56	0.35
Truncated Mean[a]	2.90	0.35
Median	1.27	0.34
10% Quantile	0.79	0.22
90% Quantile	7.35	0.48
	B. TS1-M2, True $\theta = 0$	
Mean	16.68	-0.01
Truncated Mean[a]	4.18	-0.01
Median	1.29	0.01
10% Quantile	0.38	-0.23
90% Quantile	13.46	0.19
	C. TNS2-M2, True $\theta = -1/3$	
Mean	6.03	-0.42
Truncated Mean[a]	5.82	-0.42
Median	1.24	-0.35
10% Quantile	0.00	-0.98
90% Quantile	17.95	0.09
	D. TNS1-M3, True $\theta = 1/3$	
Mean	153.96	0.66
Truncated Mean[a]	11.09	0.66
Median	1.45	0.37
10% Quantile	0.48	0.16
90% Quantile	410.29	1.16
	E. TS1-M3, True $\theta = 0$	
Mean	62.51	-0.14
Truncated Mean[a]	17.89	-0.14
Median	1.40	0.04
10% Quantile	0.03	-0.66
90% Quantile	142.79	0.26
	F. TNS2-M3, True $\theta = -1/3$	
Mean	12.98	-0.40
Truncated Mean[a]	12.78	-0.40
Median	0.35	-0.58
10% Quantile	0.01	-0.91
90% Quantile	40.99	0.17

[a]Estimates with absolute values greater than 100 were excluded from the computation of the truncated means.

