An Empirical Assessment of Spot Rate Model Stability

Lili Cai and Norman R. Swanson*

Rutgers University

April 2009

Abstract

The purpose of this paper is to add to the empirical evidence on the efficacy of alternative simulation models of the short term interest rate. This is done by constructing consistent specification tests that allow us to carry out a "horse-race" comparing various one, two, and three factor models (possibly with jumps), across multiple historical sample periods. We begin by outlining a three factor version of the simulation based specification test of Bhardwaj, Corradi and Swanson (BCS: 2008), which is based on a comparison of simulated and true conditional distributions and confidence intervals. Our evaluation involves comparing six affine models of the short rate during four historical periods, referred to as: "Post Bretton-Woods"; "Pre-1990s"; "The Stable 1990s"; and "Post 1990s". Based on the examination of Eurodollar rate data, we find that the CIR model, which is often rejected in the literature, performs best among the candidate models in "The Stable 1990s", while there is little to choose between one and two factor models when considering the "Post 1990s" period. Examination of "Pre 1990s" data, on the other hand, suggests there is little to choose between 2 and 3 factor models, and the one factor CIR model performs poorly. Moreover, under the "Post Bretton-Woods" period, the "best" performer is the three factor CHEN (1996) model examined by Andersen, Benzoni and Lund (2004). We conclude that the choice of model for simulating the future distribution of short rates is highly sample dependent.

Keywords: interest rate, multi-factor diffusion process, block bootstrap, GMM, jump process, simulation, specification test, volatility
JEL Classification: C1, C5.

——————————————————————————

* Lili Cai, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, lcai@econ.rutgers.edu. Norman R. Swanson, Department of Economics,Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. This paper is the first chapter of the first author's Ph.D. dissertation, and she is extremely grateful to her committee members, Elena Goldman, Roger Klein, and John Landon Lane for helpful comments. The authors also own many thanks to seminar participants at Rutgers University for numerous useful suggestions. Swanson gratefully acknowledges financial support from a Rutgers University Research Council grant.

1 Introduction

Diffusion processes are used in virtually all aspects of continuous time finance from yield curve to exchange rate modeling, for the purposes of prediction, simulation and pricing. This has led to many papers recently being published in the field, numerous of which are a part of ongoing "efforts" to specify models that adequately capture financial variables dynamics. In this paper we undertake a specification search of alternative short rate models. Our search centers upon the evaluation, via state of the art consistent specification tests, of simulated ex ante distributions constructed using a variety of multi factor models both with and without jumps.

One characteristic of continuous time models that is crucial to the application of such models is that only a few of those currently in use by practitioners have closed form solutions (see e.g. Vasicek model (1977), Cox, Ingersoll, and Ross (CIR: 1985), Black and Scholes (1973), and Hull and White (1990)). Many do not have closed form solutions, particularly those involving one or multiple latent variables (see e.g. the stochastic mean model of Balduzzi et al. (1998), the stochastic volatility model of Heston (1993), the three-factor model of Chen (1996), and the three-factor model with jumps discussed in the noteworthy paper by Andersen, Benzoni and Lund (ABL: 2004)). This issue has implications not only for pricing formulae derived from the models, but also for estimation. In recent years, many new methods have been developed for the estimation of continuous time models and the (often unknown in closed form) conditional densities associated with them. For example, Ait-Sahalia (1999, 2002, 2008) provides closed form approximations of (unknown) conditional densities using Hermite polynomials, for one-factor, stochastic volatility, and multi-factor models, respectively. These approximations have led to the development of numerous consistent specification tests for evaluating individual models. Another approach to the estimation problem is to use estimators of distribution functions based on historical data as well as based on simulated data, such as that discussed in Bhardwaj, Corradi and Swanson (2008) (see also Duffie and Singleton (1993)). In this paper we carry out a "horse-race" where we "select" the best model, from amongst a relatively large class of alternatives. The test statistics that we use in the "horse-race" are closely related to the discrete time, point mean square forecast error, model selection test statistics of White (2000) which are widely used in empirical finance (see e.g. Sullivan, Timmerman and White (1999, 2001)).

Indeed, the question of whether particular parametric representations of diffusion processes have acceptable explanatory power relative to the "true" underlying dynamics has led to the development of many specification tests. Such specification tests for continuous-models can be divided into multiple categories. One category focuses on nonparametric tests (see e.g. Ait-Sahalia(1996, 2002), Ait-Sahalia, Fan and Peng (2006), and Hong and Li (2005)), where tests are characterized by the nonparametric estimation (e.g. using kernels) of transition densities; and model implied transition densities are often compared with their nonparametric counterparts. Another category involves the examination of generalized cross spectra (see e.g. Hong and Li (2005)). A third category that includes papers by Andersen, Benzoni and Lund (ABL: 2004), Thompson (2004), BCS (2008), Chen and Hong (2005), and Corradi and Swanson (2008), to name but a few, uses parametric methods to examine the fit of models. The testing approach used in this paper falls within this category, and is based upon the examination of models via comparison of distributional analogs of mean square forecast errors that measure the difference between simulated conditional confidence intervals (or distributions) generated using fitted parametric models and empirical conditional confidence intervals (or distributions) estimated using historical data. In addition to directly implementing the BCS (2008) test to achieve this end, we also provide a straightforward extension of the BCS (2008) test that allows for the comparison of three factor models (as opposed to evaluating only one and two factor models as in BCS (2008)).

