A Simulation Based Specification Test for Diffusion Processes*

Geetesh Bhardwaj[1], Valentina Corradi[2], and Norman R. Swanson[3]

[1]Bates and White, [2]University of Warwick and [3] Rutgers University

this revision: February 2007

Abstract

This paper makes two contributions. First, we outline a simple simulation based framework for constructing conditional distributions for multi-factor and multi-dimensional diffusion processes, for the case where the functional form of the conditional density is unknown. The distributions can be used, for example, to form predictive confidence intervals for time period $t + \tau$, given information up to period t. Second, we use the simulation based approach to construct a test for the correct specification of a diffusion process. The suggested test is in the spirit of the conditional Kolmogorov test of Andrews (1997). However, in the present context the null conditional distribution is unknown and is replaced by its simulated counterpart. The limiting distribution of the test statistic is not nuisance parameter free. In light of this, asymptotically valid critical values are obtained via appropriate use of the block bootstrap. The suggested test has power against a larger class of alternatives than tests that are constructed using marginal distributions/densities, such as those in Aït-Sahalia (1996) and Corradi and Swanson (2005a). The findings of a small Monte Carlo experiment underscore the good finite sample properties of the proposed test, and an empirical illustration underscores the ease with which the proposed simulation and testing methodology can be applied.

JEL classification: C22, C51.
Keywords: block bootstrap, diffusion processes, parameter estimation error, simulated GMM, stochastic volatility.

----

*We would like to thank the editor, Torben Andersen, the associate editor, and two referees for numerous useful suggestions on earlier versions of this paper. Additionally, thanks are due to Marine Carrasco, Mike Pitt, Neil Shephard, Jeff Wooldridge, and seminar participants at the 2005 World Congress of the Econometric Society, the University of Montreal and Michigan State University, for useful many comments.

Geetesh Bhardwaj, Bates White, 12340 El Camino Real, suite 350, San Diego, CA 92130, USA, geetesh. bhardwaj@gmail.com. Valentina Corradi, Department of Economics, University of Warwick, Coventry CV4 7AL, UK, v.corradi@warwick.ac.uk. Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. Corradi gratefully acknowledges ESRC grant RES-000-23-0006, and Swanson acknowledges financial support from a Rutgers University Research Council grant.

1 Introduction

In this paper, we outline a simple simulation based framework for constructing conditional distributions for multi-factor and multi-dimensional diffusion processes, in the case where the functional form of the conditional density is unknown. The estimated conditional distributions can be used, for example, to form predictive confidence intervals for time period $t + \tau$, given information up to period t. Further, we use the simulation based approach to construct tests for the correct specification of a given diffusion model. What distinguishes our approach from that followed by numerous other authors is that we directly evaluate conditional distributions and confidence intervals, rather than focussing on marginal and/or joint distributions when constructing our specification test.

Precedents to our paper include Corradi and Swanson (CS: 2005a), who construct various tests, including a Kolmogorov type test, based on the comparison of the empirical cumulative distribution function and the cumulative distribution function (CDF) implied by the specification of the drift and the variance functions of a diffusion, and Aït-Sahalia (AS: 1996)), who proposes an interesting specification test based on a comparison of the marginal density of the process under the null hypothesis and its kernel counterpart. The AS and CS tests determine whether the drift and variance components of a particular single factor continuous time model are correctly specified, although the CS test is based on the comparison of CDFs, while Aït-Sahalia's is based on the comparison of densities, and uses a nonparametric density estimator. The AS test is thus characterized by a nonparametric rate, while the CS test has a parametric rate. In the case of multi-factor and multi-dimensional models characterized by stochastic volatility, say, the functional form of the invariant density of the return(s) is no longer guaranteed to be given in closed form, upon joint specification of the drift and variance terms, so that the Kolmogorov type test of CS is no longer applicable. To get around this problem, CS compare the empirical *joint* distribution of historical data with the empirical *joint* distribution of (model) simulated data (see also Corradi and Swanson (2007)).

This paper is different from CS in at least two important ways. First, we provide a correct specification test for a diffusion that is based on the evaluation of an estimate of the *conditional* distribution, rather than the joint or marginal distribution. This is a relevant departure from CS, since in the literature on the evaluation of continuous time financial models, the main goal is to construct specification tests based on the transition density associated with a model. In fact, tests based on comparisons of marginal distributions have no power against *iid* alternatives with the same marginal, for example. Second, we outline an easy to implement procedure for constructing confidence intervals for an asset price for a given period, say $t + \tau$. This second feature is important in credit risk management, for example, given that our confidence intervals are predictive intervals whenever the model is estimated using data prior to period $t+1$. On the other hand, when models are estimated using all available data, and when $t + \tau \leq T$, then the methods in this paper are strictly in-sample, so that our specification test is not ex ante in nature, where T is the sample size. For discussion of theoretical issues which arise when our specification test is used in ex ante contexts, the reader is referred to Corradi and Swanson (2006a).

If the functional form for the transition density were known, we could test the hypothesis of correct specification of a diffusion via the probability integral transform approach of Diebold, Gunther and Tay (1998), the cross spectrum approach of Hong (2001), Hong and Li (2004), Hong, Li and Zhao (2004), Thompson (2004), the test of Bai (2003) based on the joint use of a Kolmogorov test and a martingalization method, or via the normality transformation approach of Bontemps and Meddahi (2005) and Duan (2003).

Alternative tests to that proposed in this paper for the case of unknown transition density functions have recently been suggested. For example, Aït-Sahalia, Fan and Peng (2005) outline a test based on the comparison of conditional kernel density estimators and approximations of conditional densities using Hermite polynomials (see also Aït-Sahalia (1999, 2002)). In principle, one can also compare conditional kernel density estimators based on historical and simulated data, along the lines of Altissimo and Mele (2005). One feature of these tests is that they are characterized by nonparametric rates. Our test, on the other hand, is characterized by a parametric rate. This may seem surprising, since it is well known that kernel estimators of conditional distributions converge at nonparametric rates (see e.g. Fan, Yao and Tong (1996) and Hall, Wolff and Yao (1999)). Therefore, tests based on the comparison of conditional distributions of simulated and historical data are in general characterized by nonparametric rates. However, we obtain a parametric rate. This is done by simulating S paths of length τ, all having as common starting value the observable value at time t, say X_t. Now, the empirical distribution of the simulated series provides a $\sqrt{T}-$consistent estimator of the conditional distribution of the null model. We can thus construct a conditional Kolmogorov test, along the lines of Andrews (1997).

Of final note is that the null diffusion model depends on unknown parameters that need to be estimated. Parameters are assumed to be estimated via the simulated generalized method of moments (SGMM) approach of Duffie and Singleton (1993), assuming exact identification. This is because $\sqrt{T}-$consistency does not hold for overidentified (S)GMM estimators of misspecified models, as shown by Hall and Inoue (2003). Additionally, and as is common with the type of test proposed here, limiting distributions are functionals of zero mean Gaussian processes with covariance kernels that reflect the contribution of parameter estimation error (PEE). Thus, limiting distributions are not nuisance parameter free and critical values cannot be tabulated. (Note that in the special case of testing for normality, Bontemps and Meddahi (2005) provide a GMM type test based on moment conditions that is robust to parameter estimation error.) In light of this fact, we provide valid asymptotic critical values via an extension of the empirical process version of the block bootstrap which properly captures the contribution of PEE, for the case where parameters are estimated via SGMM. Our bootstrap results stem in part from the fact that when simulation error is negligible (i.e. the simulated sample grows faster than the historical sample), and given exact identification, the results of Goncalves and White (2004) for QMLE estimators extend to SGMM estimators.

The potential usefulness of our proposed bootstrap based tests is examined via a series of Monte Carlo experiments using discrete samples of 400 and 800 observations, and based on the use of bootstrap critical values constructed using as few as 100 replications. For the larger sample, rejection rates under the null are quite close to nominal values, and rejection rates under the alternative are generally high. Additionally, an empirical illustration is provided which underscores the ease with which one can apply our simulation and testing methodology.

The rest of the paper is organized as follows. In Section 2, we outline a setup which is appropriate for discussing our simulation and testing methodology, in the context of one-dimensional diffusion processes. In Section 3, we discuss a simple approach to simulation of conditional distributions and confidence intervals. Section 4 outlines our specification test, and Section 5 extends all results to the case of stochastic volatility models. Section 6 contains the results of our small Monte Carlo experiment, and discusses the findings from an empirical illustration based on the Eurodollar deposit rate. Concluding remarks are gathered in Section

7, and all proofs are collected in an appendix.

2 Setup

In this section we outline the set-up for the case of one-dimensional diffusion processes. All results carry through to the more complicated cases of multi-dimensional and multi-factor stochastic volatility models, as outlined in Section 5.

Let $X(t)$, $t \geq 0$, be a one-dimensional diffusion process solution to the following stochastic differential equation:

$$dX(t) = \mu_0(X(t), \theta_0)dt + \sigma_0(X(t), \theta_0)dW(t), \tag{1}$$

where $\theta_0 \in \Theta$, $\Theta \subset \Re^p$, and Θ is a compact set. In general, assume that the model which is specified and estimated is the same as above, but with θ_0 replaced by its pseudo true analog, $\theta^\dagger$. Namely, we consider models of the form:

$$dX(t) = \mu(X(t), \theta^\dagger)dt + \sigma(X(t), \theta^\dagger)dW(t), \tag{2}$$

Thus, correct specification of the diffusion process corresponds to $\mu(\cdot,\cdot) = \mu_0(\cdot,\cdot)$ and $\sigma(\cdot,\cdot) = \sigma_0(\cdot,\cdot)$. Note that the drift and variance terms ($\mu(\cdot)$ and $\sigma^2(\cdot)$, respectively) uniquely determine the stationary density, say $f(x, \theta^\dagger)$, associated with the invariant probability measure of the above diffusion process (see e.g. Karlin and Taylor (1981), pp. 241). In particular:

$$f(x, \theta^\dagger) = \frac{c(\theta^\dagger)}{\sigma^2(x, \theta^\dagger)} \exp\left(\int^x \frac{2\mu(v, \theta^\dagger)}{\sigma^2(v, \theta^\dagger)} dv\right), \tag{3}$$

where $c(\theta^\dagger)$ is a constant ensuring that the density integrates to one. However, knowledge of the drift and variance terms does not ensure knowledge of a closed functional form for the transition density.

Now, suppose that we observe a discrete sample (skeleton) of T observations, say $(X_1, X_2, \ldots, X_T)'$, from the underlying diffusion. Furthermore, suppose that we use these sample data in conjunction with a simulated "path" in order to construct an estimator of θ, say $\widehat{\theta}_{T,N,h}$, where N denotes the simulation path length and h is the discretization interval. For the case in which the moment conditions can be written in closed form, and so there is no need of simulations, we have that $\widehat{\theta}_{T,N,h} = \widehat{\theta}_T$. Finally, note that we use the notation $X(t)$ to denote the continuous time process, and the notation X_t to denote the skeleton (i.e. the discrete sample). In light of this, let $X^\theta_{t,h}$ denote pathwise simulated data, constructed using some $\theta \in \Theta$, and discrete interval h, and sampled at the same frequency of the data.

Assume that $\widehat{\theta}_{T,N,h}$ is the simulated generalized method of moments (SGMM) estimator, which is defined as:

$$\begin{aligned}\widehat{\theta}_{T,N,h} &= \arg\min_{\theta\in\Theta} \left(\frac{1}{T}\sum_{t=1}^{T} g(X_t) - \frac{1}{N}\sum_{t=1}^{N} g(X^\theta_{t,h})\right)' W_T \left(\frac{1}{T}\sum_{t=1}^{T} g(X_t) - \frac{1}{N}\sum_{t=1}^{N} g(X^\theta_{t,h})\right) \\ &= \arg\min_{\theta\in\Theta} G_{T,N,h}(\theta)' W_T G_{T,N,h}(\theta), \end{aligned} \tag{4}$$

where g denotes a vector of p moment conditions, $\Theta \subset \Re^p$ (so that we have as many moment conditions as parameters), and $X^\theta_{t,h} = X^\theta_{[Kth/N]}$, with $N = Kh$. Typically the p moments conditions are based on the difference between sample moments of historical and simulated data or, between sample moments and model implied moments, whenever the latter are known in closed form. Finally, W_T is the inverse of a heteroskedasticity and autocorrelation (HAC) robust covariance matrix estimator. That is:

$$W_T^{-1} = \frac{1}{T}\sum_{\nu=-l_T}^{l_T} w_\nu \sum_{t=\nu+1+l_T}^{T-l_T}\left(g(X_t)-\frac{1}{T}\sum_{t=1}^{T}g(X_t)\right)\left(g(X_{t-\nu})-\frac{1}{T}\sum_{t=1}^{T}g(X_t)\right)', \tag{5}$$

where $w_v = 1 - v/(l_T+1)$. In order to construct simulated estimators, we require simulated sample paths. If we use a Milstein scheme (see e.g. Pardoux and Talay (1985)), then:

$$\begin{aligned} X_{kh}^\theta - X_{(k-1)h}^\theta &= b(X_{(k-1)h}^\theta,\theta)h + \sigma(X_{(k-1)h}^\theta,\theta)\epsilon_{kh} - \frac{1}{2}\sigma(X_{(k-1)h}^\theta,\theta)'\sigma(X_{(k-1)h}^\theta,\theta)h \\ &\quad + \frac{1}{2}\sigma(X_{(k-1)h}^\theta,\theta)'\sigma(X_{(k-1)h}^\theta,\theta)\epsilon_{kh}^2, \end{aligned} \tag{6}$$

where $\left(W_{kh}-W_{(k-1)h}\right)=\epsilon_{kh}\overset{iid}{\sim}N(0,h)$, $k=1,\ldots,K$, $Kh=N$, and $\sigma(\cdot,\cdot)'$ is defined below (see Assumption A). Hereafter, X_{kh}^θ denotes the values of the diffusion at time kh, simulated under θ, and with a discrete interval equal to h, and so is a fine grain analogous of $X_{t,h}^\theta$. Further, the pseudo true value, $\theta^\dagger$, is defined to be:

$$\theta^\dagger = \arg\min_{\theta\in\Theta} G_\infty(\theta)'W_0G_\infty(\theta),$$

where $G_\infty(\theta)'W_0G_\infty(\theta) = \text{plim}_{N,T\to\infty,h\to0}\, G_{T,N,h}(\theta)'W_TG_{T,N,h}(\theta)$; and where $\theta^\dagger=\theta_0$, if the model is correctly specified. Note that the reason why we limit our attention to the exactly identified case is that this ensures that $G_\infty(\theta^\dagger)=0$, even when the model used to simulate the diffusion is misspecified, in the sense of differing from the underlying DGP (see e.g. Hall and Inoue (2003) for discussion of the asymptotic behavior of misspecified overidentified GMM estimators). Note also that first order conditions imply that:

$$\nabla_\theta G_\infty(\theta^\dagger)'W^\dagger G_\infty(\theta^\dagger)=0.$$

However, in the case for which the number of parameters and the number of moment conditions is the same, $\nabla_\theta G_\infty(\theta^\dagger)'W^\dagger$ is invertible, and so the first order conditions also imply that $G_\infty(\theta^\dagger)=0$.

In the sequel, we shall rely on the following assumption.

Assumption A:

(i) $X(t), t\in\Re^+$, is a strictly stationary, geometric ergodic diffusion.