Figure 5.1
Criterion Functions
Monthly Log-Normal Model, No Time Averaging

_._._Iterative, _ _ _ Two-Step, __ Continuous-Updating

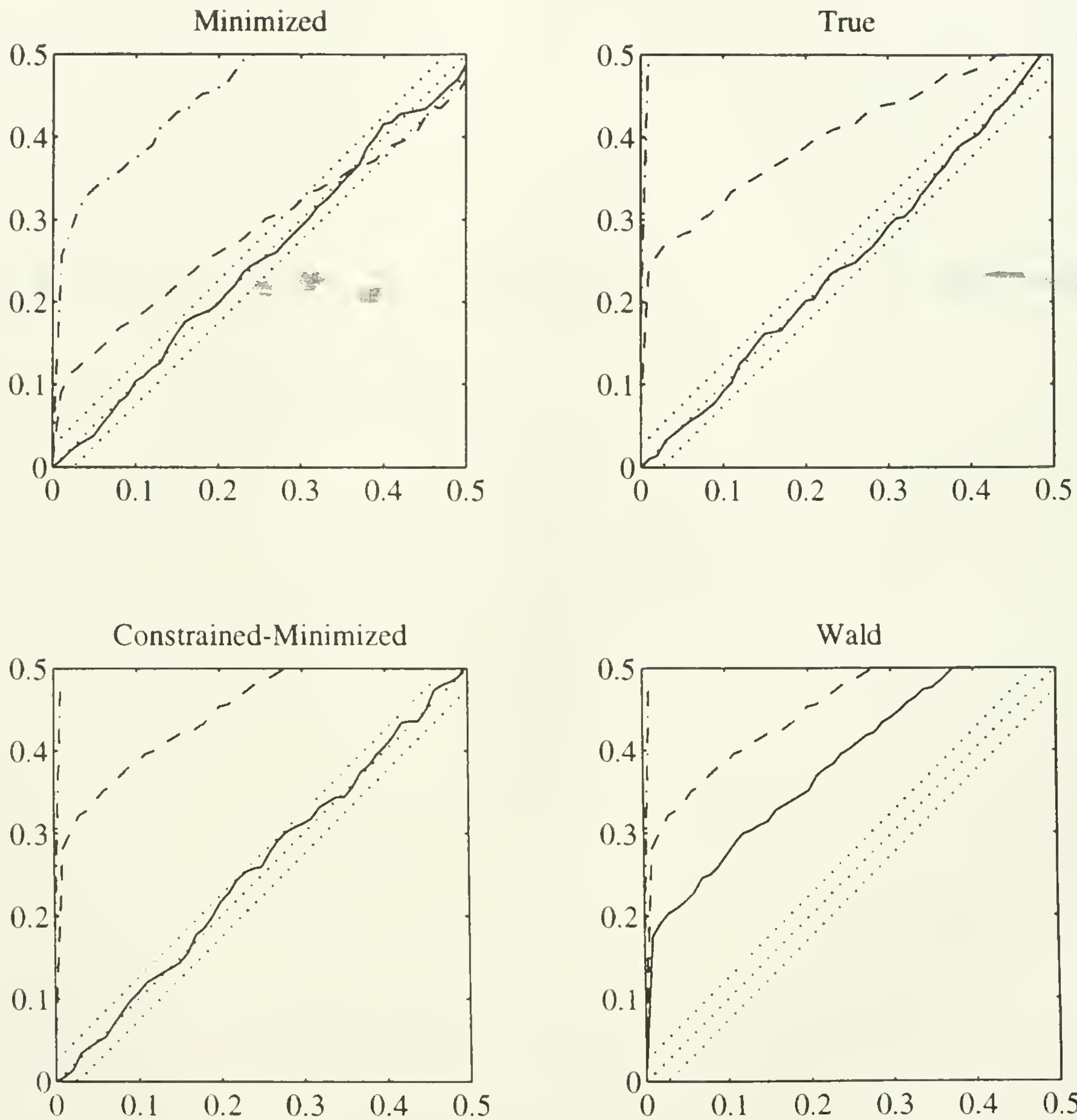


Figure 5.2
Criterion Functions
Monthly Log-Normal Model, Time Averaging

_._._.Iterative, _ _ _ Two-Step, __ Continuous-Updating

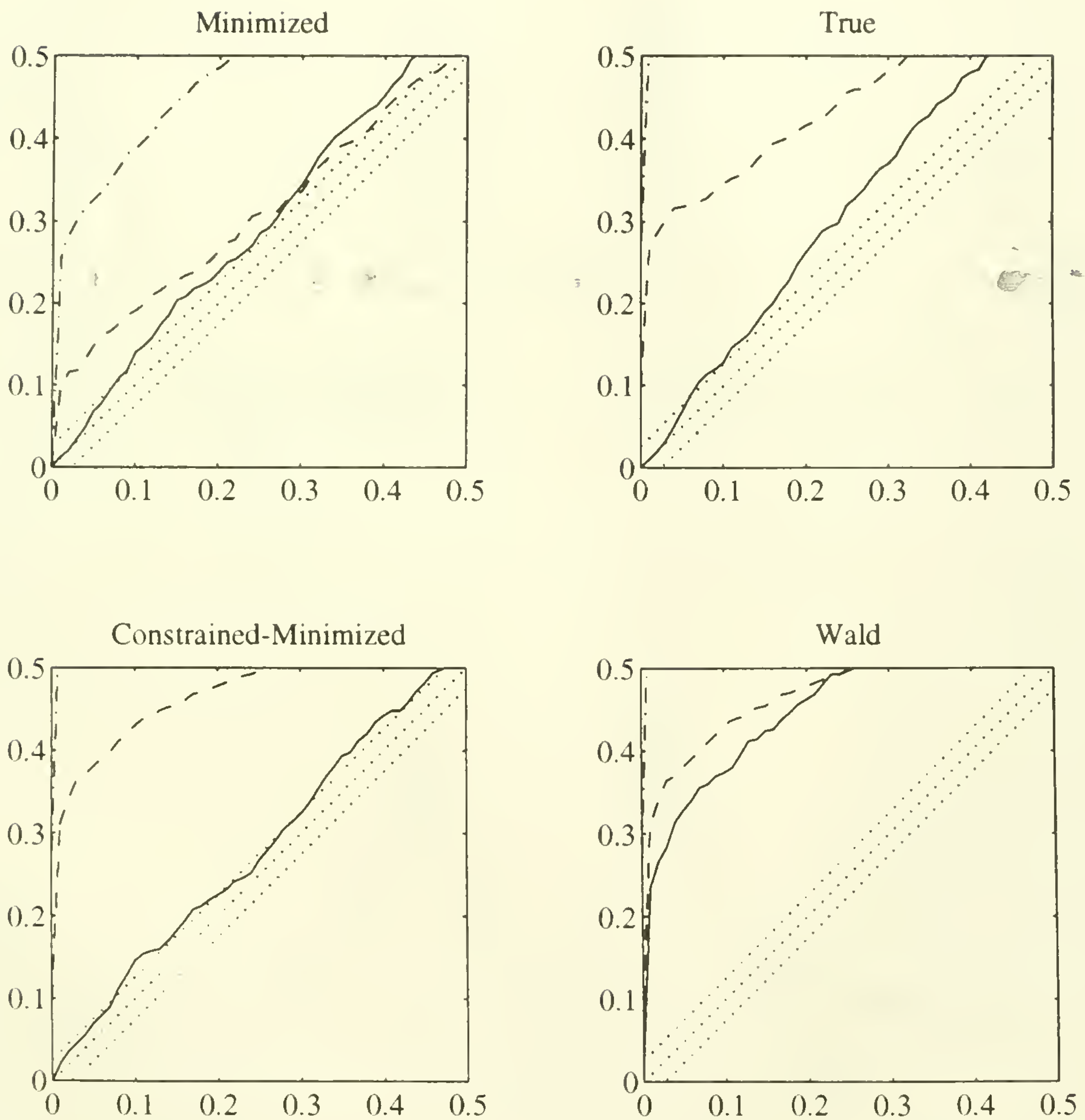


Figure 5.3
Smoothed Distribution of Estimated Angle implied by $(1, \gamma_T)$
Monthly Log-Normal Model, No Time Averaging

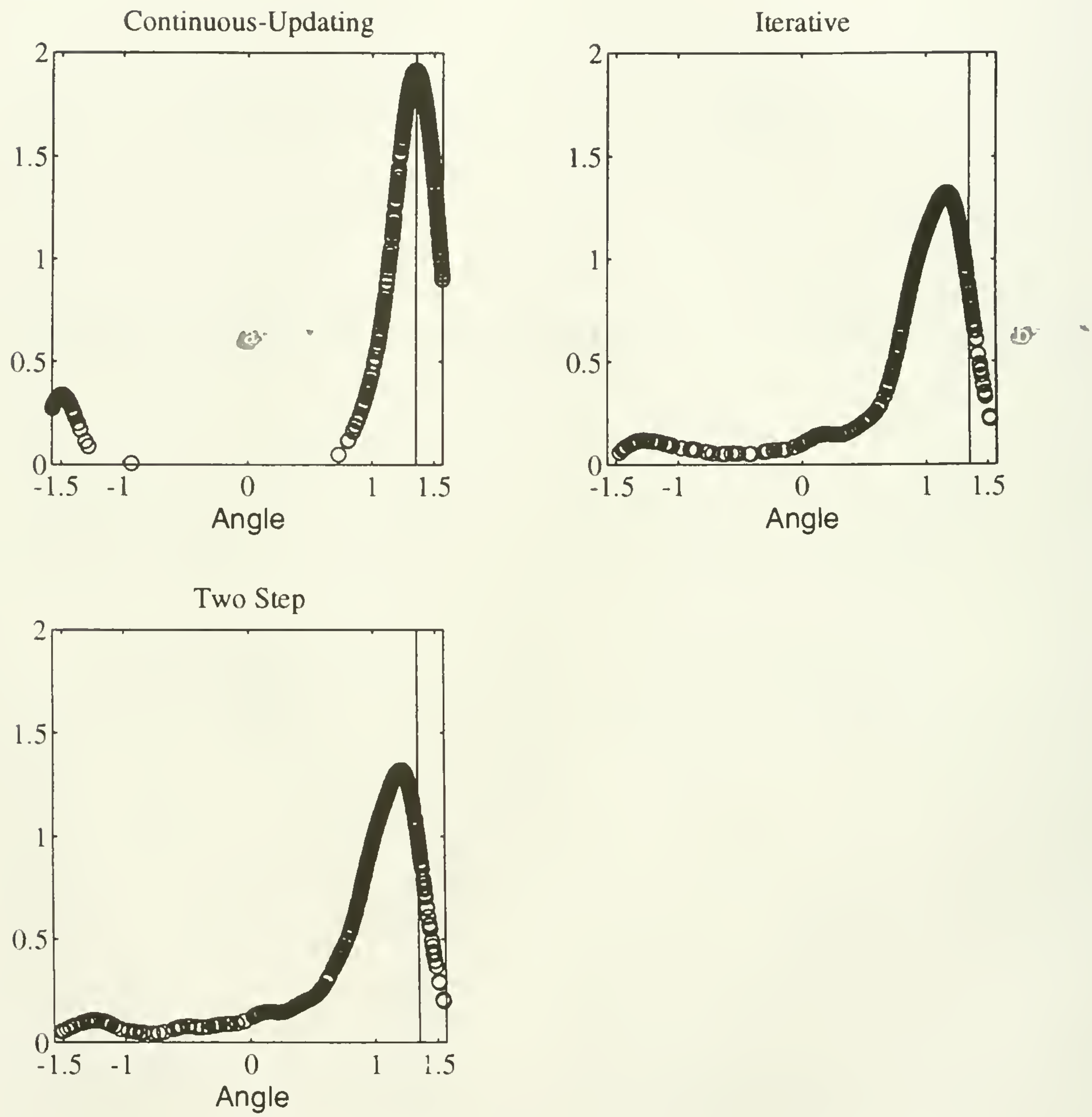


Figure 5.4
Smoothed Distribution of Estimated Angle implied by $(1, \gamma_T)$
Monthly Log-Normal Model, Time Averaging

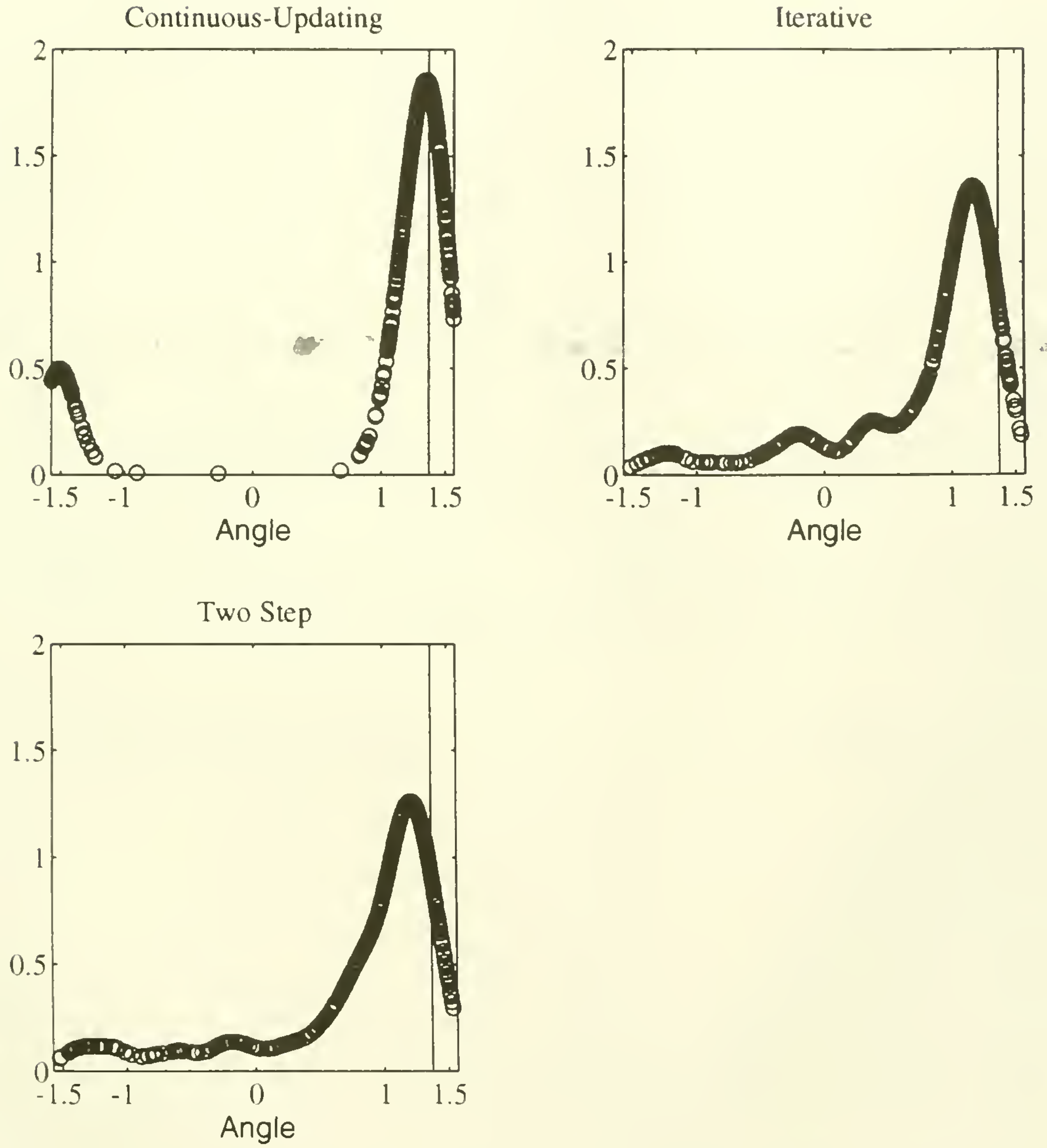


Figure 5.5
Criterion Function for Continuous-Updating Estimator
Monthly Log-Normal Model, No Time Averaging