One feature of the current literature is that the application of different specification tests and different data sets has led to a variety of different conclusions. For example, Ait-Sahalia (1996) test fails to reject the Chan, Karolyi, Longstaff and Sanders (CKLS:1992) model and the nonlinear drift model (Ait-Sahalia, 1996). On the other hand, Hong and Li (2005) strongly reject all univariate affine models of the Euro dollar rate, and suggest that even very sophisticated models (including GARCH, regime switching, and jumps) do not adequately capture interest rate dynamics. BCS (2008) reject the CIR (1985) model, and conclude that stochastic volatility models are superior to the CIR model. To some extent, one might argue that the mixed evidence in the extant literature can be attributed to the fact that numerous analyses have been carried out using (relatively) small numbers of models and varying data samples, many of which can be tied to particular historical "episodes". In light of this, we examine multiple time periods and a relatively rich set of models in this paper. This is an important element of our analysis not only because of the reasons discussed above, but also because structural breaks in the true underlying dynamics might be prevalent of sample periods are not carefully defined, and such breaks might be expected to have important effects on tests used to assess alternative term structure models.

Our empirical evaluation involves comparing six affine models of the short rate during four historical periods, referred to as: "Post Bretton-Woods", "Pre-1990s", "The Stable 1990s", and "Post 1990s". Our findings are based on an evaluation of the one month Eurodollar deposit rate collected at a weekly frequency, and can be summarized as follows. The CIR model, which is often rejected in the literature, performs best among the candidate models in "The Stable 1990s", while there is little to choose between one and two factor models when considering the "Post 1990s" period. Examination of "Pre 1990s" data, on the other hand, suggests there is little to choose between 2 and 3 factor models, and the one factor CIR model performs poorly. Interestingly, when our entire sample of weekly data from 1970 to 2008 is used to evaluate the competing models, the "best" performer is the three factor CHEN (1996) model examined by ABL (2004). We conclude that the choice of model for simulating the future distribution of short rates is highly sample dependent; and when long samples of data are used which likely span different economic regimes, more complicated models are likely to be selected as "best" than is actually warranted, particularly if the current regime is expected to continue throughout the length of the simulation period of interest. Put differently, structural breaks appear to be an important component of simple diffusion models of the short rate (same as that in the concluding part).

The rest of this paper is organized as follows. In Section 2, we review the specification test. Section 3 presents the short term interest rate models considered in this paper. Section 4 provides the data. Empirical results are given in section 5. Section 6 summarizes the results and concludes.

2 Consistent Specification Tests

The specification test we adapt is the BSC (2008) test, which is a Kolmogorov type test utilizing a simple simulation based approach to construct conditional distributions when the functional form of the conditional density is unknown. The distributions are in turn used to form predictive confidence intervals for time period $t + \tau$, given information up to period t. Thereafter, for a given diffusion model the unknown conditional distribution is replaced by the simulated counter part. The specification test measures the difference between simulated conditional confidence distributions generated using fitted parametric models and empirical conditional distributions implied by historical data. The approach of BCS builds upon specification testing work discussed in Corradi and Swanson (2005), Ait-Sahalia (1996, 2006) and Hong and Li (2005, 2007).

2.1 The BCS (2008) test

The BCS test is based on the comparison of the empirical cumulative distribution function (CDF) and the cumulative distribution function implied by the specification of the drift and the variance under a given null model. To illustrate the idea, consider a parametric diffusion process:

$$dX_t = b(X_t, \theta^{\dagger})dt + \sigma(X_t, \theta^{\dagger})dW_t, \tag{1}$$

where W_t is a Brownian motion, and the true parameter vector is $\theta_0 = (b_0', \sigma_0')' \in \Theta$, Θ is a compact subset of $\Re^K$. Under correct specification of the diffusion process, we have that $b(\cdot,\cdot) = b_0(\cdot,\cdot)$ and $\sigma(\cdot,\cdot) = \sigma_0(\cdot,\cdot)$, that is $\theta^{\dagger} = \theta_0$. Note that the stationary density, $f(x, \theta^{\dagger})$, and its associated invariant probability measure are uniquely determined by $b(\cdot)$ and $\sigma^2(\cdot)$ (the drift and variance terms in the model) (BCS(2008)). The alternative hypothesis is that the parameters in the above diffusion process do not coincide with the true parameters. Instead of comparing parameters directly (see Ait-Sahalia et al. (2006)), we compare the cumulative distribution function. The null and alternative hypotheses 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.