(ii) $\mu(\cdot,\theta^\dagger)$ and $\sigma(\cdot,\theta^\dagger)$, as defined in (2), are twice continuously differentiable. Also, $\mu(\cdot,\cdot), \mu(\cdot,\cdot)', \sigma(\cdot,\cdot)$, and $\sigma(\cdot,\cdot)'$ are Lipschitz, with Lipschitz constant independent of θ, where $\mu(\cdot,\cdot)'$ and $\sigma(\cdot,\cdot)'$ denote derivatives with respect to the first argument of the function.

(iii) For any fixed h and $\theta\in\Theta$, X_{kh}^θ is geometrically ergodic and strictly stationary.

(iv) $W_T \overset{a.s.}{\to} W_0 = \sum_0^{-1}$, where, $\sum_0 = \sum_{j=-\infty}^{\infty} E\left((g(X_1)-E(g(X_1)))(g(X_{1+j})-E(g(X_{1+j})))'\right)$.

(v) Hereafter, let $\nabla_\theta X_{t,h}^\theta$ be the vector of derivatives with respect to θ of $X_{t,h}^\theta$. For $\theta\in\Theta$ and for all h, $||g(X_{t,h}^\theta)||_{2+\delta} < C < \infty$, $g(X_{t,h}^\theta)$ is Lipschitz, uniformly on Θ, $\theta\to E(g(X_{t,h}^\theta))$ is continuous, and $g(X_t), g(X_{t,h}^\theta), \nabla_\theta X_{t,h}^\theta$ are $2r-$dominated (the last two also on Θ) for $r>3/2$.

(vi) Unique identifiability: $G_\infty(\theta^\dagger)'W_0G_\infty(\theta^\dagger) < G_\infty(\theta)'W_0G_\infty(\theta)$, $\forall\theta\neq\theta^\dagger$.

(vii) $\widehat{\theta}_{T,N,h}$ and $\theta^\dagger$ are in the interior of Θ, $g(X_t^\theta)$ is twice continuously differentiable in the interior of Θ; and $D^\dagger = E(\partial g(X_t^\theta)/\partial\theta|_{\theta=\theta^\dagger})$ exists and is of full rank, p.

Assumption A is rather standard. A(ii) ensures that under both hypotheses there is a unique solution to the stochastic differential equation in (2). A(i) and A(iii)-A(vii) ensure consistency and asymptotic normality of the SGMM estimator.

All results stated in the sequel rely on the following Lemma.

Lemma 1: Let Assumption A hold. Assume that $T, N \to \infty$. Then, if $h \to 0$, $T/N \to 0$, and $h^2 T \to 0$, then:

$$\sqrt{T}\left(\widehat{\theta}_{T,N,h} - \theta^{\dagger}\right) \xrightarrow{d} N(0, \left(D^{\dagger\prime} W_0 D^{\dagger}\right)^{-1}),$$

where W_0 and $D^{\dagger}$ are defined in Assumption A(iv) and A(vii). The above normality result for the SGMM estimator can be extended in a straightforward manner to the case of the EMM estimator of Gallant and Tauchen (1996) (see also the sequential partial indirect inference approach of Dridi, Guay and Renault (2006)).

3 Simulated Conditional Distributions

In this section, we outline how to construct in-sample $\tau-$step ahead simulated conditional distributions, when the functional form of the conditional distribution is not known in closed form. Conditional confidence interval construction follows immediately, and is discussed in Section 6. Let S be the number of simulated paths. Then, for $s = 1, ...S$, $t \geq 1$, and $k = 1, ..., \tau/h$, define:

$$\begin{aligned} & X_{s,t+kh}^{\widehat{\theta}_{T,N,h}} - X_{s,t+(k-1)h}^{\widehat{\theta}_{T,N,h}} \\ = \; & \mu(X_{s,t+(k-1)h}^{\widehat{\theta}_{T,N,h}}, \widehat{\theta}_{T,N,h})h + \sigma(X_{s,t+(k-1)h}^{\widehat{\theta}_{T,N,h}}, \widehat{\theta}_{T,N,h})\epsilon_{s,t+kh} \\ & -\frac{1}{2}\sigma(X_{s,t+(k-1)h}^{\widehat{\theta}_{T,N,h}}, \widehat{\theta}_{T,N,h})'\sigma(X_{s+(k-1)h}^{\widehat{\theta}_{T,N,h}}, \widehat{\theta}_{T,N,h})h \\ & +\frac{1}{2}\sigma(X_{s,t+(k-1)h}^{\widehat{\theta}_{T,N,h}}, \widehat{\theta}_{T,N,h})'\sigma(X_{s,t+(k-1)h}^{\widehat{\theta}_{T,N,h}}, \widehat{\theta}_{T,N,h})\epsilon_{s,t+kh}^2, \end{aligned} \tag{7}$$

where $\epsilon_{s,t+kh} \overset{iid}{\sim} N(0, h)$. That is, we simulate S paths of length τ, with τ finite, all having the same starting value, X_t. This allows for the preservation of the starting value effect on the finite length simulation paths. It should be stressed, however, that the simulated diffusion is ergodic. Thus, the effect of the starting value approaches zero at an exponential rate, as $\tau \to \infty$. Hereafter, $X_{s,t+kh}^{\widehat{\theta}_{T,N,h}}$ denotes the value for the diffusion at time $t + kh$, at simulation s, using parameters $\widehat{\theta}_{T,N,h}$, a discrete interval equal to h, and using, as a starting value at time t, X_t; $X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}$ is defined analogously, by setting $kh = \tau$.

Now, for any given starting value, X_t, the simulated randomness is assumed to be independent across simulations, so that $E(\epsilon_{s,t+kh}\epsilon_{j,t+kh}) = 0$, for all $s \neq j$. On the other hand, it is important to retain the same simulated randomness across different starting values (i.e. use the same set of random errors for each starting value), so that $E(\epsilon_{s,t+kh}\epsilon_{s,l+kh}) = h$, for any t, l.

Note that the test discussed in the next section requires the construction of multiple paths of length τ, for a number of different starting values, and hence we have attempted to underscore the importance of keeping the random errors used in the construction of each set of paths for each starting value the same.

As an estimate for the distribution, at time $t+\tau$, conditional on X_t, define:

$$\widehat{F}_\tau(u|X_t,\widehat{\theta}_{T,N,h}) = \frac{1}{S}\sum_{s=1}^{S} 1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\}$$

As stated in Proposition 2 below, if the model is correctly specified, then $\frac{1}{S}\sum_{s=1}^{S} 1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\}$ provides a consistent estimate of the "true" conditional distribution. A related approach for approximating distribution functions has been suggested by Thompson (2004), in the context of tests for the correct specification of term structure models. Hereafter, let $F_\tau(u|X_t,\theta^\dagger)$ denote the conditional distribution of $X_{t+\tau}^{\theta^\dagger}$, given $X_t^{\theta^\dagger}=X_t$, i.e. $F_\tau(u|X_t,\theta^\dagger) = \Pr\left(X_{t+\tau}^{\theta^\dagger} \leq u|X_t^{\theta^\dagger}=X_t\right)$.

The statement of Proposition 2 requires the following assumption.

Assumption B: (i) $F_\tau(u|X_t,\theta)$ is twice continuously differentiable in the interior of Θ. Also, $\nabla_\theta F_\tau(u|X_t,\theta)$ and $\nabla_\theta^2 F_\tau(u|X_t,\theta)$ are jointly continuous in the interior of Θ, almost surely, and $2r$-dominated on Θ, $r>2$. (ii) $X_{s,t+\tau}^\theta$ is continuously differentiable in the interior of Θ, for $s=1,...,S$; and $\nabla_\theta X_{s,t+\tau}^\theta$ is $2r-$dominated in Θ, uniformly in s for $r>2$.

Proposition 2: Let Assumptions A and B hold. Assume that $T,N,S\to\infty$. Then, if $h\to 0$, $T/N\to 0$, and $h^2T\to 0$, $T^2/S\to\infty$, the following result holds for any X_t, $t\geq 1$, uniformly in u :

$$\widehat{F}_\tau(u|X_t,\widehat{\theta}_{T,N,h}) - F_\tau(u|X_t,\theta^\dagger) \overset{pr}{\to} 0, \tag{8}$$

In addition, if the model is correctly specified (i.e. if $\mu(\cdot,\cdot)=\mu_0(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)=\sigma_0(\cdot,\cdot)$) then:

$$\widehat{F}_\tau(u|X_t,\widehat{\theta}_{T,N,h}) - F_{0,\tau}(u|X_t,\theta_0) \overset{pr}{\to} 0, \tag{9}$$

where $F_{0,\tau}(u|X_t,\theta_0) = \Pr\left(X_{t+\tau}\leq u|X_t\right)$.

In practice, once we have an estimate for the τ period ahead conditional distribution, we still do not know whether our estimate is based on the correct model or not. This suggests that one use of the test for correct specification outlined in the next section is to assess the "relevance" of the distribution estimator.

4 Specification Testing

In the first sub-section, we outline the test. The second sub-section discusses bootstrapping procedures for obtaining asymptotically valid critical values.

4.1 The Test

The test outlined in this section can be viewed as a simulation-based extension of Andrews (1997) and Corradi and Swanson (2005b,2006b), which has been adapted to the use of continuous time models. Of additional note is that in the Andrews and Corradi-Swanson papers, the functional form of conditional distribution under the null hypothesis is assumed to be known, while in the current context we replace the conditional distribution (or conditional mean) with a simulation based estimator. Furthermore, in the Andrews paper, a conditional-Kolmogorov test is developed under the assumption of *iid* data, so that the bootstrap used in that paper is not applicable in our context.

Recall that in the previous section we discussed the construction of simulation paths for a given starting value. In order to carry out a specification test, however, we now require the construction of a sequence of $T-\tau$ conditional distributions that are $\tau-$steps ahead, say. In this way, S paths of length τ are thus available for each starting value from $t=1,...,T-\tau$.

The hypotheses of interest are:

$H_0: F_\tau(u|X_t,\theta^\dagger)=F_{0,\tau}(u|X_t,\theta_0)$, for all u, a.s.

$H_A: \Pr\left(F_\tau(u|X_t,\theta^\dagger)-F_{0,\tau}(u|X_t,\theta_0)\neq 0\right)>0$, for some $u\in U$, with non-zero Lebesgue measure.

Thus, the null hypothesis coincides with that of correct specification of the conditional distribution, and is implied by the correct specification of the drift and variance terms used in simulating the paths. The alternative is simply the negation of the null. In practice, we observe neither $F_\tau(u|X_t,\theta^\dagger)$ nor $F_{0,\tau}(u|X_t,\theta_0)$. However, we can define the following statistic:

$$V_T=\sup_{u\times v\in U\times V}|V_T(u,v)| \tag{10}$$

where

$$V_T(u,v)=\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}, \tag{11}$$

with U and V compact sets on the real line.

The statistic defined in (10) is a simulation-based version of the conditional Kolmogorov test of Andrews (1997). As stated in Proposition 2, $\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}$ is a consistent estimator of $F_{0,\tau}(X_t\leq u|X_t,\theta_0)$, which is the conditional distribution implied by the null model. Thus, we compare the joint empirical distribution $\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}1\{X_{t+\tau}\leq u\}1\left\{X_t\leq v\right\}$ with its semi-empirical/semi-parametric analog given by the product of $\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}F_{0,\tau}(u|X_t,\theta_0)1\left\{X_t\leq v\right\}$. Intuitively, if the null model used for simulating the data is correct, then the difference between the two approaches zero, and has a well-defined limiting distribution when properly scaled. The asymptotic behavior of V_T is described by the following theorem.

Theorem 3: Let Assumptions A and B hold. Assume that $T,N,S\rightarrow\infty$. Then, if $h\rightarrow 0$, $T/N\rightarrow 0$, $T/S\rightarrow 0$, $T^2/S\rightarrow\infty$, $Nh\rightarrow 0$, and $h^2T\rightarrow 0$, the following result holds under H_0 :

$$V_T\xrightarrow{d}\sup_{u\times v\in U\times V}|Z(u,v)|,$$

where $Z(u,v)$ is a Gaussian process with covariance kernel $K(u,u',v,v')$ given by:

$$\begin{aligned}K(u,u',v,v') &= \sum_{j=-\infty}^{\infty}E\left(\left(F_{0,\tau}(u|X_1,\theta_0)-1\{X_{1+\tau}\leq u\}\right)1\left\{X_1\leq v\right\}\right.\\ &\quad\left.\times\left(F_{0,\tau}(u'|X_{1+j},\theta_0)-1\{X_{1+\tau+j}\leq u'\}\right)1\left\{X_{1+j}\leq v'\right\}\right)\\ &\quad+\mu f_{0,\tau}(u,v)'(D^{0\prime}W_0D^0)^{-1}\mu f_{0,\tau}(u,v)\end{aligned}$$

$$-2\mu f_{0,\tau}(u,v)'(D^{0\prime}W_0D^0)^{-1}D^{0\prime}W^0$$
$$\sum_{j=-\infty}^{\infty}\left(g(X_{1+j})-E(g(X_1))\right)\left(F_{0,\tau}(u|X_{1+j},\theta_0)-1\{X_{1+\tau+j}\leq u'\}\right)1\left\{X_{1+j}\leq v'\right\}$$

where

$$\mu f_{0,\tau}(u,v) = E_X \left(f_{0,\tau}(u|X_1,\theta_0) E_s \left(\nabla_{\theta_0} X_{s,1+\tau}^{\theta_0} \right) 1\left\{X_1 \leq v\right\} \right),$$

E_X denotes the expectation under the law governing the sample and E_S denotes the expectation under the measure governing the simulated randomness.

Furthermore, under H_A, there exists some $\varepsilon > 0$ such that:

$$\lim_{P\to\infty} \Pr\left(\frac{1}{\sqrt{T}} V_T > \varepsilon\right) = 1.$$

Notice that the limiting distribution is a zero mean Gaussian process, with a covariance kernel given by $K(u,u',v,v')$. The first term is the long-run variance we would have if we knew $F_{0,\tau}(u|X_1,\theta_0)$; the second term captures the contribution of parameter estimation error; and the third term captures the correlation between the first two. As $T/S \to 0$, the contribution of simulation error is asymptotically negligible. The limiting distribution is not nuisance parameter free and hence critical values cannot be tabulated. In the next section we thus outline a bootstrap procedure for calculating asymptotically valid critical values for V_T.

4.2 Bootstrap Critical Values

Given that the limiting distribution of V_T is not nuisance parameter free, our approach is to construct bootstrap critical values for the test. In order to show the first order validity of the bootstrap, we shall obtain the limiting distribution of the bootstrapped statistic and show that it coincides with the limiting distribution of the actual statistic, under H_0. Then, a test with correct asymptotic size and unit asymptotic power can be obtained by comparing the value of the original statistic with bootstrapped critical values.

Asymptotically valid bootstrap critical values for the test should be constructed as follows:

Step 1*:* At each replication, draw b blocks (with replacement) of length l, where $bl = T$. Thus, each block is equal to $X_{i+1},...,X_{i+l}$, for some $i = 0,...,T-l$, with probability $1/(T-l+1)$. More formally, let I_k, $k = 1,...,b$ be *iid* discrete uniform random variables on $[0,1,...,T-l]$. Then, the resampled series, X_t^*, is such that $X_1^*, X_2^*, ..., X_l^*, X_{l+1}^*, ..., X_T^* = X_{I_1+1}, X_{I_1+2}, ..., X_{I_1+l}, X_{I_2}, ..., X_{I_b+l}$, and so a resampled series consists of b blocks that are discrete *iid* uniform random variables, conditional on the sample. Use these data to construct $\widehat{\theta}^*_{T,N,h}$. For $N/T \to \infty$, GMM and simulated GMM are asymptotically equivalent. Thus, we do not need resample simulated moment conditions. More precisely, define:

$$\widehat{\theta}^*_{T,N,h} = \arg\min_{\theta\in\Theta} \left(\frac{1}{T}\sum_{t=1}^{T} g(X_t^*) - \frac{1}{N}\sum_{t=1}^{N} g(X_{t,h}^{\theta}) \right)' W_T \left(\frac{1}{T}\sum_{t=1}^{T} g(X_t^*) - \frac{1}{N}\sum_{t=1}^{N} g(X_{t,h}^{\theta}) \right),$$

where W_T and $g(\cdot)$ are defined in (4).