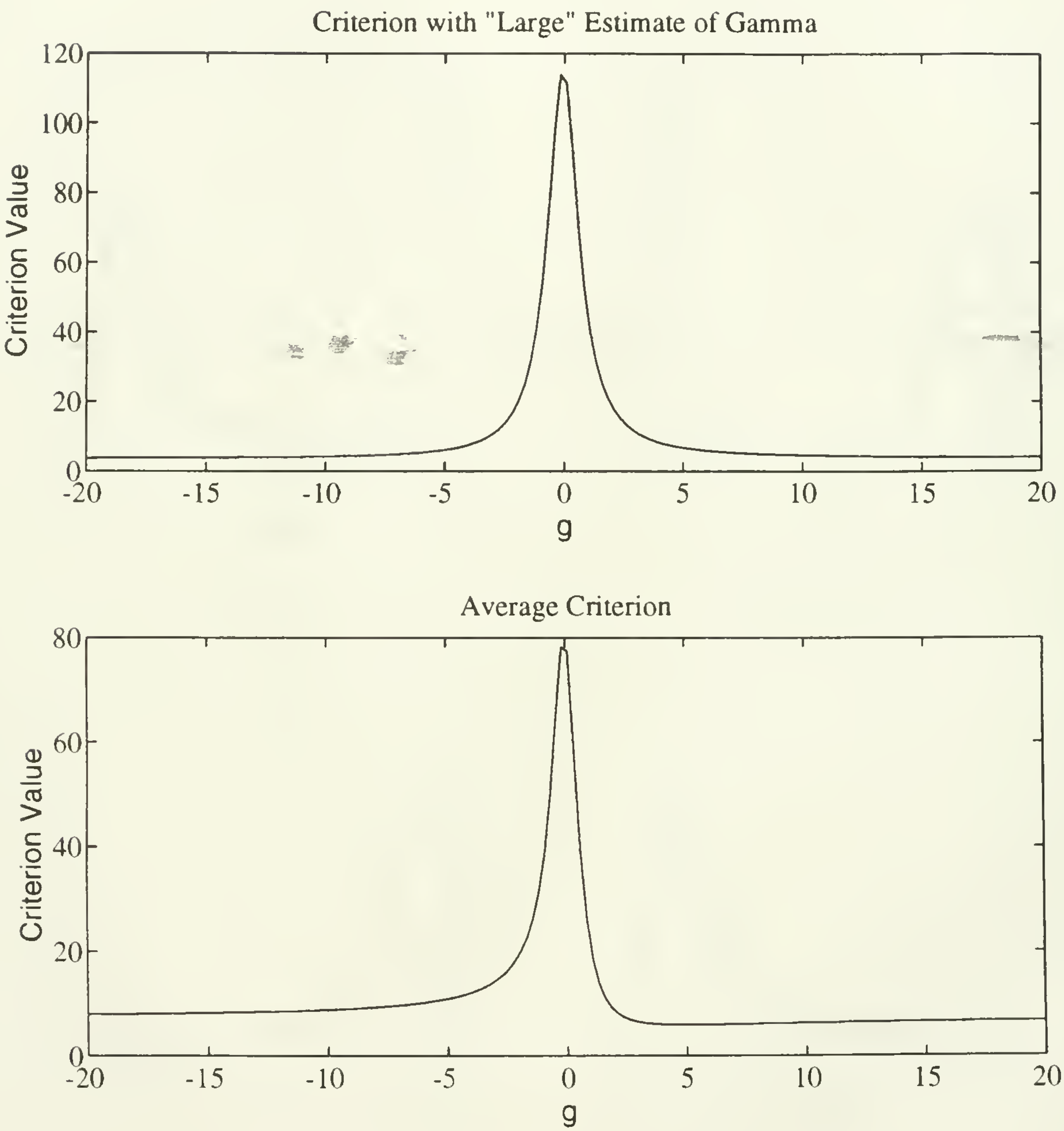


Figure 6.1
Criterion Functions
Annual Markov Chain Model
$\beta = .97$, $\gamma = 1.3$, $\theta = 0$. Moment Conditions M1.

_._._.Iterative, _ _ _ Two-Step, __ Continuous-Updating

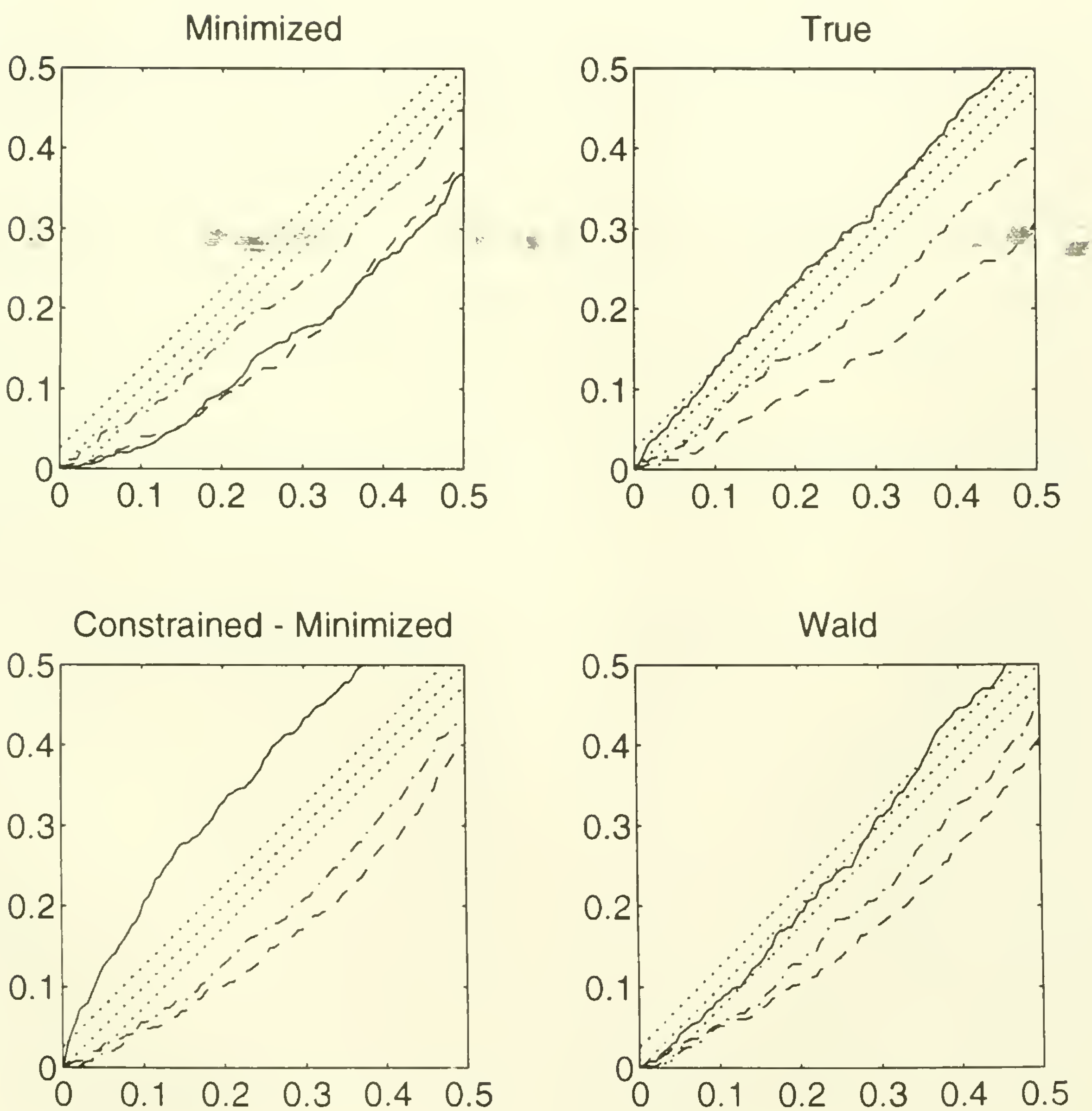


Figure 6.2
Criterion Functions
Annual Markov Chain Model
$\beta = .97$, $\gamma = 1.3$, $\theta = 0$. Moment Conditions M2.