To construct a specification test, we follow BCS (2008) by defining the $\tau-step$ ahead conditional distribution of $X_{t+\tau}^{\theta^{\dagger}}$, given $X_t^{\theta^{\dagger}} = X_t$, as:

$$F_{\tau}(u|X_t, \theta^{\dagger}) = \Pr\left(X_{t+\tau}^{\theta^{\dagger}} \leq u | X_t^{\theta^{\dagger}} = X_t\right), \tag{2}$$

where $t = 1, 2, 3, ..., T-\tau$. Instead of comparing $F_{\tau}(u|X_t, \theta^{\dagger})$ and $F_{0,\tau}(u|X_t, \theta_0)$, we need to replace $F_{\tau}(u|X_t, \theta^{\dagger})$ with its simulated counterpart. Namely:

$$\widehat{F}_{\tau}(u|X_t, \widehat{\theta}_{T,N,h}) = \frac{1}{S}\sum_{s=1}^{S} 1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\} \tag{3}$$

where $\widehat{\theta}_{T,N,h}$ is estimated by using the whole sample of T observations. Here, $\widehat{\theta}_{T,N,h}$ converges to $\theta^{\dagger}$, and S is the number of simulation paths. BCS (2008) show that $\frac{1}{S}\sum_{s=1}^{S} 1\left\{X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u\right\}$ is a consistent estimate of $F_{\tau}(u|X_t, \theta^{\dagger})$.

In order to test the above hypotheses, we measure the departure from the null hypothesis by defining the test statistic $Z_T = \sup_{u \times v \in U \times V} |Z_T(u,v)|$, where

$$Z_T(u,v) = \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{S} \sum_{s=1}^{S} 1 \left\{ X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u \right\} - 1\{X_{t+\tau} \leq u\} \right) 1 \{X_t \leq v\}, \tag{4}$$

and U and V are compact sets on the real line. BCS outline block-bootstrap methods (see details in CS (2005, 2007, 2008)) for constructing critical values for this test. Specially, the bootstrap statistic is $Z_T^* = \sup_{u \times v \in U \times V} |Z_T^*(u,v)|$, where

$$\begin{aligned} Z_T^*(u,v) \quad &= \quad \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{S} \sum_{s=1}^{S} 1 \left\{ X_{s,t+\tau}^{\widehat{\theta}^*_{T,N,h}} \leq u \right\} - 1\{X^*_{t+\tau} \leq u\} \right) 1 \{X^*_t \leq v\} \\ &\quad - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{S} \sum_{s=1}^{S} 1 \left\{ X_{s,t+\tau}^{\widehat{\theta}_{T,N,h}} \leq u \right\} - 1\{X_{t+\tau} \leq u\} \right) 1 \{X_t \leq v\} \end{aligned} \tag{5}$$

and where X_t^* is a resampled series constructed using standard block-bootstrap methods, $\widehat{\theta}^*_{T,N,h}$ is estimated parameter using the resampled data, X_t^*, and is in turn used to construct simulated paths, $X_{s,t+\tau}^{\widehat{\theta}^*_{T,N,h}}$, $s = 1, ..., S$ and $t = 1, ..., T-\tau$. In order to generate the empirical distribution of Z_T^*, one performs B bootstrap replications (B large). Then, one compares Z_T with the percentiles of the empirical distribution of Z_T^*, and rejects H_0 if Z_T is greater than the $(1-\alpha)th$-percentile. Otherwise, one fails to reject. BCS (2008) has proved that the test carried out in this manner is correctly asymptotically sized, and has unit asymptotic power.

2.2 Multi factor versions of the BCS test

For two-factor models (e.g. stochastic mean and stochastic volatility models, where $X_t = \left(X_t^1, X_t^2\right)'$), the difficulty lies in dealing with the initial value for the simulation process, given that the latent variable in X_t is unobservable. BCS (2008) integrate out this effect by first simulating a long path of length N observations for latent variable X_t^2. Second, they take the simulated values in the first step as starting values for the latent variable and simulate $S \times N$ paths in order to form, $\widehat{F}_\tau(u|X_t, \widehat{\theta}_{T,N,h})$. The associated test statistic in Eq(4) is:

$$Z_T(u,v) = \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\frac{1}{NS} \sum_{j=1}^{N} \sum_{s=1}^{S} 1 \left\{ X_{j,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}} \leq u \right\} - 1\{X^1_{t+\tau} \leq u\} \right) 1 \{X_t^1 \leq v\}. \tag{6}$$

Similarly, the bootstrap statistic analogous to that given in Eq(5) is $Z_T^* = \sup_{u\times v\in U\times V} |V_T^*(u,v)|$, where

$$Z_T^*(u,v) = \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S}1\left\{X_{j,s,t+\tau}^{1,\widehat{\theta}^*_{T,N,h}} \leq u\right\} - 1\{X_{t+\tau}^{1*} \leq u\}\right)1\left\{X_t^{1*} \leq v\right\}$$
$$-\frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{s=1}^{S}1\left\{X_{j,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}} \leq u\right\} - 1\{X_{t+\tau}^{1} \leq u\}\right)1\left\{X_t^{1} \leq v\right\}. \quad (7)$$

Of note that we use $X_{t+\tau}^1$ and $X_{j,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}}$ to construct the conditional interval because only X_t^1 is observable in X_t.