Step 2*:* Using the same set of random errors used in the construction of the actual statistic, construct $X_{s,t+\tau,*}^{\widehat{\theta}^*_{T,N,h}}$, $s = 1,...,S$, and $t = 1,...,T-\tau$. Note that we do not resample the simulated series (as $S/T \to \infty$, simulation error is asymptotically negligible). Instead, simply simulate the series using bootstrap estimators and using bootstrapped values as starting values.

Step 3*:* Construct the following bootstrap statistic, which is the bootstrap counterpart of V_T:

$$V_T^* = \sup_{u\times v\in U\times V} |V_T^*(u,v)|, \qquad (12)$$

where

$$V_T^*(u,v) = \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau,*}^{\widehat{\theta}_{T,N,h}^*}\leq u\right\}-1\{X_{t+\tau}^*\leq u\}\right)1\{X_t^*\leq v\}$$
$$-\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\{X_t\leq v\} \quad (13)$$

Step 4*:* Repeat *Steps 1-3* B times, and generate the empirical distribution of the B bootstrap statistics.

Note that the first term of on the RHS of (13) is the bootstrap analog of (11). However, simulation error is asymptotically negligible (as $T/S \to 0$), so that there is no need to resample the simulated data. Though, in order to properly mimic the contribution of parameter estimation error, data are simulated using the bootstrap estimator. Also, we need to use resampled observed values as starting values for the simulated series. The second term on the RHS of (13) is the mean of the first, computed under the law governing the bootstrap, and conditional on the sample. It plays the role of a recentering term.

For the *iid* case when the conditional distribution, $F_0(y_t|X_t)$, is known in closed form, Andrews (1997) suggests using a parametric bootstrap procedure (i.e. resample y_t^*, say, from $F(u|X_t,\widehat{\theta}_T)$, keeping the conditioning variates, X_t, fixed). In our dynamic context, the conditioning variables cannot be kept constant. One possibility, for the case where $\tau=1$, would be to draw X_2^* from $\frac{1}{S}\sum_{s=1}^{S}1\left\{X_s^{\widehat{\theta}_{T,N,h}}(X_1)\leq u\right\}$, simulate again using X_2^* as the initial value and draw X_3^* from $\frac{1}{S}\sum_{s=1}^{S}1\left\{X_s^{\widehat{\theta}_{T,N,h}}(X_2^*)\leq u\right\}$, and so on; forming the resampled series $X_2^*,...,X_T^*$. Then, construct the analog of (10). However, this bootstrap procedure may not properly mimic the contribution of parameter estimation error. We leave this issue to future research.

Theorem 4: Let Assumptions A and B hold. Assume that $T,N,S\to\infty$. Then, if $h\to0$, $T/N\to0$, $T/S\to0$, $T^2/S\to\infty$, $Nh\to0$, $h^2T\to0$, $l\to\infty$, and $l^2/T\to0$, the following result holds:

$$P\left[\omega:\sup_{x\in\Re}\left|P^*\left(V_T^*(\omega)\leq x\right)-P\left((V_T-E(V_T))\leq x\right)\right|>\varepsilon\right]\to0,$$

where P^* denotes the probability law of the resampled series, conditional on the sample.

The above results suggest proceeding in the following manner. For any bootstrap replication, compute the bootstrap statistic, V_T^*. Perform B bootstrap replications (B large) and compute the quantiles of the empirical distribution of the B bootstrap statistics. Reject H_0 if V_T is greater than the $(1-\alpha)th$-percentile. Otherwise, do not reject. Now, for all samples except a set with probability measure approaching zero, V_T has the same limiting distribution as the corresponding bootstrapped statistic, ensuring asymptotic size equal to α. Under the alternative, V_T diverges to (plus) infinity, while the corresponding bootstrap statistic has a well defined limiting distribution, ensuring unit asymptotic power.

5 Stochastic Volatility Models

We now focus our attention on stochastic volatility models. Extension to general multi-factor and multi-dimensional models follows directly. Consider the following model:

$$\begin{pmatrix} dX(t) \\ dV(t) \end{pmatrix} = \begin{pmatrix} \mu_1(X(t),\theta^\dagger) \\ \mu_2(V(t),\theta^\dagger) \end{pmatrix}dt + \begin{pmatrix} \sigma_{11}(V(t),\theta^\dagger) \\ 0 \end{pmatrix}dW_1(t) + \begin{pmatrix} \sigma_{21}(V(t),\theta^\dagger) \\ \sigma_{22}(V(t),\theta^\dagger) \end{pmatrix}dW_2(t), \quad (14)$$

where $W_{1,t}$ and $W_{2,t}$ are independent standard Brownian motions, and where $\mu_1(\cdot,\cdot)$, $\mu_2(\cdot,\cdot)$, $\sigma_{11}(\cdot,\cdot)$, $\sigma_{21}(\cdot,\cdot)$, $\sigma_{22}(\cdot,\cdot)$, and $\theta^\dagger$ are replaced with $\mu_{1,0}(\cdot,\cdot)$, $\mu_{2,0}(\cdot,\cdot)$, $\sigma_{11,0}(\cdot,\cdot)$, $\sigma_{21,0}(\cdot,\cdot)$, $\sigma_{22,0}(\cdot,\cdot)$, and θ_0, respectively, under correct specification.

In the one-dimensional case, $X(t)$, can be expressed as a function of the driving Brownian motion, $W(t)$. However, in the multi-dimensional case, i.e. $X(t) \in R^p$, $X(t)$ cannot in general be expressed as a function of the p driving Brownian motions, but is instead a function of $(W_j(t), \int_0^t W_j(s)dW_i(s))$, $i,j = 1,...p$ (see e.g. Pardoux and Talay (1985), pp. 30-32). For this reason, simple approximation schemes like the Euler and Milstein schemes, which do not involve approximations of stochastic integrals, may not be adequate. One case in which the Milstein scheme does straigthforwardly generalize to the multidimensional case is when the diffusion matrix is commutative. Namely, write the diffusion matrix $\Sigma(X)$ as:

$$\Sigma(X) = \begin{pmatrix} \sigma_1(X) & . & . & . & \sigma_P(X) \end{pmatrix},$$

where $\sigma_i(X)$ is a $p \times 1$ vector, for $i = 1,...,p$. If for all $i,j = 1,...,p$,

$$\begin{pmatrix} \frac{\partial\sigma_j(X)}{\partial X_1} & . & . & . & \frac{\partial\sigma_j(X)}{\partial X_p} \end{pmatrix} \sigma_i(X) = \begin{pmatrix} \frac{\partial\sigma_i(X)}{\partial X_1} & . & . & . & \frac{\partial\sigma_i(X)}{\partial X_p} \end{pmatrix} \sigma_j(X),$$

then $\Sigma(X)$ is commutative.

However, almost all frequently used stochastic volatility (SV) models violate the commutativity property (see e.g. Heston (1993), the GARCH diffusion model of Nelson (1990), and general the eigenfunction stochastic volatility models of Meddahi (2001)). A simple way of imposing commutativity is to assume no leverage. Now, the non-leverage assumption is suitable for exchange rates, but not for stock returns. In this situation, more "sophisticated" approximation schemes are necessary.

It is immediate to see that the diffusion in (14) violates the commutativity property. Now, let

$$\mu(\cdot,\cdot) = \begin{pmatrix} \mu_1(\cdot,\cdot) \\ \mu_2(\cdot,\cdot) \end{pmatrix}, \; \sigma = \begin{pmatrix} \sigma_{11}(\cdot,\cdot) & 0 \\ \sigma_{21}(\cdot,\cdot) & \sigma_{22}(\cdot,\cdot) \end{pmatrix}, \tag{15}$$

and define the following generalized Milstein scheme (see eq. (3.3), p. 346 in Kloeden and Platen (1999)),

$$\begin{aligned} X^\theta_{(k+1)h} &= X^\theta_{kh} + \widetilde{\mu}_1(X^\theta_{kh},\theta)h + \sigma_{11}(V^\theta_{kh},\theta)\epsilon_{1,(k+1)h} + \sigma_{12}(V^\theta_{kh},\theta)\epsilon_{2,(k+1)h} \\ &\quad + \frac{1}{2}\sigma_{22}(V^\theta_{kh},\theta)\frac{\partial\sigma_{12}(V^\theta_{kh},\theta)}{\partial V}\epsilon^2_{2,(k+1)h} \\ &\quad + \sigma_{22}(V^\theta_{kh},\theta)\frac{\partial\sigma_{11}(V^\theta_{kh},\theta)}{\partial V}\int_{kh}^{(k+1)h}\left(\int_{kh}^{s} dW_{2,\tau}\right)dW_{1,s} \end{aligned} \tag{16}$$

$$\begin{aligned} V^\theta_{(k+1)h} &= V^\theta_{kh} + \widetilde{\mu}_2(V^\theta_{kh},\theta)h + \sigma_{22}(V^\theta_{kh},\theta)\epsilon_{2,(k+1)h} \\ &\quad + \frac{1}{2}\sigma_{22}(V^\theta_{kh},\theta)\frac{\partial\sigma_{22}(V^\theta_{kh},\theta)}{\partial V}\epsilon^2_{2,(k+1)h} \end{aligned} \tag{17}$$

where $h^{-1/2}\epsilon_{i,kh} \sim N(0,1)$, $i = 1,2$, $E(\epsilon_{1,kh}\epsilon_{2,mh}) = 0$ for all k and m, and

$$\widetilde{\mu}(V,\theta) = \begin{pmatrix} \widetilde{\mu}_1(V,\theta) \\ \widetilde{\mu}_2(V,\theta) \end{pmatrix} = \begin{pmatrix} \mu_1(V,\theta) - \frac{1}{2}\sigma_{22}(V,\theta)\frac{\partial\sigma_{12}(V,\theta)}{\partial V} \\ \mu_2(V,\theta) - \frac{1}{2}\sigma_{22}(V,\theta)\frac{\partial\sigma_{22}(V,\theta)}{\partial V} \end{pmatrix}.$$

The last terms on the RHS of (16) involve stochastic integrals and cannot be explicitly computed. However, they can be approximated, up to an error of order $o(h)$ by (see eq. (3.7), p. 347 in Kloeden and Platen (1999)):

$$\int_{kh}^{(k+1)h}\left(\int_{kh}^{s}dW_{2,\tau}\right)dW_{1,s}\approx h\left(\frac{1}{2}\xi_1\xi_2+\sqrt{\rho_p}\left(v_{2,p}\xi_1-v_{1,p}\xi_2\right)\right)$$
$$+\frac{h}{2\pi}\sum_{r=1}^{p}\frac{1}{r}\left(\varsigma_{2,r}\left(\sqrt{2}\xi_1+\eta_{1,r}\right)-\varsigma_{1,r}\left(\sqrt{2}\xi_2+\eta_{2,r}\right)\right),$$

where for $j=1,2$, $\xi_j, v_{j,p}, \varsigma_{j,r}, \eta_{j,r}$ are $iidN(0,1)$ with $\xi_j=h^{-1/2}\epsilon_{j,(k+1)h}$, $\rho_p=\frac{1}{12}-\frac{1}{2\pi^2}\sum_{r=1}^{p}\frac{1}{r^2}$, and p is such that as $h\to 0$, $p\to\infty$.

In order to construct conditional distributions, given the observable state variables, we need to perform the following steps.

Step 1: Simulate a path of length N using the schemes in (16),(17) and estimate θ by SGMM, as in (4). Also, retrieve $V_{kh}^{\widehat{\theta}_{T,N,h}}$, for $k=1/h,...,K/h$, with $Kh=N$, and hence obtain $V_{j,h}^{\widehat{\theta}_{T,N,h}}$, $j=1,...,N$ (i.e. sample the simulated volatility at the same frequency as the data).

Step 2: Using the schemes in (16),(17), simulate $S\times N$ paths of length τ, setting the initial value for the observable state variable to be X_t. As we do not observe data on volatility, use the values simulated in the previous step as the initial value for the volatility process (i.e. as initial values for the unobservable state variable, use $V_{j,h}^{\widehat{\theta}_{T,N,h}}$, $j=1,...,N$). Also, keep the simulated randomness (i.e. $\epsilon_{1,kh},\epsilon_{2,kh}$, $\int_{kh}^{(k+1)h}\left(\int_{kh}^{s}dW_{1,\tau}\right)dW_{2,s}$) constant across j and t (i.e. constant across the different starting values for the unobservable and observable state variables). Define $X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}}$ to be the simulated $\tau-$step ahead value for the return series at replication s, and using initial values X_t and $V_{j,h}^{\widehat{\theta}_{T,N,h}}$.

Step 3: As an estimator of $F_\tau(u|X_t,\theta^\dagger)$, construct $\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S}1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}$. Note that, by averaging over the initial value of the volatility process, we have integrated out its effect. In other words, $\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}$ is an estimate of $F_\tau(u|X_t,V_{j,h}^{\theta^\dagger},\theta^\dagger)$.

Step 4: Construct the statistic of interest:

$$SV_T=\sup_{u\times v\in U\times V}\left|SV_T(u,v)\right|,$$

where

$$SV_T(u,v)=\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S}1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}, \qquad (18)$$

Assumption A′: This assumption is the same as Assumption A, except that Assumption A(ii) is replaced by (ii′): $\mu(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$ (as defined in (14) and (15)), and $\sigma_{lj}(V,\theta)\frac{\partial\sigma_{k\iota}(V,\theta)}{\partial V}$ are twice continuously differentiable, Lipschitz, with Lipschitz constant independent of θ, and grow at most at a linear rate, uniformly in Θ, for $l,j,k,\iota=1,2$.

All of the results outlined in Sections 3 and 4 generalize to the current setting. In particular, the following results hold.

Proposition 5: Let Assumptions A$'$ and B hold. Assume that $T, N, S \to \infty$. Then, if $h \to 0$, $T/N \to 0$, $T^2/S \to \infty$, and $h^2 T \to 0$, the following result holds for any X_t, $t \geq 1$:

$$\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} - F_\tau(u|X_t, \theta^\dagger) \xrightarrow{pr} 0, \text{ uniformly in } u.$$

In addition, if the model is correctly specified, i.e. if $\mu(\cdot,\cdot) = \mu_0(\cdot,\cdot)$ and $\sigma(\cdot,\cdot) = \sigma_0(\cdot,\cdot)$, then

$$\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} - F_{0,\tau}(u|X_t, \theta_0) \xrightarrow{pr} 0, \text{ uniformly in } u.$$

In the sequel, consider testing the hypotheses given in Section 4.1. However, in the present context correct specification of the conditional distribution of $X_{t+\tau}|X_t$ no longer implies correct specification of the underlying diffusion process. Indeed, correct specification of the stochastic volatility process is equivalent to correct specification of $X_{t+\tau}, V_{t+\tau}|X_t, V_t$. Also, while X_t and V_t are jointly Markovian, X_t is no longer Markov. Therefore, by using only X_t as a conditioning variable, we implicitly allow for dynamic mispecification under the null. Furthermore, it thus follows that the test has no power against diffusion processes having the same transition density as $X_{t+\tau}|X_t$.