_._._.Iterative, _ _ _ Two-Step, __ Continuous-Updating

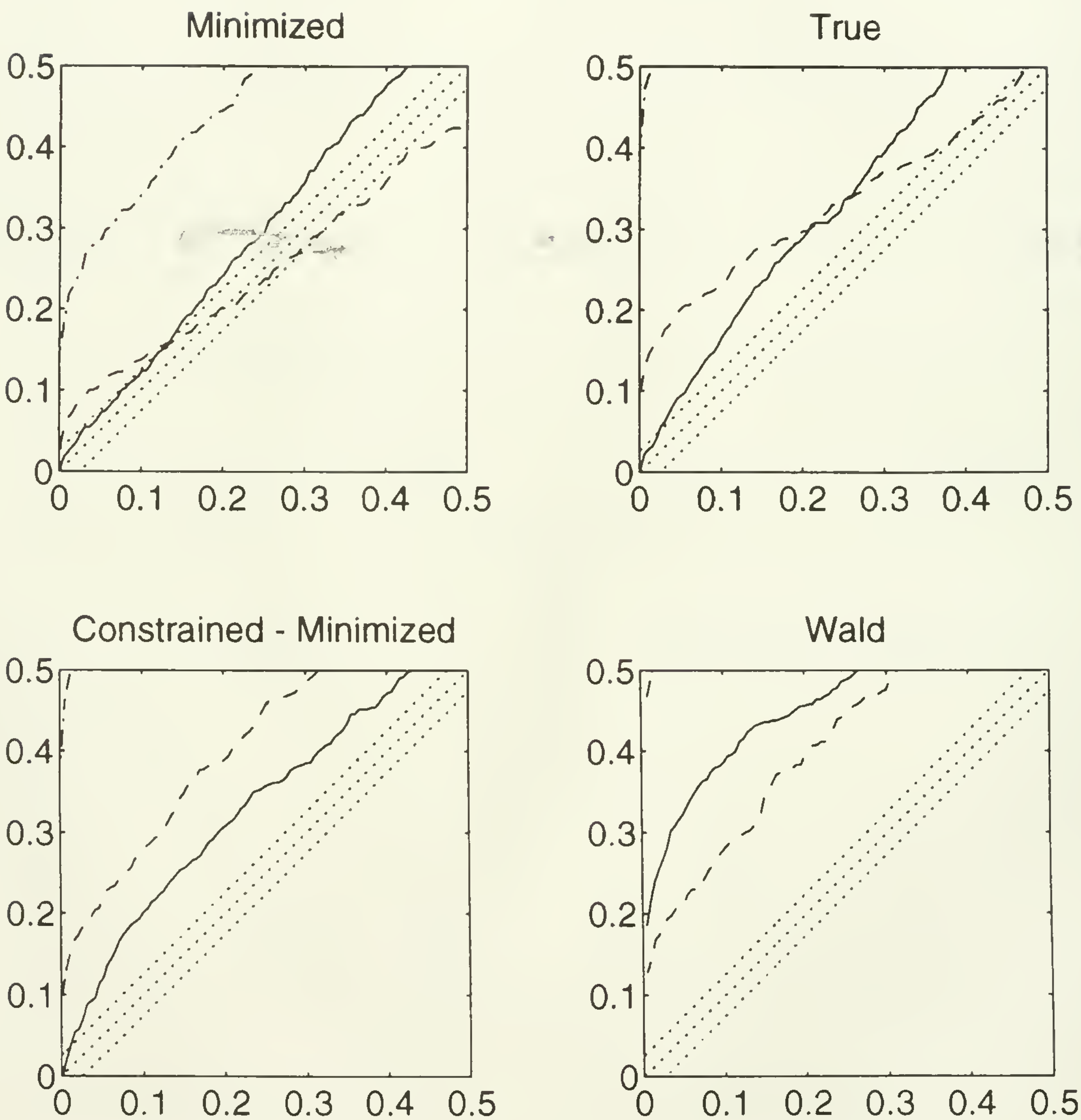


Figure 6.3
Criterion Functions
Annual Markov Chain Model
$\beta = .97$, $\gamma = 1.3$, $\theta = 0$. Moment Conditions M3.

_._._Iterative, _ _ _ Two-Step, __ Continuous-Updating

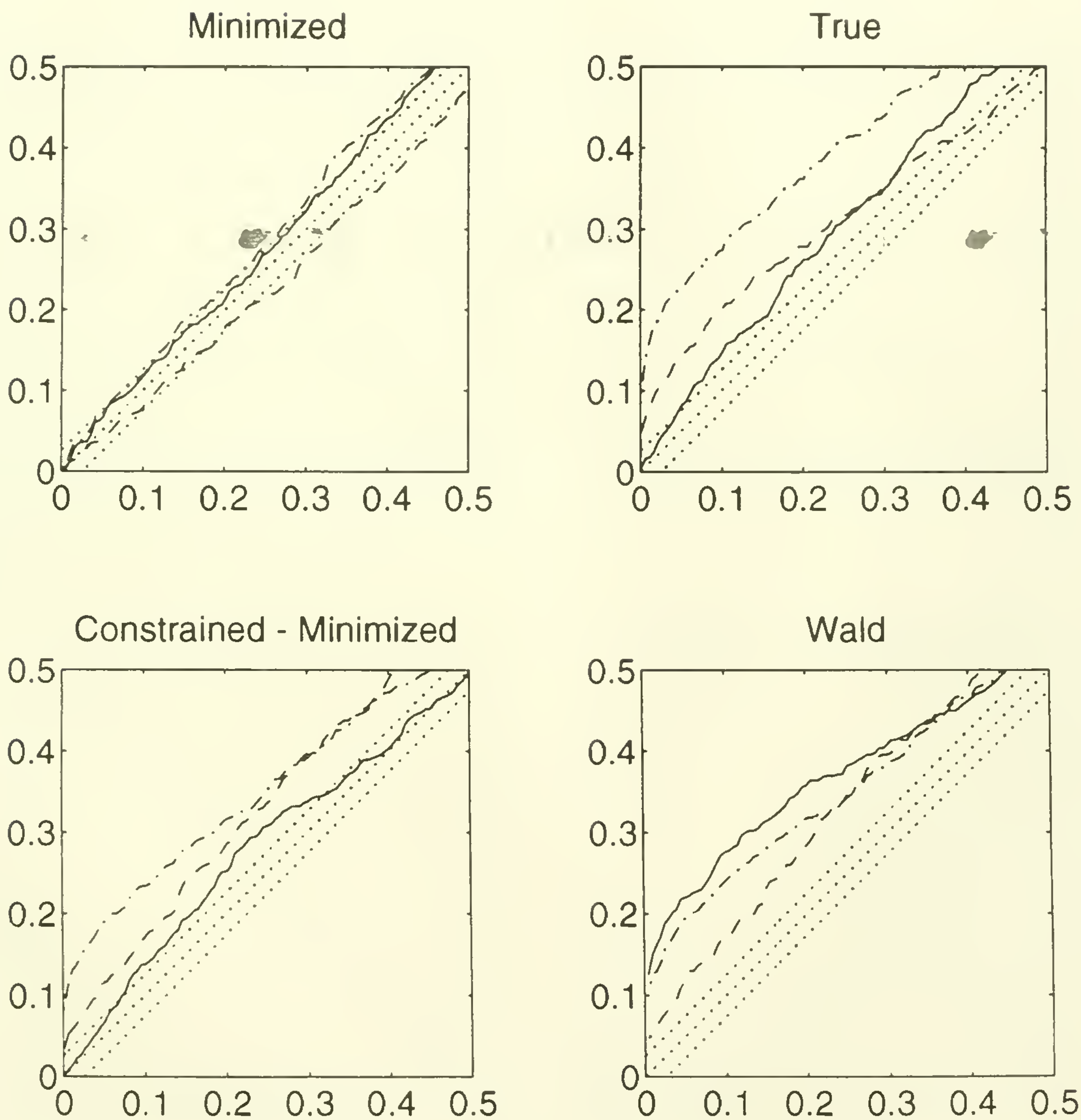


Figure 6.4
Criterion Functions
Annual Markov Chain Model
$\beta = .97$, $\gamma = 1.3$, $\theta = 0$. Moment Conditions M4.