Now, consider a three-factor model (see e.g. the "CHEN" and "CHENJ" models discussed below), where $X_t = \left(X_t^1, X_t^2, X_t^3\right)'$, and $W_t = \left(W_t^1, W_t^2, W_t^3\right)$ are mutually independent standard Brownian motions in Eq(1). The key issue concerns how to construct the conditional distribution $F_\tau(u|X_t,\theta^\dagger) = \Pr\left(X_{t+\tau}^{\theta^\dagger} \leq u | X_t^{\theta^\dagger} = X_t\right)$ without knowing the starting value of X_t^2 and X_t^3. To deal with this issue, we propose a simple simulation based method that is an immediate consequence of the approach discussed in BCS for one and two factor models:

Step 1: Given the estimated parameter $\widehat{\theta}_{T,N,h}$, *generate a path of length N (a large number) for* $X_t^{\widehat{\theta}_{T,N,h}}$*.The trick is to use the mean of stochastic volatility and the mean of stochastic mean in* $\widehat{\theta}_{T,N,h}$ *as the initial start values for these two latent variables. Retrieve* $X_t^{2,\widehat{\theta}_{T,N,h}}$ *and* $X_t^{3,\widehat{\theta}_{T,N,h}}$, $t = 1,2,...,N$ *from the path.*

Step 2: Given the observable X_t^1 *and the* $N \times N$ *simulated latent paths (*$X_j^{2,\widehat{\theta}_{T,N,h}} \times X_m^{3,\widehat{\theta}_{T,N,h}}$ $j,m = 1,...,N$*) as the start values , we simulate* τ*-step ahead* $X_{t+\tau}^{1,\widehat{\theta}_{T,N,h}}$*. Since the start values for the two latent variables are* $N \times N$ *length, so for each* X_t^1 *we have* N^2 *path, that is* $F_{\tau,i}(u|X_t,\widehat{\theta}) = \frac{1}{N^2}\sum_{m=1}^{N}\sum_{j=1}^{N}1\left\{X_{j,m,t+\tau}^{1,\widehat{\theta}_{T,N,h}} \leq u\right\}$, *where* i *denotes the ith simulation.*

Step 3: Simulate $X_{t+\tau}^{1,\widehat{\theta}_{T,N,h}}$ *S times, that is to repeat step 2 S times.* $\frac{1}{S}\sum_{i=1}^{S}F_{\tau,i}(u|X_t,\widehat{\theta}_{T,N,h})$ *is the estimator of* $F_\tau(u|X_t,\theta^\dagger)$.

Step 4: Construct the statistic for the null of correct specification of the conditional distribution:

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

where

$$Z_T(u,v) = \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{N^2S}\sum_{m=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{S}1\left\{X_{j,m,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}} \leq u\right\} - 1\{X_{t+\tau}^1 \leq u\}\right)1\left\{X_t^1 \leq v\right\}. \quad (8)$$

All of the results outlined in BCS (2008) generalize immediately to the current setting. In particular, the following results hold.

Proposition 1 *(follows immediately from Proposition 2 in BCS (2008)): Assume that* $T, N, S \to \infty$*. Then, if* $h \to 0$, $T/N \to 0$, $T/S \to 0$, $h^2 T \to 0$, *and the model is correctly specified, the following result holds for any* X_t^1, $t \geq 1$:

$$\frac{1}{N^2 S}\sum_{m=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,m,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}} \leq u\right\} - F_0(u|X_t,\theta_0) \xrightarrow{p} 0, \text{uniformly in } u.$$

Similarly, we can implement the same bootstrap method as that outlined in BCS (2008) in order to form the resampled series, X_t^{1*} and construct bootstrap statistics. Namely:

Step 1: Resample X_t^1*. In particular, we draw* b *blocks (with replacement) of length* l*, where* $bl = T$*. Thus, each block is equal to* $X_{i+1}^1, ..., X_{i+l}^1$*, for some* $i = 0, ..., T-l+1$*, with probability* $1/(T-l+1)$*. More formally, let* I_k*,* $k = 1, ..., b$ *be iid discrete uniform random variables on* $[0, 1, ..., T-l+1]$*. Then, the resampled series,* X_t^{1*} *is such that* $\{X_1^{1*}, X_2^{1*}, ..., X_l^{1*}, X_{l+1}^{1*}, ..., X_T^{1*}\} = \{X_{I_1+1}^1, X_{I_1+2}^1, ..., X_{I_1+l}^1, X_{I_2}^1, ..., X_{I_b+l}^1\}$*, 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* $\widehat{\theta}_{T,N,h}^*$.