Theorem 6: Let Assumptions A$'$ and B hold. Assume that $T, N, S \to \infty$. Then, if $h \to 0$, $T/N \to 0$, $T/S \to 0$, $T^2/S \to \infty$, $Nh \to 0$, and $h^2 T \to 0$, the following result holds under H_0 :

$$SV_T \xrightarrow{d} \sup_{u \times v \in U \times V} |SZ(u,v)|,$$

where $SZ(v)$ is a Gaussian process with covariance kernel $SK(v, v')$ given by:

$$\begin{aligned} SK(u,u',v,v') &= \sum_{k=-\infty}^{\infty} E\left((F_{0,\tau}(u|X_1,\theta_0) - 1\{X_{1+\tau} \leq u\})\, 1\{X_1 \leq v\}\right. \\ &\qquad \left.(F_{0,\tau}(u'|X_{1+k},\theta_0) - 1\{X_{1+\tau+k} \leq u'\})\, 1\{X_{1+k} \leq v'\}\right) \\ &\quad + \mu_{SV} f_{0,\tau}(u,v)'(D^{0\prime}W_0 D^0)^{-1}\mu_{SV} f_{0,\tau}(u,v) \end{aligned}$$

$$-2\mu_{SV} f_{0,\tau}(u,v)'(D^{0\prime}W_0D^0)^{-1}D^{0\prime}W^0$$
$$\sum_{j=-\infty}^{\infty}(g(X_{1+j}) - E(g(X_1)))\,(F_{0,\tau}(u|X_{1+j},\theta_0) - 1\{X_{1+\tau+j} \leq u'\})\, 1\{X_{1+j} \leq v'\}$$

where

$$\mu_{SV} f_{0,\tau}(u,v) = E\left(f_{0,\tau}(u|X_1,\theta_0)E_{j,s}\left(\nabla_{\theta_0} X_{j,s,1+\tau}^{\theta_0}\right) 1\{X_1 \leq v\}\right)$$

and $E_{j,s}$ denotes the expectation with respect to the joint probability measure governing the simulated randomness and the volatility process, $V_{h,j}^{\theta_0}$, conditional on the sample.

Furthermore, under H_A, there exists some $\varepsilon > 0$ such that:

$$\lim_{P\to\infty} \Pr\left(\frac{1}{\sqrt{T}} SV_T > \varepsilon\right) = 1.$$

Note that in a Bayesian context, Chib, Kim and Shephard (1998), Elerian, Chib and Shephard (2001) suggest likelihood ratio tests for the comparison of a stochastic volatility model against a specific alternative model, for the discete and continuous cases, respectively. In their tests, the likelihood is constructed via the joint use of probability integral transform and particle filtering (see also Pitt and Shephard (1999) and Pitt (2005)). These tests differ from ours, as we are interested in assessing whether the conditional density of $X_{t+\tau}|X_t$ implied by the null model is correct or not.

In order to obtain asymptotically valid critical values, resample as discussed above for the one-factor model, and note that there is no need to resample $V^{\theta}_{h,j}$. In particular, form bootstrap statistics as follows:

$$SV_T^* = \sup_{u \times v \in U \times V} \left| SV_T^*(u,v) \right|, \tag{19}$$

where

$$\begin{aligned} SV_T^*(u,v) &= \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau,*}^{\widehat{\theta}^*_{i,T,N,h}} \leq u \right\} - 1\{X^*_{t+\tau} \leq u\} \right) 1\left\{X_t^* \leq v\right\} \\ &\quad - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau}^{\widehat{\theta}_{i,T,N,h}} \leq u \right\} - 1\{X_{t+\tau} \leq u\} \right) 1\left\{X_t \leq v\right\}, \end{aligned}$$

and where $X_{j,s,t+\tau,*}^{\widehat{\theta}^*_{i,T,N,h}}$ is the simulated value at simulation s, constructed using $\widehat{\theta}^*_{i,T,N,h}$ and using X_t^* and $V_{j,h}^{\widehat{\theta}^*_{i,T,N,h}}$ as initial values. This allows us to state the following theorem.

Theorem 7: Let Assumptions A$'$ and B hold. Assume that $T, N, S \to \infty$. Then, if $h \to 0$, $T/N \to 0$, $T/S \to 0$, $T^2/S \to \infty$, $Nh \to 0$, $h^2 T \to 0$, $l \to \infty$, and $l^2/T \to 0$, the following result holds under H_0 :

$$P\left[\omega : \sup_{x \in \Re} \left| P^*\left(SV_T^*(\omega) \leq x\right) - P\left(\left(SV_T - E(SV_T)\right) \leq x\right)\right| > \varepsilon\right] \to 0.$$

As in the single-factor case, the above results suggest proceeding in the following manner. For any bootstrap replication, compute the bootstrap statistic, SV_T^*. Perform B bootstrap replications (B large) and compute the quantiles of the empirical distribution of the B bootstrap statistics. Reject H_0, if SV_T is greater than the $(1-\alpha)th$-percentile. Otherwise, do not reject. Now, under the null, for all samples except a set with probability measure approaching zero, SV_T and SV_T^* have the same limiting distribution, ensuring asymptotic size equal to α. Under the alternative, SV_T diverges to (plus) infinity, while the corresponding bootstrap statistic has a well defined limiting distribution, ensuring unit asymptotic power.

6 Experimental and Empirical Results

In this section, we briefly illustrate the above testing methodology for the case in which we are interested in specification testing from the perspective of conditional confidence intervals. In particular, consider $V_T = \sup_{v \in V} |V_T(v)|$,where

$$V_T(v) = \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{S} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq \overline{u} \right\} - 1\{\underline{u} \leq X_{t+\tau} \leq \overline{u}\} \right) 1\left\{X_t \leq v\right\}.$$

Additionally, define the following bootstrap statistic: $V_T^* = \sup_{v \in V} |V_T^*(v)|$, where

$$\begin{aligned} V_T^*(v) &= \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{S} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{s,t+\tau,*}^{\widehat{\theta}^*_{T,N,h}} \leq \overline{u} \right\} - 1\{\underline{u} \leq X_{t+\tau}^* \leq \overline{u}\} \right) 1\left\{X_t^* \leq v\right\} \\ &\quad - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{S} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq \overline{u} \right\} - 1\{\underline{u} \leq X_{t+\tau} \leq \overline{u}\} \right) 1\left\{X_t \leq v\right\}. \end{aligned}$$

It is immediate to see that the above statistic is a version of the distributional test discussed above, so that all of the theoretical results outlined in Section 4 and 5 hold. For the case of stochastic volatility models, consider the statistic $SV_T = \sup_{v \in V} |SV_p(v)|$, where

$$SV_T(v) = \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq \overline{u} \right\} - 1\{\underline{u} \leq X_{t+\tau} \leq \overline{u}\} \right) 1\left\{X_t \leq v\right\};$$

and its bootstrap analog $SV_T^* = \sup_{v \in V} |SV_T^*(v)|$, where

$$\begin{aligned} SV_T^*(v) &= \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{j,s,t+\tau,*}^{\widehat{\theta}^*_{i,T,N,h}} \leq \overline{u} \right\} - 1\{\underline{u} \leq X_{t+\tau}^* \leq \overline{u}\} \right) 1\left\{X_t^* \leq v\right\} \\ &\quad - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{j,s,t+\tau}^{\widehat{\theta}_{i,T,N,h}} \leq \overline{u} \right\} - 1\{\underline{u} \leq X_{t+\tau} \leq \overline{u}\} \right) 1\left\{X_t \leq v\right\}. \end{aligned}$$

In the sequel we shall study a common version of the square root process discussed in Cox, Ingersoll, and Ross (CIR: 1985), the square root stochastic volatility model of Heston (1993), called SV, and a stochastic volatility model with jumps , called SVJ. The specifications of these models are as follows:

CIR: $dr(t) = k_r(\overline{r} - r(t))\,dt + \sigma_r\sqrt{r(t)}dW_r(t)$, where $k_r > 0$, $\sigma_r > 0$ and $2k_r\overline{r} \geq \sigma_r^2$,

SV: $dr(t) = k_r(\overline{r} - r(t))\,dt + \sqrt{V(t)}dW_r(t)$, and $dV(t) = k_v(\overline{v} - V(t))\,dt + \sigma_v\sqrt{V(t)}dW_v(t)$, where $W_r(t)$ and $W_v(t)$ are independent Brownian motions, and where $k_r > 0$, $\sigma_r > 0$, $k_v > 0$, $\sigma_v > 0$, and $2k_v\overline{v} > \sigma_v^2$.

SVJ: $dr(t) = k_r(\overline{r} - r(t))\,dt + \sqrt{V(t)}dW_r(t) + J_u dq_u - J_d dq_d$, and $dV(t) = k_v(\overline{v} - V(t))\,dt + \sigma_v\sqrt{V(t)}dW_v(t)$, where $W_r(t)$ and $W_v(t)$ are independent Brownian motions, and where $k_r > 0$, $\sigma_r > 0$, $k_v > 0$, $\sigma_v > 0$, and $2k_v\overline{v} > \sigma_v^2$. Further q_u and q_d are Poisson processes with jump intensity λ_u and λ_d, and are independent of the Brownian motions $W_r(t)$ and $W_v(t)$. Jump sizes are *iid* and are controlled by jump magnitudes $\zeta_u, \zeta_d > 0$, which are drawn from exponential distributions, with densities: $f(J_u) = \frac{1}{\zeta_u}\exp\left(-\frac{J_u}{\zeta_u}\right)$ and $f(J_d) = \frac{1}{\zeta_d}\exp\left(-\frac{J_d}{\zeta_d}\right)$. Here, λ_u is the probability of a jump up, $\Pr(dq_u(t) = 1) = \lambda_u$, and jump up size is controlled by J_u; while λ_d and J_d control jump down intensity and size. Note that the case of Poisson jumps with constant intensity and jump size with exponential density is covered by the assumptions stated in the previous sections.

6.1 Monte Carlo Experiment

In this subsection, parameterizations for the above models were selected via examination of models estimated using the interest rate data discussed in the next subsection, and include: *CIR* - $(k_r, \overline{r},\ \sigma_r) =$

$\{(0.15, 0.05, 0.10), (0.30, 0.05, 0.10), (0.50, 0.05, 0.10)\}$; SV - $(k_r, \overline{r}, k_v, \overline{v}, \sigma_v) = (0.30, 0.05, 5.00, 0.0005, 0.01)$; SVJ - $(k_r, \overline{r}, k_v, \overline{v}, \sigma_v, \lambda_u, \zeta_u, \lambda_d, \zeta_d) = (0.30, 0.05, 2.5, 0.0005, 0.01, 8.00, 0.001, 1.50, 0.001)$.Note that k_r in the data was estimated to be approximately 0.30, and our DGPs set $k_r = 0.15, 0.30, 0.50$. Thus, we include a case where there is substantially less mean reversion than that observed in the data. For a discussion of the crucial issue of persistence and its effect on tests related to that discussed in this paper, the reader is referred to Pritsker (1998). Note that the parameters estimated for the three models, ordered as CIR,SV, and SVJ are: $k_r = 0.31, 0.24, 0.50$; $\overline{r} = 0.068, 0.059, 0.068$; $\sigma_r = 0.11$; $k_v = 5.12, 1.56$; $\overline{v} = 0.0003, 0.0009$; $\sigma_v = 0.014, 0.029$; $\lambda_u = 7.86$; $\zeta_u = 0.0004$; $\lambda_d = 2.08$; and $\zeta_d = 0.0008$. Data were generated using the Milstein scheme discussed above with h=$1/T$, for $T = \{400, 800\}$. The jump component can be simulated without any error, because of the constancy of the intensity parameter.

The three models we study (i.e. the CIR, SV and SVJ models) fall in the class of affine diffusions. Therefore, it is possible to compute in closed form the conditional characteristic function; for the CIR model we follow Singleton (2001), for the SV model Jiang and Knigth (2002), and for the SVJ model Chacko and Viceira (2003). Given the conditional characteristic function, we compute as many conditional moments as the number of parameters to be estimated and thus we implement exactly identified GMM. (Note that the criticism of Andersen and Sorensen (1996) about the use of exact GMM does not necessarily apply to the case of conditional moments obtained via the conditional characteristic function.)

In our experiment, all empirical bootstrap distributions were constructed using 100 bootstrap replications, and critical values were set equal to the 90^{th} percentile of the bootstrap distributions. For the bootstrap, block lengths of 5, 10, 20 and 50 were tried. Additionally, we set $S = \{10T, 20T\}$, and for model SV and SVJ we set $N = S$. Tests were carried out using $\tau - step$ ahead confidence intervals, for $\tau = \{1, 2, 4, 12\}$, and we set $(\underline{u}, \overline{u})$; $\overline{X} \pm 0.5\sigma_X$, and $\overline{X} \pm \sigma_X$, where $\overline{X}$ and σ_X are the mean and variance of an initial sample of data. Finally, $[\underline{v}, \overline{v}]$ was set equal to $[X_{\min}, X_{\max}]$, and a grid of 100 equally spaced values for v across this range was used, where $X_{\min}$ and $X_{\max}$ are again fixed using the initial sample. All results are based on 500 Monte Carlo iterations.

Results are gathered in Tables 1-3. Tables 1-2 report results for the size of the test while Tables 3 reports results based on power experiments. In the size experiments, estimated models are the same as the models used for data generation, while in the power experiments, we estimate the CIR model when data are generated according to either the SV or SVJ model (Bhardwaj, Corradi and Swanson (2005) also consider Ohrstein-Uhlenbeck (OU) models); and we estimate an OU model when data are generated according to a CIR model. Finally, for the power experiment where data are generated according to a CIR model, we report results for the parameterization where $(k_r, \overline{r}, \sigma_r) = (0.30, 0.05, 0.10)$ (additional results are available upon request).

Each of the panels in the tables contains results for T =400 and T =800. Table 1 reports the empirical size for the three CIR parameterizations, while Table 2 reports the empirical size for the SV and SVJ models. Turning first to Tables 1-2, note that for $T = 400$, empirical size ranges from 13% to 30% (the nominal size is 10%). Additionally, the test seems to perform slightly worse for the SVJ model, while there are no relevant differences across the three CIR parameterizations and the SV model. When the sample is increased to $T = 800$ observations, empirical size is rather closed to nominal size, ranging from 10% to the 15% (again, there are no obvious differences across different models and parameterizations). Turning now to Table 3, note that rejection rates are rather similar for both cases where data are generated using the SV and the

SVJ model. Namely, for $T = 400$ rejection rates range from 35% to 60%, while for $T = 800$ rejection rates range from 55% to 75%. Interestingly, these findings seem to be quite robust to the choice of bootstrap block length, as well as to the choice of S and τ.

6.2 Empirical Illustration

In this subsection, the CIR, SV and SVJ models discussed above were fit to the one-month Eurodollar deposit rate for the period January 6, 1971 - September 30, 2005 (1,813 weekly observations). Other interest rate datasets examined in the literature include the monthly federal funds rate (Ait-Sahalia (1999)), the weekly 3-month T-bill rate (Andersen, Benzoni and Lund (2004)), and the weekly US dollar swap rate (Dai and Singleton (2000)), to name but a few.