_._._.Iterative, _ _ _ Two-Step, __ Continuous-Updating

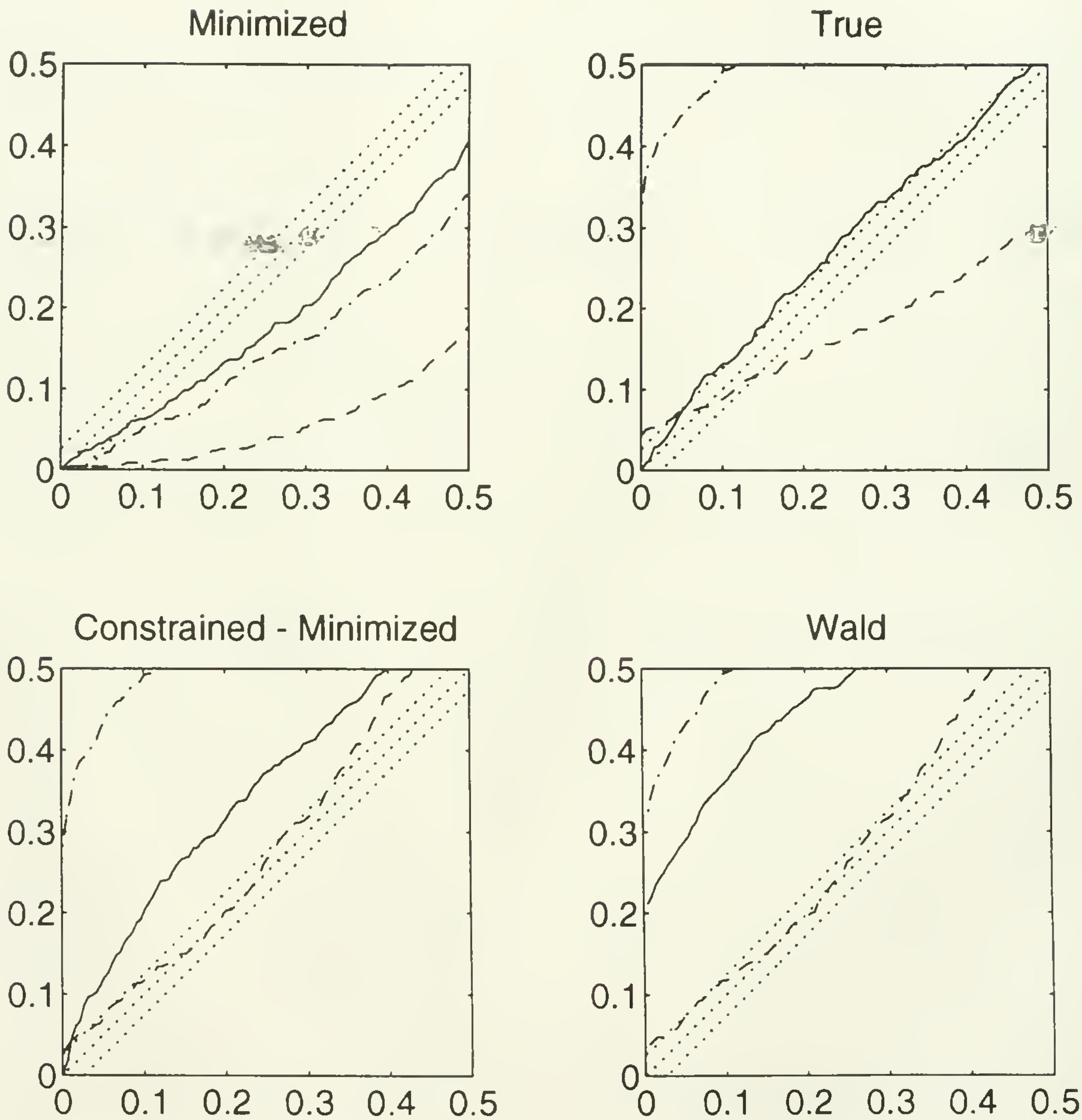


Figure 6.5
Criterion Functions
Annual Markov Chain Model
$\beta = 1.139$, $\gamma = 13.7$, $\theta = 0$. Moment Conditions M3.

_._._.Iterative, _ _ _ Two-Step, __ Continuous-Updating

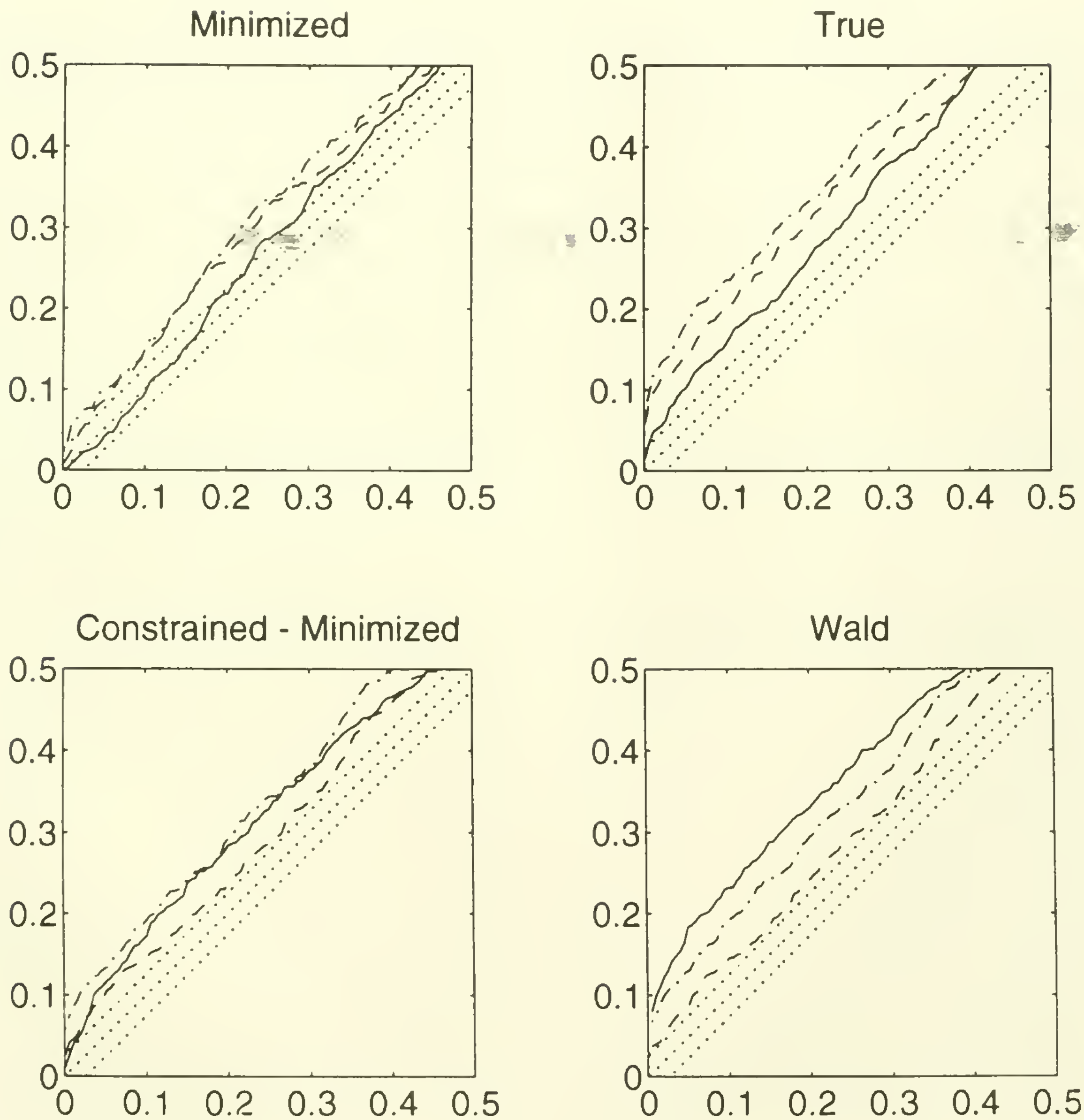


Figure 6.6
Criterion Functions
Annual Markov Chain Model
$\beta = 1.139$, $\gamma = 13.7$, $\theta = 0$. Moment Conditions M4.