Step 2: Repeat Steps 1-3 in constructing $Z_T(u,v)$, but we use X_t^{1*} and $\widehat{\theta}_{T,N,h}^*$ to replace X_t^1 and $\widehat{\theta}_{T,N,h}$ respectively, to construct the conditional distribution for $X_{t+\tau}^{1*}$. *Particularly,* $X_{j,s,m,t+\tau}^{1,\widehat{\theta}_{T,N,h}^*}$ *is the simulated value at simulation* s*, constructed using* $\widehat{\theta}_{T,N,h}^*$*, and* $X_t^{1,\widehat{\theta}_{T,N,h}^*}$*,* $X_{j,h}^{2,\widehat{\theta}_{T,N,h}^*}$*,* $X_{m,h}^{3,,\widehat{\theta}_{T,N,h}^*}$ *as initial value.* Of note that we use the same set of random errors used in $X_{j,m,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}}$to construct $X_{j,m,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}^*}$.

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

$$Z_T^* = \sup_{u \times v \in U \times V} |Z_T^*(u,v)|, \tag{9}$$

where

$$\begin{aligned} Z_T^*(u,v) &= \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{N^2 S}\sum_{m=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,m,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}^*} \leq u\right\} - 1\{X_{t+\tau}^{1*} \leq u\}\right) 1\left\{X_t^{1*} \leq v\right\} \\ &\quad - \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}\left(\frac{1}{N^2 S}\sum_{m=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{S} 1\left\{X_{j,m,s,t+\tau}^{1,\widehat{\theta}_{T,N,h}} \leq u\right\} - 1\{X_{t+\tau}^{1} \leq u\}\right) 1\left\{X_t^{1} \leq v\right\}, \end{aligned}$$

Step 4: Repeat step 1-3 B times to generate the empirical distribution of the B bootstrap statistics.

The first order asymptotic validity of inference carried out using bootstrap statistics formed as outlined above follows immediately from BCS (2008).

3 The Models

We take the CIR model as our univariate benchmark model, although much of the focus of this paper is on multifactor models with one or two latent variables. Our multifactor models include a stochastic mean model, stochastic volatility model, and three-factor model. Although modelling financial variables using continuous time diffusion processes simplifies the analytical work, empirical observation suggests that there are often violent movements in underlying measures of these variables, and hence models with jumps have become important. For example, such models often better explain excess kurtosis and skewness. Thus, we also consider models with jumps, as suggested in the key "continuous-time model horse-race" paper by ABL (2004). However, our models differ somewhat from ABL (2004). For example, we consider the jump processes that are driven by two separate Poisson draws (jump up and jump down) with different intensities, and different jump magnitudes, as documented in Chacko and Das (2002). In total, we consider six short rate affine models, all of which are outlined briefly below.

The Cox-Ingersoll-Ross (CIR) Model: We follow CIR (1985) and posit that:

$$dr(t) = \kappa_r (\theta - r(t))\, dt + \sigma_r \sqrt{r(t)} dW_r(t), \tag{10}$$

where we assume that $W_r(t)$ is Brownian motion, θ is the long-run mean of interest rate, κ_r is the mean-reversion speed, and σ_r is the standard deviation of the interest rate, which we assume it is a constant. Different from Vasicek model (1977), the process $r(t)$ is a square-root diffusion process and it can not take negative values ($2\kappa_r\theta > \sigma_r^2$).

Stochastic Mean Model (SM): We generalize the above model by allowing for a time varying mean, $\theta(t)$, where $\theta(t)$ is itself is a mean reverting process that converges to its unconditional mean:

$$\begin{aligned} dr(t) &= \kappa_r (\theta(t) - r(t))\, dt + \sigma_r dW_r(t), \\ d\theta(t) &= \kappa_\theta \left(\overline{\theta} - \theta(t)\right) dt + \sigma_\theta \sqrt{\theta(t)} dW_\theta(t), \end{aligned} \tag{11}$$

where we assume that $W_r(t)$ and $W_\theta(t)$ are independent Brownian motions. $\overline{\theta}$ and σ_θ are the mean and standard deviation of the stochastic mean process $\theta(t)$, respectively. Of note is that the mean process $\theta(t)$ can not take negative values provided that $2\kappa_\theta\overline{\theta} > \sigma_\theta^2$. Stationarity requires that κ_r and κ_θ (which controls mean reversion speed) are greater than zero.

Stochastic Volatility Model (SV): We consider the prototypical stochastic volatility (SV) model proposed by Heston (1993) that has been extensively examined in the literature (see e.g. Chen (1996), Andersen, and Lund (1997), ABL (2004), and Ait-Sahalia and Kimmel (2007)). Namely:

$$\begin{aligned} dr(t) &= \kappa_r(\overline{r} - r(t))\,dt + \sqrt{V(t)}dW_r(t), \\ dV(t) &= \kappa_v(\overline{v} - V(t))\,dt + \sigma_v\sqrt{V(t)}dW_v(t), \end{aligned} \tag{12}$$

where κ_r and κ_v are the mean-reversion speeds of the interest rate and the variance thereof, respectively; which are required to be greater than zero to avoid nonstationarity. Additionally, $\overline{v}$ is the mean of the variance, and σ_v is the volatility of variance. $W_r(t)$ and $W_v(t)$ are scalar Brownian motions in some probability measure. We assume that $W_r(t)$ and $W_v(t)$ are not independent, but correlated. In particular, $dW_r(t)\; dW_v(t) = \rho dt$, where the correlation ρ is some constant in [-1,1]. Finally, Volatility is a square-root diffusion process, which implies that $2\kappa_v\overline{v} > \sigma_v^2$.