We use the specification tests outlined in Section 6.1. Intervals examined, block lengths considered, simulation samples sizes used, bootstrap replications, and values of τ considered are the same as discussed the previous subsection. For example, we again consider $\tau = \{1, 2, 4, 12\}$, corresponding to one week, two week, one month, and one quarter ahead intervals. The exception is that the interval endpoints are chosen using actual interest rate data (i.e. $\overline{X}$, σ_X, $X_{\min}$, and $X_{\max}$ are constructed using the historical data). Note, for example, that using this approach, $\overline{X} \pm 0.5\sigma_X$ and $\overline{X} \pm \sigma_X$ correspond to intervals with 46.3% and 72.4% coverage, respectively. In the tables, test statistic values (denoted by V_T for CIR and SV_T for SV and SVJ) and 5%, 10%, and 20% nominal size bootstrap critical values are given. Single, double, and triple starred entries denote rejection using $20\%, 10\%$, and 5% size tests, respectively. Not surprisingly, the CIR models is rejected using 5% size tests in almost all cases. The same finding has been found for example by Aït-Sahalia (1996) and Bandi (2002). On the other hand, for the SV and SVJ models the results are more mixed. Rejections tends to occur only for the smaller confidence interval. Additionally, the SVJ model appears to be rejected slightly more frequently.

7 Concluding Remarks

In this paper we outline a simple simulation based framework for constructing conditional distributions for multi-factor and multi-dimensional diffusion processes, for the case where the functional form of the conditional density is unknown. In a Monte Carlo experiment and an empirical illustration, we show how the estimated distributions can be used, for example, to form conditional confidence intervals for time period $t+\tau$, say, given information up to period t. In addition, we use the simulation based framework to construct a test for the correct specification of a diffusion process, and establish the asymptotic validity of the block bootstrap for use in the construction of critical values.

This work represents a starting point in our investigation of the usefulness of simulation based methods for examining continuous time financial models. From a theoretical perspective, it is of interest to establish whether or not the simulation methodology discussed herein can be extended to contexts in which recursively constructed predictions are evaluated, and are used in model selection tests. From an empirical perspective, it remains to compare the finite sample performance of the specification test proposed here with alternative tests available in the literature. For example, it should be of interest to compare the recent Aït-Sahalia et al. (2005) and Altissimo et al. (2005) tests with that proposed in this paper, in the context of finite sample test performance.

8 Appendix

Proof of Lemma 1: Immediate from the proof of Lemma A1 in Corradi and Swanson (2005a).

Proof of Proposition 2: Given Assumption A, by Lemma 1, $\sqrt{T}\left(\widehat{\theta}_{T,N,h}-\theta^{\dagger}\right)=O_P(1)$. Given B(ii) and noting that the indicator function is a measurable function, by the same argument as in Proposition 1 in Corradi and Swanson (2007),

$$\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}=\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta^{\dagger}}\leq u\right\}+o_P(1).$$

Finally, for any given X_t, $X_{s,t+\tau}^{\theta^{\dagger}}$ is identically distributed and independent of $X_{j,t+\tau}^{\theta^{\dagger}}$, for all $j\neq s$. The statement in (8) then follows from the uniform law of large number for *iid* random variables. In fact, conditional on X_t, $X_{s,t+\tau}^{\theta^{\dagger}}$ is *iid*, as randomness is independent across simulations. Finally, the statement in (9) follows immediately, as in this case $F_\tau(u|X_t,\theta^{\dagger})=F_{0,\tau}(u|X_t,\theta_0)$.

Proof of Theorem 3: We begin by showing convergence in distribution under the null, pointwise in u and v. Recalling that $\theta^{\dagger}=\theta_0$ under H_0,

$$\begin{aligned}V_T(u,v) &= \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta_0}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}\\ &\quad+\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta_0}\leq u\right\}\right)1\left\{X_t\leq v\right\}\end{aligned}\tag{20}$$

Recalling that $F_\tau(u|X_t,\theta^{\dagger})=F_{0,\tau}(u|X_t,\theta_0)$, under H_0, the first term on the RHS can be written as:

$$\begin{aligned}&\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta_0}\leq u\right\}-F_{0,\tau}(u|X_t,\theta_0)\right)1\left\{X_t\leq v\right\}\\ &-\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(1\left\{X_{t+\tau}\leq u\right\}-F_{0,\tau}(u|X_t,\theta_0)\right)1\left\{X_t\leq v\right\}\end{aligned}\tag{21}$$

We now show that the first term on the RHS of (21) is $o_P(1)$, uniformly in u. Given that $1\left\{X_t\leq v\right\}$ is either 0 or 1, it suffices to show that $\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta_0}\leq u\right\}-F_{0,\tau}(u|X_t,\theta_0)\right)=o_P(1)$, uniformly in u. By Chebyshev's inequality, it thus suffices to show that

$$Var\left(\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta_0}\leq u\right\}-F_{0,\tau}(u|X_t,\theta_0)\right)\right)\to 0,\ \text{as } T,S\to\infty.\tag{22}$$

First note that,

$$\begin{aligned}&E\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta_0}\leq u\right\}-F_{0,\tau}(u|X_t,\theta_0)\right)\\ =\ &\frac{1}{S}\sum_{s=1}^{S}E_X\left(E_S\left(1\left\{X_{s,t+\tau}^{\theta_0}\leq u\right\}-F_{0,\tau}(u|X_t,\theta_0)\right)|X_t\right)=0,\end{aligned}$$

where E_X denotes expectation with respect to the probability law governing the sample, and E_S denotes expectation with respect to the probability law governing the simulated randomness, conditional on the sample. Hereafter, let $\phi_s(X_t) = 1\left\{X^{\theta_0}_{s,t+\tau} \leq u\right\} - F_{0,\tau}(u|X_t, \theta_0)$, and note that,

$$E\left(\phi_s(X_t)\phi_j(X_l)\right) = 0, \text{ for all } s \neq j,$$

given that:

$$\begin{aligned} & E\left(1\left\{X^{\theta_0}_{s,t+\tau} \leq u\right\} 1\left\{X^{\theta_0}_{j,l+\tau} \leq u\right\}\right) = E_X\left(E_S\left(1\left\{X^{\theta_0}_{s,t+\tau} \leq u\right\} 1\left\{X^{\theta_0}_{j,l+\tau} \leq u\right\}\right)\right) \\ = \; & E_X\left(F_{0,\tau}(u|X_t, \theta_0)F_{0,\tau}(u|X_l, \theta_0)\right) \end{aligned}$$

Thus, (22) can be written as:

$$\frac{1}{TS^2}\sum_{t=1}^{T-\tau}\sum_{l=1}^{T-\tau}\sum_{s=1}^{S} E\left(\phi_s(X_t)\phi_s(X_l)\right) = O\left(\frac{T}{S}\right) = o(1),$$

for $T/S \to 0$. This establishes that the first term on the RHS of (21) is $o_P(1)$, pointwise in u. Uniformity in u follows from the stochastic equicontinuity of $1\left\{X^{\theta_0}_{s,t+\tau} \leq u\right\} - F_{0,\tau}(u|X_t, \theta_0)$.

Now, the second term on the RHS of (20), can be written as:

$$\begin{aligned} & \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S} 1\left\{X^{\theta_0}_{s,t+\tau} \leq u - \left(X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau} - X^{\theta_0}_{s,t+\tau}\right)\right\} - F_{0,\tau}\left(u - \left(X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau} - X^{\theta_0}_{s,t+\tau}\right)|X_t, \theta_0\right)\right) \\ & \times 1\left\{X_t \leq v\right\} \\ & + \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\frac{1}{S}\sum_{s=1}^{S}\left(F_{0,\tau}\left(u - \left(X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau} - X^{\theta_0}_{s,t+\tau}\right)|X_t, \theta_0\right) - F_{0,\tau}\left(u|X_t, \theta_0\right)\right) 1\left\{X_t \leq v\right\}. \end{aligned} \tag{23}$$

The first term of (23) is $o_P(1)$, uniformly in u and v, by an argument analogous to that used in the proof of Theorem 3 in Corradi and Swanson (2005a). Given Assumption B(i)-(ii), it thus follows that:

$$\begin{aligned} & V_T(u, v) \\ = \; & -\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(1\left\{X_{t+\tau} \leq u\right\} - F_{0,\tau}(u|X_t, \theta_0)\right) 1\left\{X_t \leq v\right\} \\ & + \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\frac{1}{S}\sum_{s=1}^{S}\left(F_{0,\tau}\left(u - \left(X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau} - X^{\theta_0}_{s,t+\tau}\right)|X_t, \theta_0\right) - F_{0,\tau}\left(u|X_t, \theta_0\right)\right) 1\left\{X_t \leq v\right\} + o_P(1) \\ = \; & -\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(1\left\{X_{t+\tau} \leq u\right\} - F_{0,\tau}(u|X_t, \theta_0)\right) 1\left\{X_t \leq v\right\} \\ & + \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\left(\left(\frac{1}{S}\sum_{s=1}^{S}\left(f_{0,\tau}\left(u - \left(X^{\overline{\theta}_{T,N,h}}_{s,t+\tau} - X^{\theta_0}_{s,t+\tau}\right)|X_t\right)\nabla_{\theta_i} X^{\theta}_{s,t+\tau}|_{\theta=\overline{\theta}_{T,N,h}}\right)'\right. \\ & \left. 1\left\{X_t \leq v\right\}\sqrt{T-\tau}\left(\widehat{\theta}_{T,N,h} - \theta_0\right)\right) + o_P(1), \end{aligned} \tag{24}$$

where the $o_P(1)$ term holds uniformly in u and v. Finally, recalling Lemma 1 and Assumption A1(v), the second term on the RHS of the last equality in (24) can be appropriately restated, so that:

$$\begin{aligned} V_T(u, v) \; = \; & -\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(1\left\{X_{t+\tau} \leq u\right\} - F_{0,\tau}(u|X_t, \theta_0)\right) 1\left\{X_t \leq v\right\} \\ & + \frac{1}{T-\tau}\sum_{t=1}^{T-\tau} f_{0,\tau}\left(u|X_t\right) E_S\left(\nabla_{\theta_i} X^{\theta_0}_{s,1+\tau+j}\right)' 1\left\{X_t \leq v\right\}\sqrt{T-\tau}\left(\widehat{\theta}_{T,N,h} - \theta_0\right) \\ & + o_P(1), \text{ uniformly in } u \text{ and } v, \end{aligned}$$

where $E_S\left(\nabla_{\theta_i} X^{\theta_0}_{s,1+\tau+j}\right)$ is a measurable function of X_t.
The covariance expression in the statement then follows by noting that:

$$\sqrt{T}\left(\widehat{\theta}_{T,N,h}-\theta_0\right)=\left(-\nabla_\theta G_{T,N,h}(\widehat{\theta}_{T,N,h})' W_T \nabla_\theta G_{T,N,h}(\overline{\theta}_{T,N,h})\right)^{-1}\nabla_\theta G_{T,N,h}(\widehat{\theta}_{T,N,h})' W_T \sqrt{T} G_{T,N,h}(\theta_0)$$

Proof of Theorem 4: Consider the bootstrap statistic:

$$V_T^* = \sup_{u\times v\in U\times V}\left|V_T^*(u,v)\right|,$$

where

$$\begin{aligned} V_T^*(u,v) &= \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau,*}\leq u\right\}-1\{X^*_{t+\tau}\leq u\}\right)1\left\{X^*_t\leq v\right\} \\ &\quad -\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\} \\ &\quad +\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\widehat{\theta}^*_{T,N,h}}_{s,t+\tau,*}\leq u\right\}-\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau,*}\leq u\right\}\right)1\left\{X^*_t\leq v\right\} \end{aligned} \tag{25}$$

The term in the first two lines in (25) has the same limiting distribution as:

$$\begin{aligned} &\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\theta^\dagger}_{s,t+\tau}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)\right. \\ &\left.-E\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\theta^\dagger}_{s,t+\tau}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)\right)1\left\{X_t\leq v\right\}, \end{aligned}$$

where $\theta^\dagger=\theta_0$ if the null is true, conditional on the sample and for all samples except a subset of probability measure approaching zero. In fact,

$$\begin{aligned} &\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}E^*\left(\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau,*}\leq u\right\}-1\{X^*_{t+\tau}\leq u\}\right)1\left\{X^*_t\leq v\right\}\right. \\ &\left.-\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X^{\widehat{\theta}_{T,N,h}}_{s,t+\tau}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}\right) \\ = &\ O\left(l/T\right),\ \ \Pr-P. \end{aligned}$$

Furthermore, given Lemma 1, and using the same arguments as those used in Theorem 4 of Corradi and

Swanson (2007):

$$
\begin{aligned}
& Var^*\left(\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau,*}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-1\{X_{t+\tau}^*\leq u\}\right)1\left\{X_t^*\leq v\right\}\right.\right. \\
& \left.\left.-\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}\right)\right) \\
= & \frac{1}{T-\tau}\sum_{t=l}^{T-\tau-l}\sum_{k=1}^{l}\left(\left(\left(\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}-\mu(u,v)\right)\right.\right. \\
& \times\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau+k}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-1\{X_{t+\tau+k}\leq u\}\right)1\left\{X_{t+k}\leq v\right\}-\mu(u,v)\Bigg)+O\left(l/T^{1/2}\right),\ \Pr-P \\
= & Var\left(\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta^\dagger}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}\right)+O(l/T^{1/2}),\ \Pr-P,
\end{aligned}
$$

where $\mu(u,v) = E\left(\left(\frac{1}{S}\sum_{s=1}^{S}1\left\{X_{s,t+\tau}^{\theta^\dagger}\leq u\right\}-1\{X_{t+\tau}\leq u\}\right)1\left\{X_t\leq v\right\}\right)$. Thus, it remains to show that the term in the last line of (25) properly captures the contribution of parameter estimation error. Now, conditional on the resampling variability:

$$
\frac{1}{S}\sum_{s=1}^{S}\left(1\left\{X_{s,t+\tau,*}^{\widehat{\theta}^*_{T,N,h}}\leq u\right\}-F\left(u|X_t^*,\widehat{\theta}^*_{T,N,h}\right)\right)=O_S\left(S^{-1/2}\right),\ \Pr-P^*.
$$

Furthermore, conditional on sample and resampling variability:

$$
\frac{1}{S}\sum_{s=1}^{S}\left(1\left\{X_{s,t+\tau,*}^{\widehat{\theta}_{T,N,h}}\leq u\right\}-F\left(u|X_t^*,\widehat{\theta}_{T,N,h}\right)\right)=O_S\left(S^{-1/2}\right),\ \Pr-P^*,\ \Pr-P,
$$

where the subscript S denotes convergence in terms of the probability law governing the simulated randomness, which is independent of sampling and resampling (bootstrap) variability. Note that $F_\tau = F_{0,\tau}$ when the null is true. Thus, as $S/T\to\infty$, the last term on the RHS of (25) can be written as:

$$
\begin{aligned}
& -\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T}\left(F_\tau\left(u|X_t^*,\widehat{\theta}^*_{T,N,h}\right)-F_\tau\left(u|X_t^*,\widehat{\theta}_{T,N,h}\right)\right)1\{X_t^*\leq v\}+o_S(1)\ \Pr-P,\ \Pr-P^* \\
= & -\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T}f_\tau(u|X_t^*,\overline{\theta}^*_{T,N,h})\nabla'_\theta X_t^{*,\overline{\theta}^*_{T,N,h}}1\{X_t^*\leq v\}\left(\widehat{\theta}^*_{T,N,h}-\widehat{\theta}_{T,N,h}\right)+o_S(1)\ \Pr-P,\ \Pr-P^* \\
= & -E_X\left(f_{0,\tau}(u|X_1,\theta_0)E_s\left(\nabla_{\theta_0}X_{s,1+\tau}^{\theta_0}\right)1\{X_t\leq v\}\right)'\sqrt{T-\tau}\left(\widehat{\theta}^*_{T,N,h}-\widehat{\theta}_{T,N,h}\right)+o(1),\ \Pr-P,\ \Pr-P^*,
\end{aligned}
$$

where $\overline{\theta}^*_{T,N,h}\in\left(\widehat{\theta}^*_{T,N,h},\widehat{\theta}_{R,N,h}\right)$. As by Theorem 2.2 in Goncalves and White (2004), $\sqrt{T-\tau}\left(\widehat{\theta}^*_{T,N,h}-\widehat{\theta}_{T,N,h}\right)$ has the same limiting distribution as $\sqrt{T-\tau}\left(\widehat{\theta}_{T,N,h}-\theta^\dagger\right)$, the statement in the theorem then follows from Lemma 1.