_._._.Iterative, _ _ _ Two-Step, __ Continuous-Updating

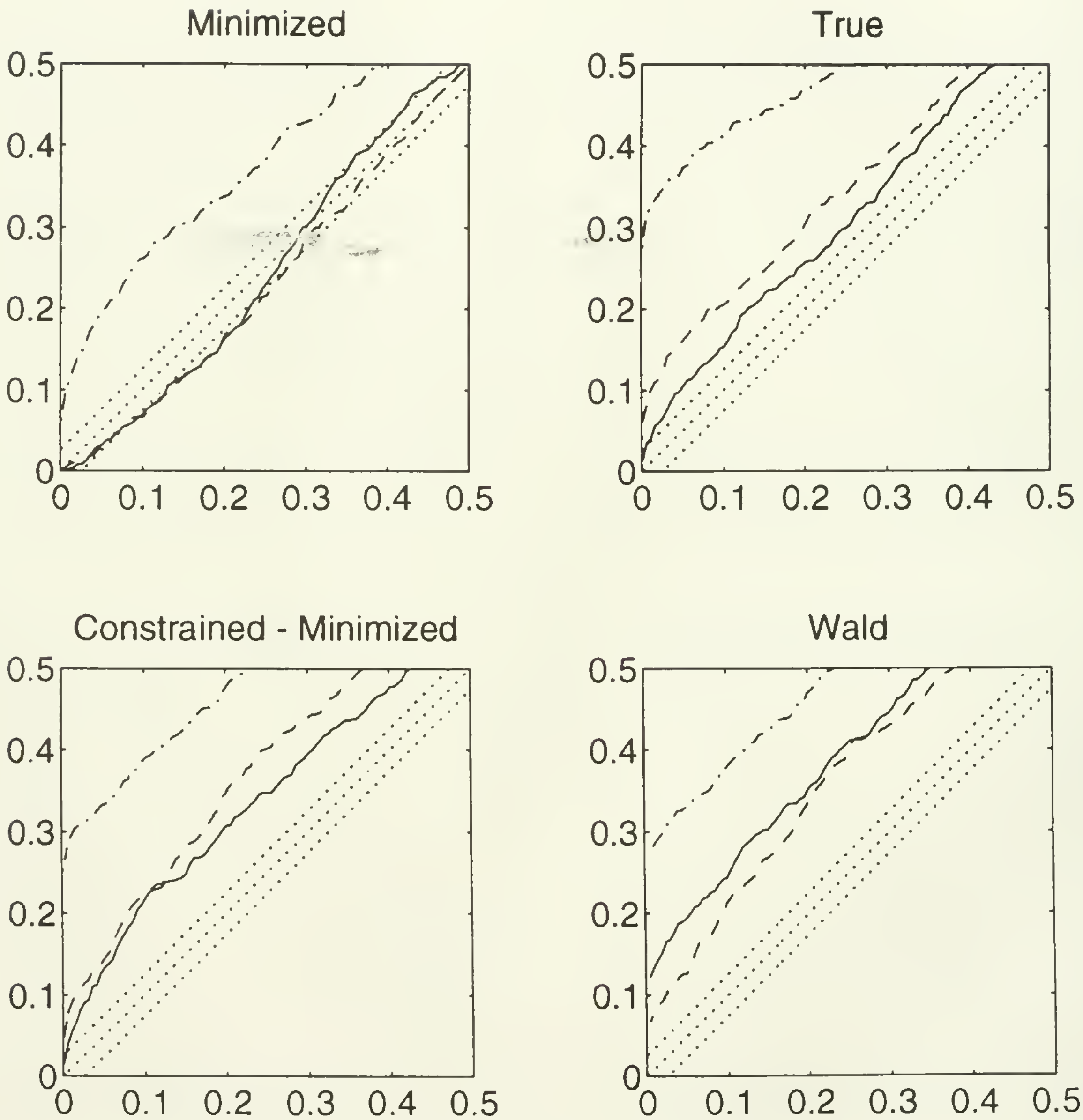


Figure 6.7
Criterion Functions
Monthly Markov Chain Model
$\beta = .97^{1/12}$, $\gamma = 1.3$, $\theta = 0$. Moment Conditions M2.

_._._Iterative, _ _ _ Two-Step, __ Continuous-Updating

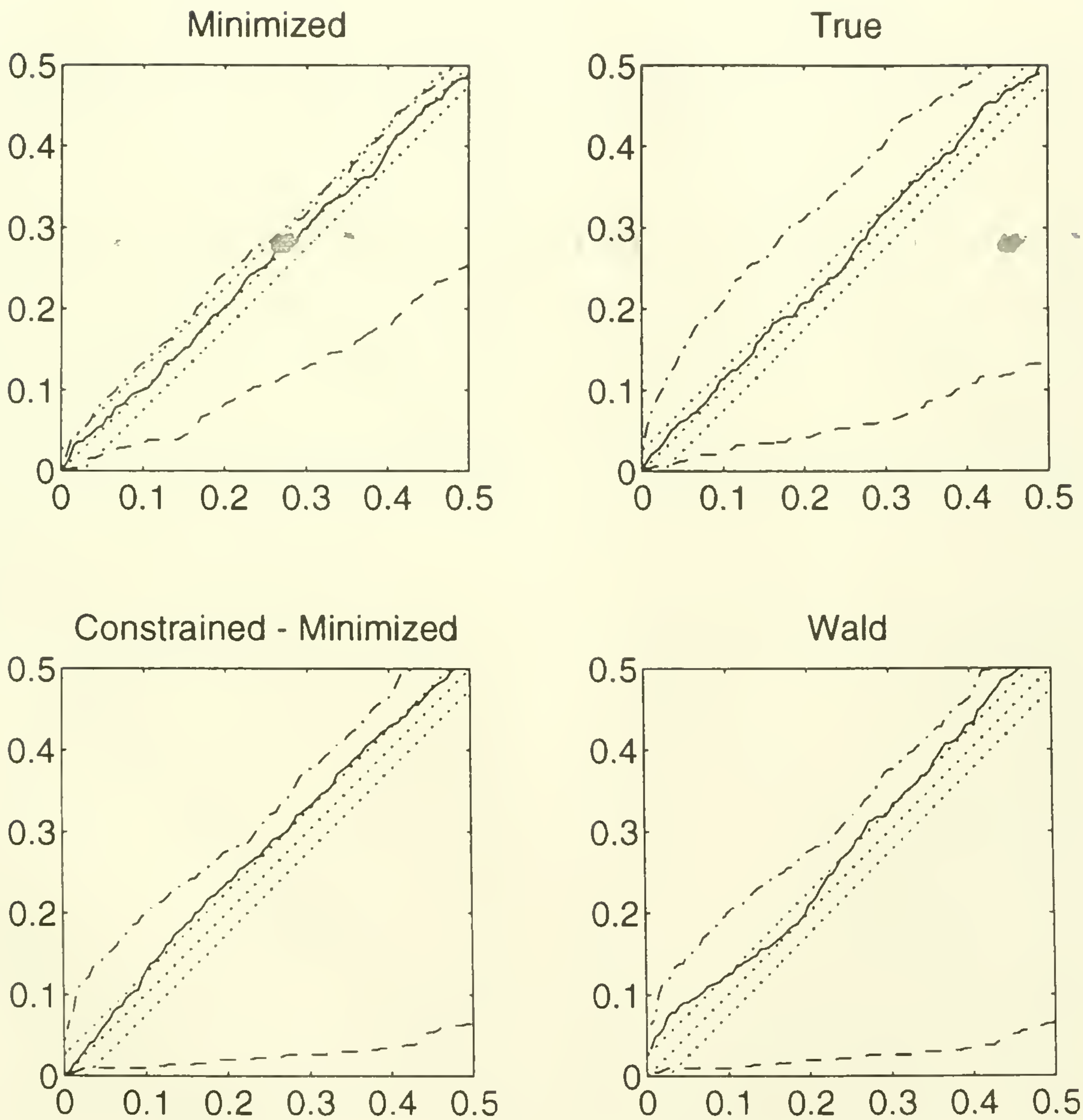


Figure 6.8
Criterion Functions
Monthly Markov Chain Model
$\beta = .97^{1/12}$, $\gamma = 1.3$, $\theta = 0$. Moment Conditions M3.