Stochastic Volatility Model with Jumps (SVJ): We relax the SV by allowing for discontinuous dynamics. In particular, we add Poisson-exponential jumps to the model, as follows:

$$\begin{aligned} dr(t) &= \kappa_r(\overline{r} - r(t))\,dt + \sqrt{V(t)}dW_r(t) + J_u dq_u - J_d dq_d, \\ dV(t) &= \kappa_v(\overline{v} - V(t))\,dt + \sigma_v\sqrt{V(t)}dW_v(t), \end{aligned} \tag{13}$$

where q_u and q_d are Poisson processes with jump intensity λ_u and λ_d respectively, and are independent of the Brownian motions $W_r(t)$ and $W_v(t)$. In particular, λ_u is the probability of a jump up, $\Pr(dq_u(t) = 1) = \lambda_u$ and λ_d is the probability of a jump down, $\Pr(dq_d(t) = 1) = \lambda_d$. J_u and J_d are jump up and jump down sizes and have exponential distributions: $f(J_u) = \frac{1}{\zeta_u}\exp\left(-\frac{J_u}{\zeta_u}\right)$ and $f(J_d) = \frac{1}{\zeta_d}\exp\left(-\frac{J_d}{\zeta_d}\right)$, where ζ_u, $\zeta_d > 0$ are the jump magnitudes, which are the means of the jumps, J_u and J_d.

Three Factor Model (CHEN): We combine various features of the above models, by considering a version of the oft examined 3-factor model due to Chan, Karolyi, Longstaff and Sanders (1992), which is discussed in detail in Dai and Singleton (2000). In particular, we consider the Chen (1996) 3-factor model:

$$\begin{aligned} dr(t) &= \kappa_r(\theta(t) - r(t))\,dt + \sqrt{V(t)}dW_r(t), \\ dV(t) &= \kappa_v(\overline{v} - V(t))\,dt + \sigma_v\sqrt{V(t)}dW_v(t), \\ d\theta(t) &= \kappa_\theta\left(\overline{\theta} - \theta(t)\right)dt + \sigma_\theta\sqrt{\theta(t)}dW_\theta(t), \end{aligned} \tag{14}$$

where $W_r(t)$, $W_v(t)$ and $W_\theta(t)$ are independent Brownian motions, and V and θ are the stochastic volatility and stochastic mean of short rate r, respectively. As discussed above, non-negativity for $V(t)$ and $\theta(t)$ requires that $2\kappa_v\overline{v} > \sigma_v^2$ and $2\kappa_\theta\overline{\theta} > \sigma_\theta^2$.

Three Factor Jump Diffusion Model (CHENJ): ABL (2004) extend the three factor Chen (1996) model by incorporating jumps in the short rate process, hence improving the ability of the model to capture the effect of outliers, and to address the finding by Piazzesi (2004, 2005) that violent discontinuous movements in underlying measures may arise from monetary policy regime changes. The model is defined as follows:

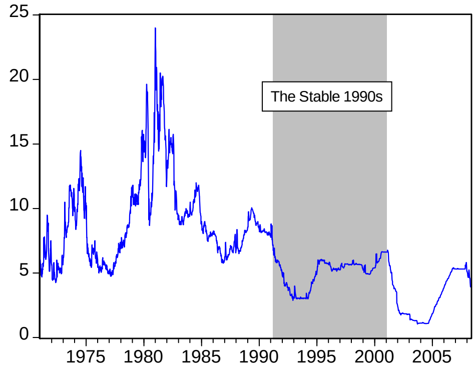
$$
\begin{aligned}
dr(t) &= \kappa_r(\theta(t) - r(t))\,dt + \sqrt{V(t)}dW_r(t) + J_u dq_u - J_d dq_d, \\
dV(t) &= \kappa_v(\overline{v} - V(t))\,dt + \sigma_v\sqrt{V(t)}dW_v(t), \\
d\theta(t) &= \kappa_\theta\left(\overline{\theta} - \theta(t)\right)dt + \sigma_\theta\sqrt{\theta(t)}dW_\theta(t),
\end{aligned}
\tag{15}
$$

where q_u and q_d are Poisson processes with jump intensities λ_u and λ_d, respectively; and are independent of the Brownian motions $W_r(t)$, $W_v(t)$ and $W_\theta(t)$. In particular, λ_u is the probability of a jump up, $\Pr(dq_u(t) = 1) = \lambda_u$ and λ_d is the probability of a jump down, $\Pr(dq_d(t) = 1) = \lambda_d$. J_u and J_d are jump up and jump down sizes and have exponential distributions $f(J_u) = \frac{1}{\zeta_u}\exp\left(-\frac{J_u}{\zeta_u}\right)$ and $f(J_d) = \frac{1}{\zeta_d}\exp\left(-\frac{J_d}{\zeta_d}\right)$, where ζ_u, $\zeta_d > 0$ are the jump magnitudes, which are the means of the jumps J_u and J_d.