Proof of Proposition 5: Note that:

$$
\begin{aligned}
& \frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} - F_\tau(u|X_t,\theta^\dagger) \\
= & \frac{1}{NS}\sum_{j=1}^{N}\left(\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} - F_\tau(u|X_t,V_{j,h}^{\widehat{\theta}_{T,N,h}},\theta^\dagger)\right) \\
& + \frac{1}{N}\sum_{j=1}^{N}\left(F_\tau(u|X_t,V_{j,h}^{\widehat{\theta}_{T,N,h}},\theta^\dagger) - F_\tau(u|X_t,\theta^\dagger)\right) = I + II
\end{aligned}
$$

Now,

$$
\begin{aligned}
II &= \frac{1}{N}\sum_{j=1}^{N}\left(F_\tau(u|X_t,V_{j,h}^{\widehat{\theta}_{T,N,h}},\theta^\dagger) - F_\tau(u|X_t,V_{j,h}^{\theta^\dagger},\theta^\dagger)\right) + \frac{1}{N}\sum_{j=1}^{N}\left(F_\tau(u|X_t,V_{j,h}^{\theta^\dagger},\theta^\dagger) - F_\tau(u|X_t,\theta^\dagger)\right) \\
&= o_P(1) + o_P(1), \text{ uniformly in } u,
\end{aligned}
$$

where the first $o_P(1)$ term follows from the fact that, given Assumption A', $\left(\widehat{\theta}_{T,N,h} - \theta^\dagger\right) = o_P(1)$; and the second $o_P(1)$ term follows from the uniform law of large numbers, given that $V_{j,h}^{\theta^\dagger}$ is a geometrically ergodic process. Also,

$$
\begin{aligned}
I &= \frac{1}{NS}\sum_{j=1}^{N}\left(\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} - F_\tau(u|X_t,V_{j,h}^{\widehat{\theta}_{T,N,h}},\theta^\dagger)\right) \\
&= \frac{1}{NS}\sum_{j=1}^{N}\left(\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\theta_i^\dagger} \leq u\right\} - F_\tau(u|X_t,V_{j,h}^{\theta^\dagger},\theta^\dagger)\right) + o_P(1),
\end{aligned}
$$

where the $o_P(1)$ term holds uniformly in u. Finally, the first term on the RHS above is $o_P(1)$, uniformly in u, by the uniform law of large numbers, given that:

$$
E\left(1\left\{X_{j,s,t+\tau}^{\theta^\dagger} \leq u\right\}|X_t,V_{j,h}^{\theta^\dagger}\right) = F_\tau(u|X_t,V_{j,h}^{\theta^\dagger},\theta^\dagger).
$$

Proof of Theorem 6: Note that:

$$
\begin{aligned}
& \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} - 1\{X_{t+\tau} \leq u\}\right) 1\{X_t \leq v\} \\
= & \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,s,t+\tau}^{\theta_0} \leq u\right\} - 1\{X_{t+\tau} \leq u\}\right) 1\{X_t \leq v\} \\
& + \frac{1}{\sqrt{T-\tau}}\sum_{t=R}^{T-\tau}\left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S}\left(1\left\{X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} - 1\left\{X_{j,s,t+\tau}^{\theta_0} \leq u\right\}\right)\right) 1\{X_t \leq v\} \\
= & I + II
\end{aligned}
$$

Recalling that as $T \to \infty$, $N/T \to \infty$ and $S/T \to \infty$, by a similar argument to that used in the proof of Theorem 3:

$$
II = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau} f_0\left(u|X_t\right) E_{j,s}\left(\nabla_{\theta_i} X_{s,1+\tau+j}^{\theta_0}\right)' 1\{X_t \leq v\}\sqrt{T-\tau}\left(\widehat{\theta}_{T,N,h} - \theta_0\right) + o_P(1),
$$

where $E_{j,s}(\cdot)$ denotes expectation with respect to the joint probability measure governing the simulated randomness and the volatility process, $V_{h,j}^{\theta_0}$, conditional on the sample. With regard to I, note that:

$$
\begin{aligned}
I &= \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau}^{\theta_0} \leq u \right\} - F_{0,\tau}\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) \right) 1\{X_t \leq v\} \\
&+ \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{N} \sum_{j=1}^{N} \left(F_{0,\tau}\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) - F_{0,\tau}\left(u|X_t, \theta_0 \right) \right) \right) 1\{X_t \leq v\} \\
&- \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(1\{X_{t+\tau} \leq u\} - F_{0,\tau}\left(u|X_t, \theta_0 \right) \right) 1\{X_t \leq v\}. \qquad (26)
\end{aligned}
$$

We need to show that the first and second terms on the RHS of (26) are $o_P(1)$. The statement in the theorem will then follow by the same argument as that used in the proof of Theorem 3. With regard to the second term on the RHS of (26), note that:

$$
\begin{aligned}
& E\left(\left(\frac{1}{N} \sum_{j=1}^{N} \left(F\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) - F_1\left(u|X_t, \theta_0 \right) \right) \right) 1\{X_t \leq v\} \right) \\
&= E\left(\frac{1}{N} \sum_{j=1}^{N} E_j\left(\left(F\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) - F\left(u|X_t, \theta_0 \right) \right) |X_t \right) 1\{X_t \leq v\} \right) = 0,
\end{aligned}
$$

where E_j denotes expectation with respect to the probability measure governing $V_{h,j}^{\theta_0}$, conditional on the sample. Also, note that $V_{h,j}^{\theta_0}$ and X_t are independent of each other, as the simulated randomness does not depend on X_t. Thus, by noting that $1\{X_t \leq v\}$ is equal to either 1 or 0, it suffices to show that:

$$
\lim_{T,N\to\infty} Var\left(\frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{N} \sum_{j=1}^{N} \left(F_\tau\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) - F_\tau\left(u|X_t, \theta_0 \right) \right) \right) \right) = 0.
$$

Now,

$$
\begin{aligned}
& Var\left(\frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{N} \sum_{j=1}^{N} \left(F_\tau\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) - F_\tau\left(u|X_t, \theta_0 \right) \right) \right) \right) \\
&= \frac{1}{TN^2} \sum_{t=1}^{T-\tau} \sum_{l=1}^{T-\tau} \sum_{j=1}^{N} \sum_{s=j}^{N} E\left(\left(F_\tau\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) - F_\tau\left(u|X_t, \theta_0 \right) \right) \left(F_\tau\left(u|X_l, V_{h,s}^{\theta_0}, \theta_0 \right) - F_\tau\left(u|X_l, \theta_0 \right) \right) \right) \\
&+ \frac{1}{TN^2} \sum_{t=1}^{T-\tau} \sum_{l=1}^{T-\tau} \sum_{j=1}^{N} \sum_{s\neq j}^{N} E\left(\left(F_\tau\left(u|X_t, V_{h,j}^{\theta_0}, \theta_0 \right) - F_\tau\left(u|X_t, \theta_0 \right) \right) \left(F_\tau\left(u|X_l, V_{h,s}^{\theta_0}, \theta_0 \right) - F_\tau\left(u|X_l, \theta_0 \right) \right) \right). \qquad (27)
\end{aligned}
$$

The first term on the RHS of (27) is $O\left(\frac{T}{N}\right) = o(1)$, for $T/N \to 0$. Given that $V_{h,j}^{\theta_0}$ is a geometric mixing process, it can be shown via standard mixing inequality arguments that the second term on the RHS of (27) is $O\left(\frac{T}{N}\right) = o(1)$, for $T/N \to 0$.

Now, consider the first term on the RHS of (26). It can be shown that this term is $o_P(1)$ by a similar argument as that used in the proof of Theorem 3. Finally, the statement under the alternative follows by the same argument as that used in the proof of Theorem 3.

Proof of Theorem 7: Note that:

$$
\begin{aligned}
& SV_T^*(u,v) \\
= \quad & \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau,*}^{\widehat{\theta}_{T,N,h}} \leq u \right\} - 1\{X_{t+\tau}^* \leq u\} \right) 1\left\{X_t^* \leq v\right\} \\
& - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u \right\} - 1\{X_{t+\tau} \leq u\} \right) 1\left\{X_t \leq v\right\} \\
& + \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau,*}^{\widehat{\theta}^*_{T,N,h}} \leq u \right\} - \frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau,*}^{\widehat{\theta}_{T,N,h}} \leq u \right\} \right) \\
& \times 1\left\{X_t^* \leq v\right\}.
\end{aligned} \tag{28}
$$

By the same argument as that used in the proof of Theorem 4, the first two terms on the RHS of (28) have the the same limiting distribution as:

$$
\begin{aligned}
& \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau}^{\theta^\dagger} \leq u \right\} - 1\{X_{t+\tau} \leq u\} \right) \right. \\
& \left. - E\left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1\left\{ X_{j,s,t+\tau}^{\theta^\dagger} \leq u \right\} - 1\{X_{t+\tau} \leq u\} \right) \right) 1\left\{X_t \leq v\right\},
\end{aligned}
$$

conditional on the sample. Furthermore, given Proposition 5, the last term on the RHS of (28) properly mimics the contribution of parameter estimation error, by the same argument as that used in the proof of Theorem 4.

9 References

Aït-Sahalia, Y., (1996), Testing Continuous Time Models of the Spot Interest Rate, *Review of Financial Studies*, 9, 385-426.

Aït-Sahalia, Y., (1999), Transition Densities for Interest Rate and Others Nonlinear Diffusions, *Journal of Finance,* LIV, 1361-1395.

Aït-Sahalia, Y., (2002), Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed Form Approximation Approach, *Econometrica,* 70, 223-262.

Aït-Sahalia, Y., J. Fan and H. Peng (2005). Nonparametric Transition-Based Tests for Diffusions. Manuscript, Princeton University.

Altissimo, F., and A. Mele, (2005), Simulated Nonparametric Estimation of Continuous Time Models of Asset Prices and Returns, Working Paper, London School of Economics.

Andersen, T.G., L. Benzoni, J. Lund, (2004), Stochastic Volatility, Mean Drift, and Jumps in the Short-Term Interest Rate, Working Paper, Northwestern University.

Andersen, T.G., and B.E. Sorensen, (1996), GMM Estimation of a Stochastic Volatility Model: a Monte Carlo Study, *Journal of Business and Economic Statistics,* 14, 328-352.

Andrews, D.W.K., (1997), A Conditional Kolmogorov Test, *Econometrica,* 65, 1097-1128.

Bai, J., (2003), Testing Parametric Conditional Distributions of Dynamic Models, *Review of Economics and Statistics,* 85, 531-549.

Bandi, F.M., (2002), Short-Term Interest Rate Dynamics: A Spatial Approach, *Journal of Financial Economics,* 65, 73-110.

Bhardwaj, G., V. Corradi, and N.R. Swanson, (2005), A Simulation Based Specification Test for Diffusion Processes, Working Paper, Rutgers University.

Bontemps, C., and N. Meddahi, (2005), Testing Normality: a GMM Approach, *Journal of Econometrics*, 124, 149-186.

Chacko, G., and L.M. Viceira, (2003), Spectral GMM estimation of Continuous-Time Processes, *Journal of Econometrics,* 116, 259-292.

Chib, S., S. Kim, and N. Shephard (1998), Stochastic Volatility: Likelihood Inference and Comparison with ARCH models, *Review of Economic Studies,* 65, 361-393.

Corradi, V. and N.R. Swanson, (2005a), A Bootstrap Specification Test for Diffusion Processes, *Journal of Econometrics,* 124, 117-148.

Corradi, V. and N.R. Swanson, (2005b), A Test for Comparing Multiple Misspecified Conditional Interval Models, *Econometric Theory*, 21, 991-1016.

Corradi, V., and N.R. Swanson, (2006a), Bootstrap Conditional Distribution Tests In the Presence of Dynamic Misspecification, *Journal of Econometrics*, 133, 779-806.

Corradi, V. and N.R. Swanson, (2006b), Simulation Based Predictive Density Estimation and Testing for Diffusion Processes, Working Paper, Rutgers University and Warwick University.

Corradi, V. and N.R. Swanson, (2007), Evaluation of Dynamic Stochastic General Equilibrium Models Bases on Distributional Comparison of Simulated and Historical Data, *Journal of Econometrics,* 136, 699-723.

Cox, J.C., J.E. Ingersoll and S.A. Ross, (1985), A Theory of the Term Structure of Interest Rates, *Econometrica,* 53, 385-407.

Dai, Q., and Kenneth J. Singleton, (2000), Specification Analysis of Affine Term Structure Models, *Journal of Finance* 55, 1943-1978.

Diebold, F.X., T. Gunther and A.S. Tay, (1998), Evaluating Density Forecasts with Applications to Finance and Management, *International Economic Review* 39, 863-883.

Dridi, R., A. Guay and E. Renault, (2006), Indirect Inference and Calibration of Dynamic Stochastic General Equilibrium Models, *Journal of Econometrics,* forthcoming.

Duan, J.C., (2003), A Specification Test for Time Series Models by a Normality Transformation, Working Paper, University of Toronto.

Duffie, D. and K. Singleton, (1993), Simulated Moment Estimation of Markov Models of Asset Prices, *Econometrica,* 61, 929-952.

Elerian, O., S. Chib, and N. Shephard, (2002), Likelihood Inference for Discretely Observed Nonlinear Diffusion, *Econometrica,* 69, 959-994.

Fan, J., Q. Yao and H. Tong (1996). Estimation of Conditional Densities and Sensitivity Measures in Nonlinear Dynamical Systems. *Biometrika,* 83, 189-206.

Gallant, A.R. and G. Tauchen, (1996), Which Moments to Match, *Econometric Theory,* 12, 657-681.

Gonçalves, S. and H. White, (2004), Maximum Likelihood and the Bootstrap for Dynamic Nonlinear Models, *Journal of Econometrics*, 119, 199-219.

Hall, P., R.C.L. Wolff and Q. Yao (1999). Methods for Estimating a Conditional Distribution Function. *Journal of the American Statistical Association,* 94, 154-163.

Hall, A.R., and A. Inoue, (2003), The Large Sample Behavior of the Generalized Method of Moments Estimator in Misspecified Models, *Journal of Econometrics,* 361-394.

Heston, S.L., (1993), A Closed Form Solution for Option with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies,* 6, 327-344.

Hong, Y., (2001), Evaluation of Out of Sample Probability Density Forecasts with Applications to S&P 500 Stock Prices, Working Paper, Cornell University.

Hong, Y.M., and H.T. Li, (2005), Nonparametric Specification Testing for Continuous Time Models with Applications to Term Structure Interest Rates, *Review of Financial Studies*, 18, 37-84.

Hong, Y.M., H. Li, and F. Zhao, (2004) Out of Sample Performance of Spot Interest Rate Models, *Journal of Business and Economic Statistics,* 22, 457-473.