_._._.Iterative, _ _ _ Two-Step, __ Continuous-Updating

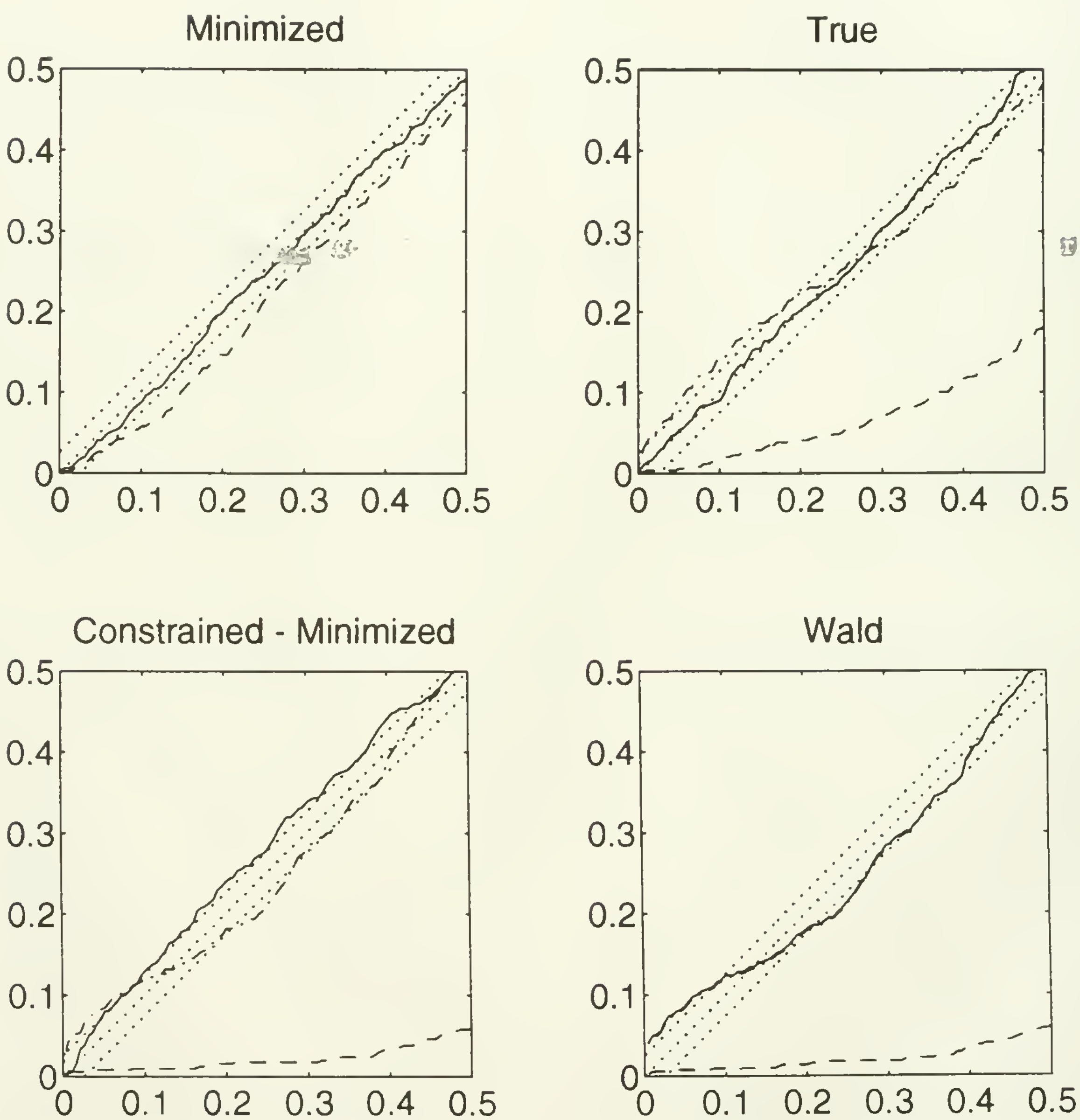


Figure 6.9
Criterion Functions
Annual Markov Chain Model, Continuous-Updating Estimator
$\beta = .97$, $\gamma = 1.3$, Time Nonseparable. Moment Conditions M3.

-.-.-. $\theta = 1/3$, - - - $\theta = 0$, __ $\theta = -1/3$, $\cdots$ $\theta = -2/3$

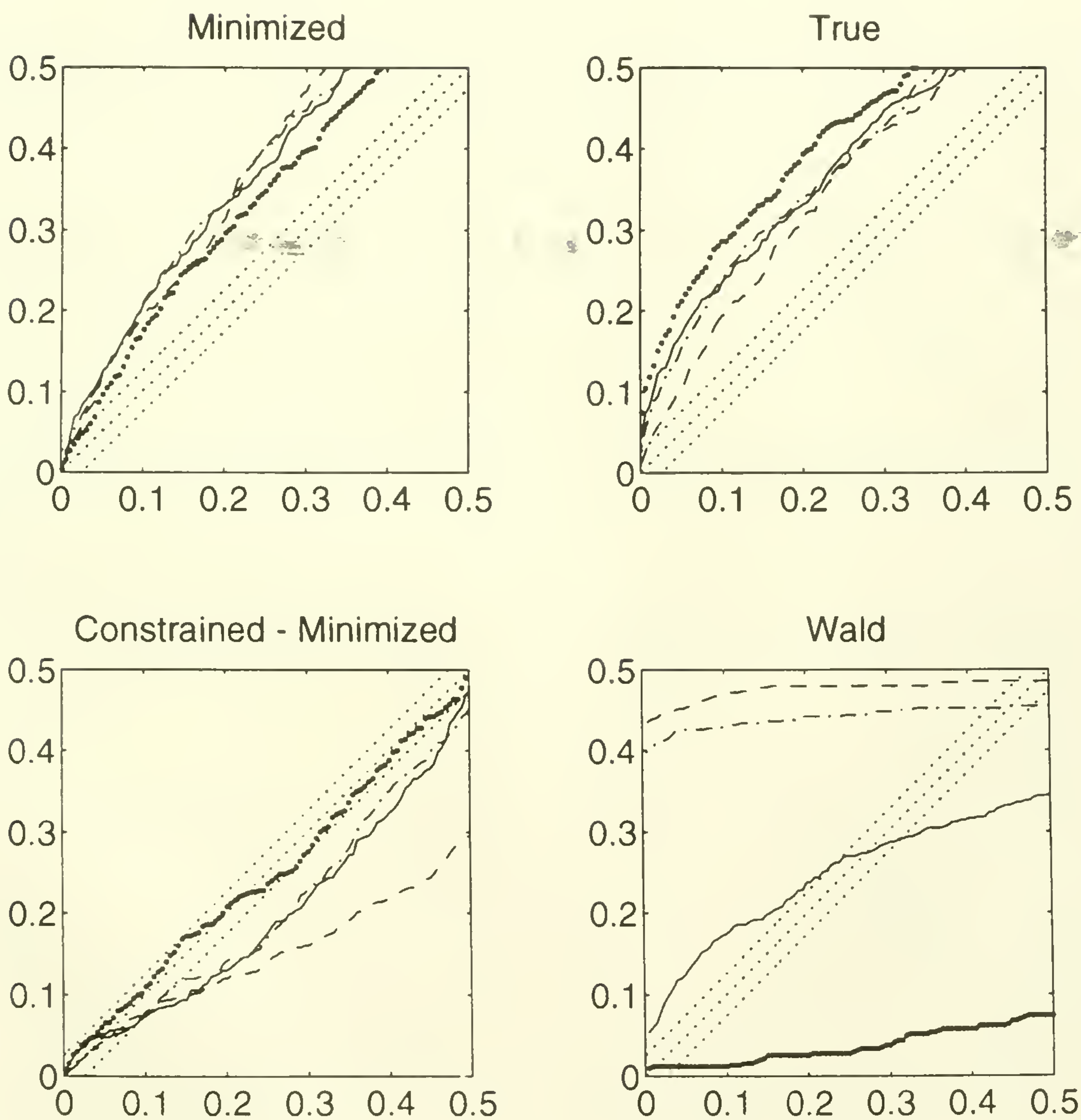
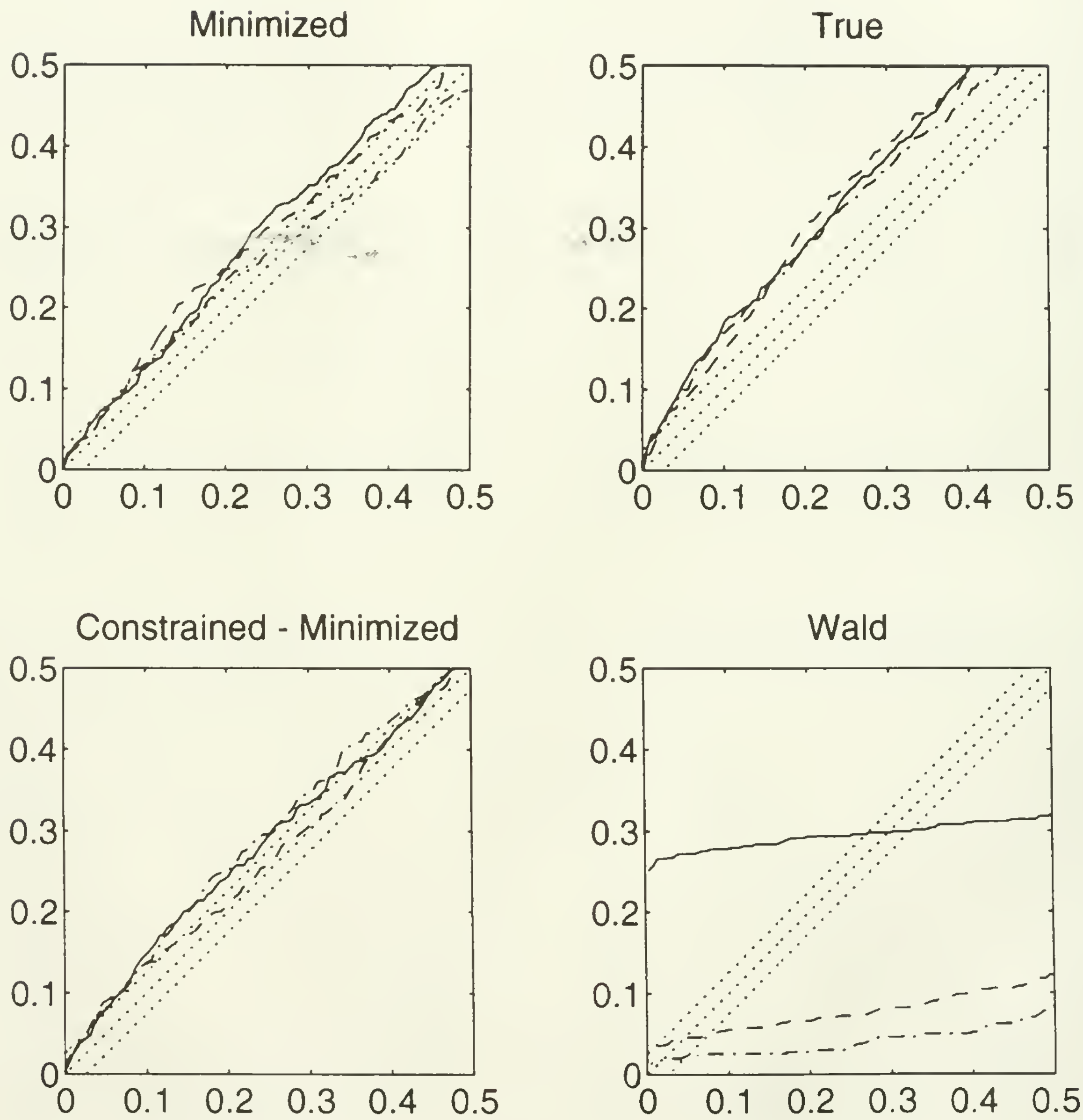