4 Data

In order to facilitate the comparison of our empirical findings with those of BCS (2008), we use the same data set as they do. Namely, we use data collected on the one-month Eurodollar deposit rate as our proxy of the short rate. Our data ranges from January 1971 to April 2008 (1,996 weekly observations). Other yields that are often considered in the literature include the monthly federal funds rate (Ait-Sahalia (1999)), monthly yields on zero-coupon bonds with different maturities (see Duffee (2002) and Diebold and Li (2006, 2007)), the weekly 3-month T-bill rate (see Anderson, Benzoni and Lund (2004) and Durham (2001)), and even yields with longer maturities like the 6-month LIBOR (see Piazzesi (2001)).

Figure 1: Eurodollar rate 01/08/1971 to 04/25/2008



Notes: This figure plots empirical weekly data on the Eurodollar rate for the period 01/08/1971 to 04/24/2008. The shadowed is the "Stable 1990s" period, which covers period 03/1991- 01/2001.

To perform our empirical analysis, we divide the sample into 3 sub-samples corresponding roughly to three historical episodes (see Figure 1). One is the "Pre 1990s" period, from January 1971 to February 1991. A second period is the "Stable 1990s" period, from March 1991 to January 2001. A final period is the "Post 1990s" from February 2001 to April 2008. Of note is that the "Stable 1990s" is the longest economic expansion that the United States has ever experienced. In early 1990s, the Federal Reserve returned to targeting the federal fund rate. The Federal Reserve fixed the federal funds rate at 3% from late 1992 until February 1994, and increased the rate to 6% by early 1995. Only in January 2001, did the Federal Reserve begin to cut rates again. Interest rate during this expansion period were quite stable. During our "Pre 1990s" sample period, the Federal Reserve targeted, at various times, monetary aggregates and interest rates, which resulted to some extent in elevated interest rate volatility relative to that experienced during the 1990s. Moreover, high inflation and recessions in 1980s also increased volatility over this period. The "Post 1990s" period has also been more volatile, as federally manipulated interest rates were first lowered in order to stimulate the economy after the high-tech bubble crash, were later increased to accommodate increasing inflation and concerns about the booming housing market, and have recently again been lowered due to global financial concerns surrounding the recent crash of the U.S. housing market and related problems associated with related derivative mortgage products.

5 Empirical Results

We first discuss broad characteristics of our data via examination of various summary statistics; and also discuss our parameter estimation results. Thereafter, we disucc the results of our "horse-race" carried out using four different sample periods: January 1971- April 2008 ("Post Bretton-Woods"), January 1971- February 1991 ("Pre-1990s"), March 1991- January 2001 ("The Stable 1990s"), and February 2001- April 2008 ("Post-1990s").

5.1 Summary statistics

Table 1: Summery statistics for different subsamples

Sample period	Date	Mean	Median	Std. Dev.	Skewness	Kurtosis	JB
Post Bretton-Woods	01/1971-04/2008	0.0687	0.0594	0.0365	1.1939	5.1710	844.9710
Pre 1990s	01/1971-02/1991	0.0908	0.0831	0.0343	1.3553	4.9309	483.6905
The Stable 1990s	03/1991-01/2001	0.0508	0.0541	0.0109	-0.6995	2.4096	49.8637
Post 1990s	02/2001-04/2008	0.0317	0.0286	0.0168	0.1909	1.4539	39.5192

Notes: This table reports summary statistics for our historical data. The"Post Bretton Wood" stands for the period from January 1971 to April 2008. The "Stable 1990" period is from March 1991 to January 2001. The "Pre 1990s" period covers January 1971 to February 1991. The "Post 1990s" period is from February 2001 to April 2008. See section 4 for details.