Jiang, G.J., and J.L. Knigth, (2002), Estimation of Continuous-Time Processes via the Empirical Characteristic Function, *Journal of Business and Economic Statistics,* 20, 198-212.

Karlin, S., and H.M. Taylor, (1981), *A Second Course in Stochastic Processes,* Academic Press, San Diego.

Kloeden, P.E., and E. Platen, (1999), *Numerical Solution of Stochastic Differential Equations*, Springer and Verlag, New York.

Meddahi, N., (2001), An Eigenfunction Approach for Volatility Modeling, Working Paper, University of Montreal.

Nelson, D.B., (1990), ARCH as Diffusion Approximations, *Journal of Econometrics*, 45, 7-38.

Pardoux, E. and D. Talay, (1985), Discretization and Simulation of Stochastic Differential Equations, *Acta Applicandae Mathematicae,* 3, 23-47.

Pitt, M.K., and N. Shephard, (1999), Filtering via Simulation-Based on Auxiliary Particle Filters, *Journal of the American Statistical Association,* 94, 590-599.

Pitt, M.K., (2005), Smooth Particle Filters for Likelihood Evaluation and Maximizations, Working Paper, University of Warwick.

Pritsker, M., (1998), Nonparametric Density Estimation and Tests of Continuous-Time Interest Rate Models, *The Review of Financial Studies*, 11, 449-87.

Singleton, K.J., (2001), Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function, *Journal of Econometrics,* 2001, 111-141.

Thompson, S.B., (2004), Identifying Term Structure Volatility from the LIBOR-Swap Curve, Working Paper, Harvard University.

Table 1: Specification Test Rejection Frequencies for the One Factor Model - Empirical Size[(*)]

Data Generated using the *CIR* $(0.15, 0.05, 0.10)$ Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
				Panel A: $T = 400$			
1	$\overline{X} \pm 0.5\sigma_X$	0.1559	0.2383	0.177	0.1503	0.2191	0.2831
	$\overline{X} \pm \sigma_X$	0.1559	0.2211	0.1638	0.1539	0.2105	0.1611
2	$\overline{X} \pm 0.5\sigma_X$	0.1796	0.2534	0.2303	0.2827	0.1492	0.1441
	$\overline{X} \pm \sigma_X$	0.1459	0.2009	0.1854	0.2256	0.1040	0.1468
4	$\overline{X} \pm 0.5\sigma_X$	0.1826	0.1963	0.1849	0.1341	0.1560	0.2107
	$\overline{X} \pm \sigma_X$	0.1933	0.1545	0.1868	0.1658	0.1158	0.1484
12	$\overline{X} \pm 0.5\sigma_X$	0.2142	0.2697	0.2372	0.283	0.3113	0.2443
	$\overline{X} \pm \sigma_X$	0.1374	0.1804	0.2836	0.2439	0.2666	0.1949
				Panel B: $T = 800$			
1	$\overline{X} \pm 0.5\sigma_X$	0.1149	0.1266	0.1267	0.1197	0.1260	0.1177
	$\overline{X} \pm \sigma_X$	0.1313	0.1199	0.1203	0.1291	0.1110	0.1513
2	$\overline{X} \pm 0.5\sigma_X$	0.1332	0.1058	0.1151	0.1291	0.1313	0.1110
	$\overline{X} \pm \sigma_X$	0.1462	0.1068	0.1217	0.1383	0.1388	0.1078
4	$\overline{X} \pm 0.5\sigma_X$	0.1089	0.1327	0.1193	0.1227	0.1274	0.1016
	$\overline{X} \pm \sigma_X$	0.1222	0.1056	0.1231	0.1071	0.1275	0.1210
12	$\overline{X} \pm 0.5\sigma_X$	0.1373	0.1409	0.132	0.1154	0.1028	0.1225
	$\overline{X} \pm \sigma_X$	0.1269	0.1369	0.134	0.1091	0.1032	0.1067

Data Generated using the *CIR* $(0.30, 0.05, 0.10)$ Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
				Panel A: $T = 400$			
1	$\overline{X} \pm 0.5\sigma_X$	0.1804	0.2433	0.1747	0.1523	0.2382	0.2848
	$\overline{X} \pm \sigma_X$	0.1744	0.2204	0.1868	0.1649	0.2325	0.1781
2	$\overline{X} \pm 0.5\sigma_X$	0.1715	0.2772	0.2502	0.2887	0.1509	0.1577
	$\overline{X} \pm \sigma_X$	0.1654	0.2084	0.2077	0.2391	0.1381	0.1714
4	$\overline{X} \pm 0.5\sigma_X$	0.1971	0.2010	0.1725	0.1562	0.1455	0.2260
	$\overline{X} \pm \sigma_X$	0.2085	0.1609	0.1998	0.1770	0.1316	0.1692
12	$\overline{X} \pm 0.5\sigma_X$	0.2199	0.2681	0.2626	0.2988	0.3112	0.2617
	$\overline{X} \pm \sigma_X$	0.1588	0.1869	0.2881	0.2478	0.2824	0.1994
				Panel B: $T = 800$			
1	$\overline{X} \pm 0.5\sigma_X$	0.1205	0.1291	0.1382	0.1203	0.1412	0.1106
	$\overline{X} \pm \sigma_X$	0.1291	0.1199	0.1253	0.1447	0.1133	0.1417
2	$\overline{X} \pm 0.5\sigma_X$	0.1354	0.1118	0.1198	0.1265	0.1390	0.1180
	$\overline{X} \pm \sigma_X$	0.1397	0.1248	0.1139	0.1313	0.1443	0.1168
4	$\overline{X} \pm 0.5\sigma_X$	0.1161	0.1251	0.1298	0.1313	0.1289	0.1259
	$\overline{X} \pm \sigma_X$	0.1363	0.1108	0.1324	0.1230	0.1189	0.1206
12	$\overline{X} \pm 0.5\sigma_X$	0.1355	0.1421	0.1297	0.1246	0.1180	0.1254
	$\overline{X} \pm \sigma_X$	0.1373	0.1474	0.1283	0.1130	0.1148	0.1348

Data Generated using the *CIR* $(0.50, 0.05, 0.10)$ Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
				Panel A: $T = 400$			
1	$\overline{X} \pm 0.5\sigma_X$	0.2011	0.247	0.1886	0.1633	0.244	0.2842
	$\overline{X} \pm \sigma_X$	0.1908	0.2383	0.1944	0.1814	0.2403	0.1756
2	$\overline{X} \pm 0.5\sigma_X$	0.1904	0.2783	0.2492	0.2978	0.1452	0.1666
	$\overline{X} \pm \sigma_X$	0.1781	0.219	0.2175	0.261	0.1547	0.181
4	$\overline{X} \pm 0.5\sigma_X$	0.2141	0.2109	0.1868	0.1783	0.1563	0.2221
	$\overline{X} \pm \sigma_X$	0.2204	0.1593	0.2075	0.2005	0.1523	0.1977
12	$\overline{X} \pm 0.5\sigma_X$	0.2257	0.2766	0.2767	0.3171	0.3345	0.2653
	$\overline{X} \pm \sigma_X$	0.162	0.1917	0.3012	0.2662	0.2962	0.2035
				Panel B: $T = 800$			
1	$\overline{X} \pm 0.5\sigma_X$	0.1261	0.1368	0.1615	0.1324	0.1568	0.1215
	$\overline{X} \pm \sigma_X$	0.118	0.1248	0.1381	0.1469	0.1225	0.1682
2	$\overline{X} \pm 0.5\sigma_X$	0.1437	0.1378	0.1231	0.1483	0.1608	0.1209
	$\overline{X} \pm \sigma_X$	0.1625	0.1429	0.1124	0.1553	0.1585	0.1202
4	$\overline{X} \pm 0.5\sigma_X$	0.1319	0.1248	0.1451	0.1525	0.1388	0.1541
	$\overline{X} \pm \sigma_X$	0.1428	0.1263	0.1517	0.1317	0.1364	0.1167
12	$\overline{X} \pm 0.5\sigma_X$	0.1406	0.1432	0.128	0.1202	0.145	0.1293
	$\overline{X} \pm \sigma_X$	0.1399	0.1688	0.1409	0.1216	0.1305	0.1294

[(*)] Notes: Entries in the table are empirical rejection frequencies for tests constructed using intervals given in the second column of the table, and for $\tau = 1, 2, 4, 12$. (S, l) combinations used in test construction are given in the second row of the table, so that simulation periods considered are $S = (10T, 20T)$ and block lengths considered are $l = (10, 20, 50)$, where T is the sample size, and $T = 400, 800$. Empirical bootstrap distributions are constructed using 100 bootstrap replications, and critical values are set equal to the 90^{th} percentile of the bootstrap distribution. Finally, $\overline{X}$ and σ_X are the mean and variance of an initial sample of data. All results are based on 500 Monte Carlo simulations. See Section 6.1 for further details.

Table 2: Specification Test Rejection Frequencies For the Two Factor Models - Empirical Size[(*)]

Data Generated using the SV Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
				Panel A: $T = 400$			
1	$\overline{X} \pm 0.5\sigma_X$	0.2995	0.1698	0.1752	0.1765	0.3532	0.2132
	$\overline{X} \pm \sigma_X$	0.2266	0.2467	0.2585	0.1550	0.1506	0.2040
2	$\overline{X} \pm 0.5\sigma_X$	0.2141	0.2856	0.2328	0.2814	0.1912	0.3257
	$\overline{X} \pm \sigma_X$	0.1396	0.2107	0.2219	0.2264	0.1597	0.2196
4	$\overline{X} \pm 0.5\sigma_X$	0.2074	0.1941	0.2803	0.1313	0.2057	0.1475
	$\overline{X} \pm \sigma_X$	0.1770	0.1117	0.2266	0.1205	0.1550	0.1974
12	$\overline{X} \pm 0.5\sigma_X$	0.1910	0.2917	0.2056	0.2317	0.2435	0.2680
	$\overline{X} \pm \sigma_X$	0.1355	0.1497	0.1748	0.1990	0.2117	0.1168
				Panel B: $T = 800$			
1	$\overline{X} \pm 0.5\sigma_X$	0.1405	0.1584	0.1299	0.1366	0.1491	0.1409
	$\overline{X} \pm \sigma_X$	0.1282	0.1140	0.1271	0.1430	0.1208	0.1192
2	$\overline{X} \pm 0.5\sigma_X$	0.1048	0.1493	0.1169	0.1228	0.1203	0.1107
	$\overline{X} \pm \sigma_X$	0.1167	0.1548	0.1159	0.1275	0.1112	0.1165
4	$\overline{X} \pm 0.5\sigma_X$	0.1035	0.1183	0.1312	0.1416	0.1055	0.1276
	$\overline{X} \pm \sigma_X$	0.1173	0.1269	0.1329	0.1196	0.1123	0.1017
12	$\overline{X} \pm 0.5\sigma_X$	0.1207	0.1324	0.1043	0.1584	0.1033	0.1104
	$\overline{X} \pm \sigma_X$	0.1178	0.1071	0.1258	0.1058	0.1121	0.1277

Data Generated using the SVJ Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
				Panel A: $T = 400$			
1	$\overline{X} \pm 0.5\sigma_X$	0.2768	0.2686	0.2340	0.2114	0.3675	0.1569
	$\overline{X} \pm \sigma_X$	0.1939	0.1626	0.2192	0.2439	0.2342	0.1355
2	$\overline{X} \pm 0.5\sigma_X$	0.2112	0.3244	0.2442	0.1924	0.1727	0.2838
	$\overline{X} \pm \sigma_X$	0.1456	0.2375	0.2978	0.1346	0.1572	0.2124
4	$\overline{X} \pm 0.5\sigma_X$	0.2078	0.2898	0.1467	0.2099	0.1929	0.1839
	$\overline{X} \pm \sigma_X$	0.2927	0.1512	0.1189	0.1381	0.2961	0.1453
12	$\overline{X} \pm 0.5\sigma_X$	0.2102	0.1667	0.1881	0.1228	0.2757	0.3071
	$\overline{X} \pm \sigma_X$	0.2029	0.1459	0.3656	0.1784	0.2463	0.2439
				Panel B: $T = 800$			
1	$\overline{X} \pm 0.5\sigma_X$	0.1533	0.1328	0.1433	0.1280	0.1320	0.1099
	$\overline{X} \pm \sigma_X$	0.1068	0.1214	0.1397	0.1151	0.1228	0.1363
2	$\overline{X} \pm 0.5\sigma_X$	0.1263	0.1259	0.1191	0.1389	0.1350	0.1187
	$\overline{X} \pm \sigma_X$	0.1179	0.1122	0.1134	0.1164	0.1115	0.1183
4	$\overline{X} \pm 0.5\sigma_X$	0.1403	0.1541	0.1595	0.1262	0.1597	0.1394
	$\overline{X} \pm \sigma_X$	0.1178	0.1248	0.1185	0.1152	0.1131	0.1130
12	$\overline{X} \pm 0.5\sigma_X$	0.1248	0.1042	0.1249	0.1432	0.1110	0.1515
	$\overline{X} \pm \sigma_X$	0.1187	0.1120	0.1135	0.1188	0.1187	0.1062

[(*)] Notes: See notes to Table 1.

Table 3: Specification Test Rejection Frequencies - Empirical Power[(*)]

Data Generated using the *CIR* Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
Panel A: $T = 400$							
1	$\overline{X} \pm 0.5\sigma_X$	0.3784	0.3425	0.4643	0.4662	0.4323	0.4582
	$\overline{X} \pm \sigma_X$	0.2547	0.2786	0.3222	0.3444	0.3443	0.3287
2	$\overline{X} \pm 0.5\sigma_X$	0.3782	0.3766	0.4483	0.4486	0.4568	0.4643
	$\overline{X} \pm \sigma_X$	0.2361	0.2786	0.3444	0.3328	0.3584	0.3587
4	$\overline{X} \pm 0.5\sigma_X$	0.3482	0.3727	0.4429	0.4584	0.4580	0.4267
	$\overline{X} \pm \sigma_X$	0.2345	0.2621	0.3189	0.3484	0.3169	0.3525
12	$\overline{X} \pm 0.5\sigma_X$	0.3747	0.3681	0.4644	0.4381	0.4248	0.4583
	$\overline{X} \pm \sigma_X$	0.2580	0.2485	0.3282	0.3161	0.3480	0.3681
Panel B: $T = 800$							
1	$\overline{X} \pm 0.5\sigma_X$	0.8544	0.9261	0.8225	0.8442	0.8023	0.9285
	$\overline{X} \pm \sigma_X$	0.7125	0.7342	0.8727	0.8164	0.9404	0.8166
2	$\overline{X} \pm 0.5\sigma_X$	0.8067	0.8164	0.8864	0.9348	0.8185	0.9028
	$\overline{X} \pm \sigma_X$	0.8382	0.7667	0.9200	0.9360	0.9465	0.9188
4	$\overline{X} \pm 0.5\sigma_X$	0.8567	0.8884	0.8181	0.8969	0.9168	0.9062
	$\overline{X} \pm \sigma_X$	0.8329	0.7768	0.9367	0.8840	0.8383	0.9183
12	$\overline{X} \pm 0.5\sigma_X$	0.8744	0.9328	0.9442	0.9024	0.8924	0.8481
	$\overline{X} \pm \sigma_X$	0.7244	0.7223	0.9383	0.8163	0.8167	0.8143