Figure 6.10
Criterion Functions
Monthly Markov Chain Model, Continuous-Updating Estimator
$\beta = .97^{1/12}$, $\gamma = 1.3$, Time Nonseparable. Moment Conditions M2.

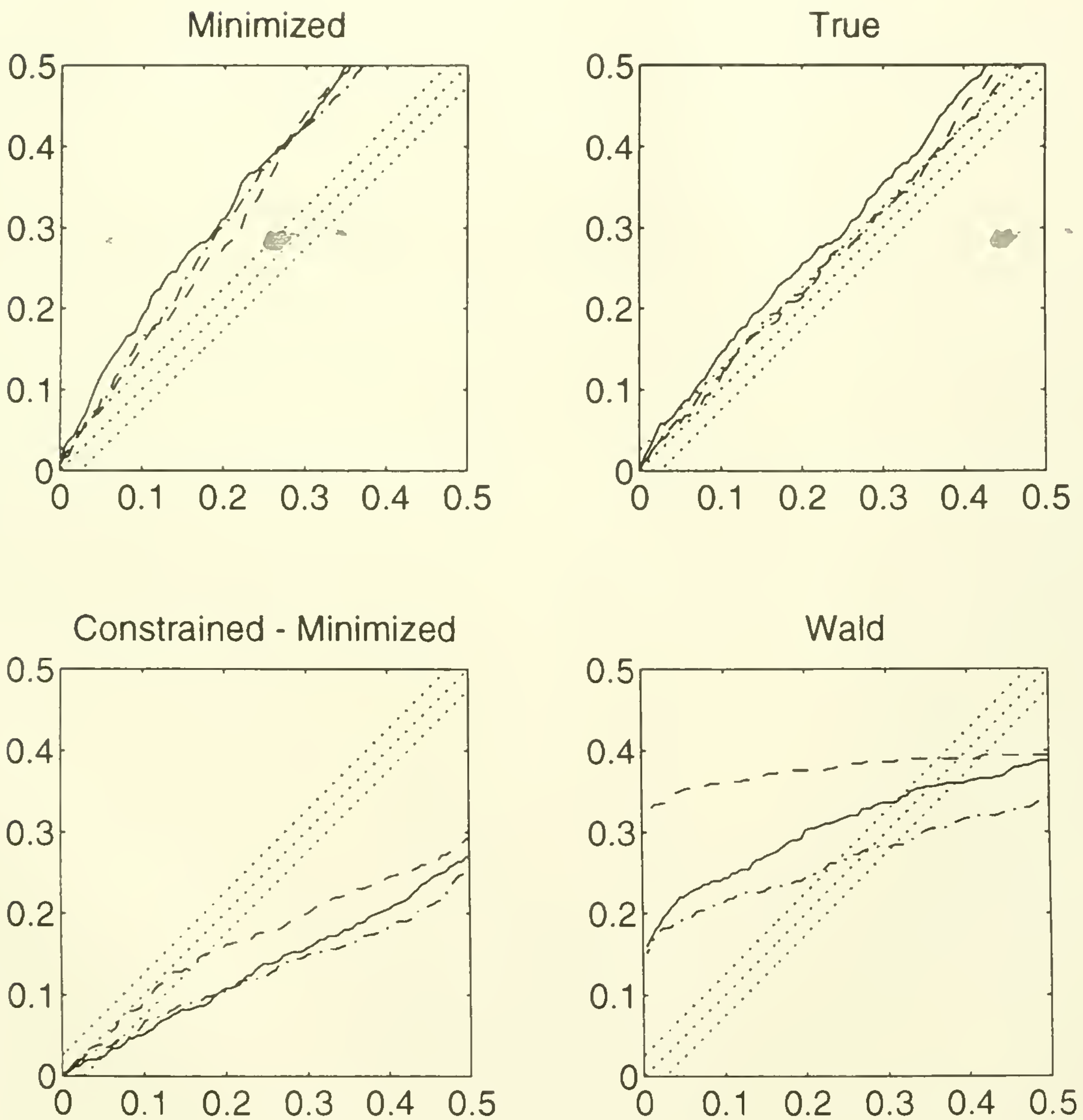
_._._ $\theta = 1/3$, _ _ _ $\theta = 0$, __ $\theta = -1/3$



MIT LIBRARIES
3 9080 00897 4716

Figure 6.11
Criterion Functions
Monthly Markov Chain Model, Continuous-Updating Estimator
$\beta = .97^{1/12}$, $\gamma = 1.3$, Time Nonseparable. Moment Conditions M3.

$\_.\_.\_.\theta = 1/3$, $\_\ \_\ \_\ \theta = 0$, $\_\_\ \theta = -1/3$



2176

Date Due

MAY 29 1998

Lib-26-67