Data Generated using the *SV* Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
Panel A: $T = 400$							
1	$\overline{X} \pm 0.5\sigma_X$	0.5611	0.5613	0.5534	0.5491	0.5299	0.5421
	$\overline{X} \pm \sigma_X$	0.4589	0.4355	0.4248	0.5007	0.4461	0.4546
2	$\overline{X} \pm 0.5\sigma_X$	0.4555	0.5396	0.5204	0.5200	0.5374	0.4363
	$\overline{X} \pm \sigma_X$	0.4528	0.3486	0.4202	0.4361	0.4018	0.4062
4	$\overline{X} \pm 0.5\sigma_X$	0.5423	0.5666	0.4707	0.4851	0.5085	0.5612
	$\overline{X} \pm \sigma_X$	0.4217	0.4226	0.4445	0.4301	0.4279	0.4461
12	$\overline{X} \pm 0.5\sigma_X$	0.4561	0.4485	0.4364	0.4820	0.4307	0.4534
	$\overline{X} \pm \sigma_X$	0.3781	0.3431	0.3937	0.3537	0.3348	0.3719
Panel B: $T = 800$							
1	$\overline{X} \pm 0.5\sigma_X$	0.6483	0.6885	0.6272	0.6962	0.6708	0.6362
	$\overline{X} \pm \sigma_X$	0.5309	0.5728	0.5139	0.5457	0.5859	0.6005
2	$\overline{X} \pm 0.5\sigma_X$	0.6324	0.6021	0.5930	0.6038	0.6048	0.6194
	$\overline{X} \pm \sigma_X$	0.5057	0.5169	0.5258	0.5160	0.5466	0.5580
4	$\overline{X} \pm 0.5\sigma_X$	0.5931	0.6143	0.6098	0.5909	0.6539	0.6115
	$\overline{X} \pm \sigma_X$	0.5039	0.5286	0.5508	0.5651	0.5783	0.5569
12	$\overline{X} \pm 0.5\sigma_X$	0.6598	0.6104	0.6242	0.6319	0.636	0.6936
	$\overline{X} \pm \sigma_X$	0.5614	0.5385	0.5725	0.5995	0.5891	0.6061

Data Generated using the *SVJ* Model

τ	$(\underline{u}, \overline{u})$	S,l					
		10T,10	20T,10	10T,20	20T,20	10T,50	20T,50
Panel A: $T = 400$							
1	$\overline{X} \pm 0.5\sigma_X$	0.4581	0.5514	0.4988	0.4942	0.4867	0.4901
	$\overline{X} \pm \sigma_X$	0.4261	0.4358	0.4080	0.4035	0.3553	0.3857
2	$\overline{X} \pm 0.5\sigma_X$	0.4724	0.4724	0.5173	0.5236	0.5477	0.5030
	$\overline{X} \pm \sigma_X$	0.3650	0.3461	0.4072	0.4182	0.4020	0.4189
4	$\overline{X} \pm 0.5\sigma_X$	0.5364	0.5659	0.5599	0.5442	0.4206	0.5604
	$\overline{X} \pm \sigma_X$	0.3912	0.3470	0.3945	0.4259	0.3202	0.4005
12	$\overline{X} \pm 0.5\sigma_X$	0.4327	0.4406	0.5387	0.4920	0.4657	0.5130
	$\overline{X} \pm \sigma_X$	0.4192	0.3343	0.4005	0.3420	0.4492	0.4466
Panel B: $T = 800$							
1	$\overline{X} \pm 0.5\sigma_X$	0.6053	0.7094	0.7019	0.6936	0.5708	0.5944
	$\overline{X} \pm \sigma_X$	0.5566	0.6406	0.5721	0.6083	0.5351	0.5765
2	$\overline{X} \pm 0.5\sigma_X$	0.6029	0.6983	0.6716	0.6260	0.5841	0.6327
	$\overline{X} \pm \sigma_X$	0.5511	0.5362	0.5762	0.5471	0.4741	0.5956
4	$\overline{X} \pm 0.5\sigma_X$	0.6176	0.6980	0.6109	0.6919	0.6546	0.6748
	$\overline{X} \pm \sigma_X$	0.4757	0.5663	0.5254	0.5675	0.4627	0.5117
12	$\overline{X} \pm 0.5\sigma_X$	0.6964	0.6304	0.5953	0.6155	0.7165	0.7053
	$\overline{X} \pm \sigma_X$	0.5657	0.5960	0.4882	0.5848	0.6158	0.6036

[(*)] Notes: See notes to Table 1.

Table 4: Empirical Illustration - Specification Testing Using 1 and 2-Factor Models[(*)]

Specification Test Results - CIR Model

τ	$(\underline{u},\overline{u})$	$S=10T$ V_T	5% CV	10% CV	20% CV	$S=20T$ V_T	5% CV	10% CV	20% CV	$S=30T$ V_T	5% CV	10% CV	20% CV
						Panel A: $l=25$							
1	$\overline{X}\pm 0.5\sigma_X$	0.5274***	0.2906	0.3545	0.3705	0.5046***	0.2858	0.3980	0.4762	0.4923**	0.3682	0.4768	0.5349
	$\overline{X}\pm \sigma_X$	0.4289***	0.2658	0.3178	0.3571	0.4524***	0.2948	0.3568	0.4448	0.4655***	0.2596	0.3635	0.4322
2	$\overline{X}\pm 0.5\sigma_X$	0.6824***	0.4291	0.4911	0.6321	0.6973***	0.4731	0.5509	0.6568	0.6075**	0.3994	0.5134	0.6866
	$\overline{X}\pm \sigma_X$	0.4897*	0.4264	0.5182	0.5695	0.4601	0.4660	0.5040	0.5868	0.4985**	0.3197	0.3560	0.5693
4	$\overline{X}\pm 0.5\sigma_X$	0.8662**	0.7111	0.8491	0.9290	0.8813**	0.6779	0.7962	0.9459	0.8726**	0.6094	0.6247	0.9702
	$\overline{X}\pm \sigma_X$	0.8539*	0.7521	0.9389	1.0655	0.8153*	0.7127	0.9330	1.0228	0.8595**	0.6875	0.8581	1.0370
12	$\overline{X}\pm 0.5\sigma_X$	1.1631*	1.0087	1.3009	2.0108	1.2236*	0.9669	1.2932	2.0232	1.2432**	0.9525	1.1562	1.9910
	$\overline{X}\pm \sigma_X$	1.0429	1.4767	2.0222	2.2371	1.0731	1.4798	2.0401	2.2706	1.0387	1.5007	2.0335	2.2841
						Panel B: $l=50$							
1	$\overline{X}\pm 0.5\sigma_X$	0.5274***	0.2924	0.3523	0.4499	0.5046***	0.3883	0.4440	0.4593	0.4923*	0.3359	0.5749	0.7234
	$\overline{X}\pm \sigma_X$	0.4289**	0.2323	0.3325	0.4544	0.4524***	0.2787	0.3584	0.4048	0.4655***	0.2167	0.297	0.4348
2	$\overline{X}\pm 0.5\sigma_X$	0.6824***	0.4043	0.4915	0.5766	0.6973***	0.4538	0.5141	0.6910	0.6075**	0.3198	0.4683	0.6987
	$\overline{X}\pm \sigma_X$	0.4897*	0.4236	0.5594	0.5771	0.4601**	0.3853	0.4574	0.5818	0.4985**	0.2488	0.3960	0.5968
4	$\overline{X}\pm 0.5\sigma_X$	0.8662*	0.7073	0.9498	1.0332	0.8813**	0.6515	0.7367	0.9357	0.8726**	0.5871	0.5917	0.9723
	$\overline{X}\pm \sigma_X$	0.8539*	0.7321	0.9371	1.1302	0.8153*	0.7065	0.9404	1.0128	0.8595**	0.7281	0.8055	1.0012
12	$\overline{X}\pm 0.5\sigma_X$	1.1631*	1.0750	1.3256	2.0626	1.2236*	0.9979	1.2570	1.9836	1.2432*	0.9378	1.2776	2.0647
	$\overline{X}\pm \sigma_X$	1.0429	1.4721	2.0165	2.2850	1.0731	1.4552	2.0157	2.3086	1.0387	1.5280	2.0071	2.2379

Specification Test Results - SV Model

τ	$(\underline{u},\overline{u})$	$S=10T$ V_T	5% CV	10% CV	20% CV	$S=20T$ V_T	5% CV	10% CV	20% CV	$S=30T$ V_T	5% CV	10% CV	20% CV
						Panel A: $l=25$							
1	$\overline{X}\pm 0.5\sigma_X$	0.9841***	0.8729	0.9031	0.9242	0.9453***	0.7485	0.7986	0.8439	0.9112**	0.8286	0.8328	1.0113
	$\overline{X}\pm \sigma_X$	0.6870	0.6954	0.7254	0.7379	0.7276	0.7345	0.7674	0.8180	0.7775**	0.6798	0.7608	0.8175
2	$\overline{X}\pm 0.5\sigma_X$	0.4113	1.3751	1.4900	1.5535	1.0265	1.3769	1.4124	1.5298	0.9641	1.4032	1.4808	1.6647
	$\overline{X}\pm \sigma_X$	0.3682	1.1933	1.2243	1.2918	0.8390	1.2019	1.4938	1.6579	0.8295	1.2534	1.5048	1.6388
4	$\overline{X}\pm 0.5\sigma_X$	1.2840	2.3297	2.6109	2.6396	1.0835	2.2812	2.5397	2.7109	1.6839	2.312	2.5685	2.7955
	$\overline{X}\pm \sigma_X$	1.0472	2.2549	2.2745	2.3670	1.0110	2.0244	2.2695	2.3931	1.1328	1.9523	2.3104	2.4084
12	$\overline{X}\pm 0.5\sigma_X$	1.7687	4.9298	5.2832	5.8204	1.7135	4.9302	5.3526	5.8164	2.6901	5.0453	5.345	5.6689
	$\overline{X}\pm \sigma_X$	1.7017	5.2601	5.6522	5.8351	1.4404	5.2686	5.6279	5.8831	1.7675	5.2943	5.6733	5.7644
						Panel B: $l=50$							
1	$\overline{X}\pm 0.5\sigma_X$	0.9841***	0.8140	0.8988	0.9228	0.9453**	0.7710	0.8093	0.9815	0.9112**	0.684	0.7988	1.0574
	$\overline{X}\pm \sigma_X$	0.6870**	0.6106	0.6597	0.7605	0.7276	0.7549	0.7903	0.8806	0.7775*	0.7118	0.8269	0.8707
2	$\overline{X}\pm 0.5\sigma_X$	0.4113	1.3606	1.4466	1.5524	1.0265	1.3149	1.3969	1.6394	0.9641	1.3923	1.4365	1.6937
	$\overline{X}\pm \sigma_X$	0.3682	1.2016	1.4673	1.6917	0.8390	1.1785	1.4975	1.7076	0.8295	1.3235	1.3444	1.5953
4	$\overline{X}\pm 0.5\sigma_X$	1.2840	2.3183	2.5657	2.6907	1.0835	2.2977	2.6108	2.8177	1.6839	2.1829	2.4884	2.666
	$\overline{X}\pm \sigma_X$	1.0472	2.0725	2.2299	2.3655	1.0110	1.9547	2.3095	2.3837	1.1328	1.9125	2.3672	2.3968
12	$\overline{X}\pm 0.5\sigma_X$	1.7687	4.9526	5.2820	5.7781	1.7135	4.9679	5.3347	5.7549	2.6901	5.0203	5.4429	5.6651
	$\overline{X}\pm \sigma_X$	1.7017	5.2707	5.6487	5.8446	1.4404	5.2283	5.6249	5.7786	1.7675	5.3325	5.6787	5.8608

Specification Test Results - SVJ Model

τ	$(\underline{u},\overline{u})$	$S=10T$ V_T	5% CV	10% CV	20% CV	$S=20T$ V_T	5% CV	10% CV	20% CV	$S=30T$ V_T	5% CV	10% CV	20% CV
						Panel A: $l=25$							
1	$\overline{X}\pm 0.5\sigma_X$	1.1319	1.8468	2.1957	2.6832	1.1787	1.9065	2.1342	2.9061	1.1655	1.6929	2.1594	2.3706
	$\overline{X}\pm \sigma_X$	1.2272*	1.1203	1.3031	1.5307	1.0220	1.0304	1.1669	1.6448	0.9906	1.1359	1.2893	1.7683
2	$\overline{X}\pm 0.5\sigma_X$	0.9615*	0.8146	1.1334	1.3568	1.0150*	0.8902	1.0677	1.4168	1.0528*	0.7372	1.0903	1.5176
	$\overline{X}\pm \sigma_X$	1.2571	1.3316	1.4096	1.8620	1.1491	1.2997	1.4264	1.8372	1.1562	1.3130	1.5162	1.8557
4	$\overline{X}\pm 0.5\sigma_X$	1.5012*	1.1188	1.6856	2.0371	1.3255	1.4806	1.6501	1.9897	1.3545*	1.2608	1.5410	2.0618
	$\overline{X}\pm \sigma_X$	0.9901*	0.9793	1.0507	1.4065	1.0180*	0.9480	1.0400	1.292	0.6941	0.7759	0.9318	1.1776
12	$\overline{X}\pm 0.5\sigma_X$	2.4237*	2.0818	3.0640	3.1095	2.3428*	2.0129	2.9880	3.1835	2.3622*	2.0331	3.0997	3.1346
	$\overline{X}\pm \sigma_X$	1.4522	1.7400	2.1684	2.5116	1.4766	1.7126	2.1625	2.483	1.4668	1.7378	2.1360	2.4869
						Panel B: $l=50$							
1	$\overline{X}\pm 0.5\sigma_X$	1.1319	1.7798	2.1109	2.6225	1.1787	1.8672	2.0323	3.084	1.1655	1.4856	2.0733	2.4134
	$\overline{X}\pm \sigma_X$	1.2272*	1.0900	1.2574	1.5941	1.0220*	0.9959	1.2657	1.725	0.9906	1.0808	1.2276	1.7426
2	$\overline{X}\pm 0.5\sigma_X$	0.9615*	0.9128	1.1571	1.3345	1.0150*	0.8352	1.0640	1.5213	1.0528**	0.674	1.0296	1.5991
	$\overline{X}\pm \sigma_X$	1.2571	1.3726	1.3863	1.7170	1.1491	1.2948	1.4236	1.839	1.1562	1.4211	1.5230	1.8761
4	$\overline{X}\pm 0.5\sigma_X$	1.5012*	1.1897	1.6471	2.0270	1.3255	1.4588	1.5650	2.0087	1.3545*	1.2548	1.5775	2.0860
	$\overline{X}\pm \sigma_X$	0.9901	0.9944	1.0296	1.2721	1.0180**	0.8692	0.9835	1.2422	0.6941	0.8057	0.8466	1.1248
12	$\overline{X}\pm 0.5\sigma_X$	2.4237*	2.0814	3.0895	3.1697	2.3428*	1.9723	3.0086	3.1404	2.3622*	2.0402	2.9668	3.0886
	$\overline{X}\pm \sigma_X$	1.4522	1.7225	2.2239	2.5189	1.4766	1.7186	2.1317	2.5227	1.4668	1.7319	2.2222	2.4098

[(*)] Notes: See notes to Table 1. Tabulated entries are test statistics and 5%, 10% and 20% level critical values. Test intervals are given in the second column of the table, for $\tau = 1, 2, 4, 12$. All tests are carried out using historical one-month Eurodollar deposit rate data for the period January 1971 - September 2005, measured at a weekly frequency. Single, double, and triple starred entries denote rejection at the 20%,10%, and 5% levels, respectively. Additionally, $\overline{X}$ and σ_X are the mean and standard deviation of the historical data. See Section 6.2 for complete details.