Table 1 reports various summary statistics for the data, including mean, median, variance, skewness, kurtosis and Jarque-Bera test statistic. The "Post Bretton-Woods" data has a mean of 0.0687 (i.e. 6.87%), standard deviation of 0.0365 (3.65%), positive skewness of 1.1939, and a distribution that is peaked (leptokurtic) relative to a normal distribution, with a kurtosis value of 5.1710. The Jarque-Bera test indicates that the sample does not follow a normal distribution. The three sub-samples are also not normally distributed; and they have quite different distributional properties. The "Pre-1990s" period is quite volatile, and has the highest mean (9.08%), and the highest standard deviation (3.43%). This is due to the high inflation/interest rate regime in the 1980s (the maximum spot interest rate during the period was 24%). In contrast, the "Stable 1990s" period is much more stable, with a mean of 5.08%, standard deviation of 1.09%, and smaller skewness and kurtosis. Of note is that the Federal Reserve Bank fixed the federal funds rate at 3% from late 1992 until 02/1994, and then increased the rate to 6% by early 1995, keeping it at this level until January 2001. This may explain why the"Stable 1990s" period is negatively skewed (-0.699), although all other samples have significant positive skewness. As opposed to the other two sub-samples, the "Post-1990s" period demonstrates a bimodal distribution (see discussion in Section 4 for details). Of note is that in 2004 the Eurodollar rate reached a low of 1.04%. Moreover, it is not surprising that the Eurodollar rate data that we examine shares the same patterns of increase and decrease as the federal funds rate, which explains the sharp decreases and increases in the Eurodollar rate in the "Post-1990s" period. Compared with other samples, the "Post-1990s" period has the lowest mean (3.17%), but has a relatively high standard deviation (1.68%). These results suggest that interest rate models that are "regime-dependent" may provide a better fit to data generated during a particular regime. Of course, a difficulty with modelling individual regimes will be ascertaining whether the period to be simulated can be expected to remain within the regime.

5.2 Estimation results

One of the many available estimation method for models of the variety considered here is efficient method of moments (EMM) as proposed by Gallant, and Tauchen (1996, 1997), which calculates moment functions by simulating the expected value of the score implied by an auxiliary model. Parameters are then computed by minimizing a chi-square criterion function. An alternative estimation procedure based on method of moments is generalized method of moments (GMM), as documented in Jiang and Knight (2002) (see also Chacko and Viceira (2003)), which can be used to easily estimate relatively complicated processes including all of the models considered in this paper, but which requires the existence of a closed form expression for the empirical conditional characteristic function. Since the conditional characteristic functions and associated moments are of the underlying continuous time process, and are not discrete approximation, estimation is free from discretization error associated with simulation based estimators. In our empirical analysis, we have found that these GMM estimates appear to be stable and robust.

Estimation results for the different sample periods are summarized in Table 2, where parameter estimates are annualized (this is done for weekly data by setting $\tau = \frac{1}{52}$), to accord with the extant literature. Panel A of table 2 reports the estimation results for the whole sample. The CIR model has an estimate of $\overline{r} = 6.57\%$, which is very close to the observed sample mean of 6.87%. Of note is that the CIR model has a very high standard deviation of 11.8%, which is much higher than the observed sample standard deviation of 3.65%. This discrepancy is likely driven by the model itself, which assumes that volatility is constant, and is one reason why the CIR model is often rejected in literature.[1]

The SM model extends the CIR model by allowing for a time varying drift. The long term mean, $\overline{\theta}$, is estimated to be 5.58%. Further, our estimates of the speed of mean reversion (i.e. κ_r and κ_θ) are 0.2169 and 0.2926, respectively. Compared with the estimation results reported by ABL (2004), the short rate $r(t)$ reverts slowly to its time-varying mean $\theta(t)$, and $\theta(t)$ reverts a little bit faster

[1] see e.g. Ait-Sahalia (1996), Tauchen (1997), Bandi (2002), and BCS (2008).

to its long-run mean, $\overline{\theta}$. Of note is that some of the SM parameters are not highly significant (e.g. the mean reversion parameter), perhaps accounting to some extent of the SM model's inability to fully capture the fat tailed behavior of our data.

The SV model extends the CIR by allowing for time varying volatility. As shown in Panel A of Table 2, the unconditional mean is $\overline{r} = 5.55\%$, and the mean reversion speed is 0.210. The half life of shocks in this model is three and a half years (i.e. the first order autoregressive coefficient is $e^{-\frac{\kappa_r}{52}} = 0.996$, at the weekly level), which corresponds to the strong serial dependence observed in the data. Furthermore, the first order autoregressive coefficient for the volatility process in the SV model is $e^{-\frac{\kappa_v}{52}} = 0.951$, which is again indicative of strong serial dependence. In this model, the estimate of ρ is -0.192, suggesting small negative conditional correlation between the short rate and its stochastic volatility.

Our SVJ model estimation results are as expected. In this model, the estimates of μ and κ_r are very close to the corresponding values in the SV model. Examination of the jump coefficients suggests that there is a clear asymmetry in the magnitude and intensity of up and down jumps. Our estimates indicate that the average size of an up-jump is small, at only 7 basis points; but with a high intensity of 5.49. As parameters are annualized, this means that there are roughly 5.5 up-jumps of average size 7 basis points, in a year. On the contrary, we estimate roughly 3.2 down-jumps per year, but with a much larger average size of 11.4 basis points.[2] Thus, the total number of jumps (up and down) to the interest rate process is estimated to be roughly 8-9. This result is in line with the observation of Piazzesi (2005) that jumps in the short rate are linked closely to the Federal Open Market Committee (FOMC) meetings; and the FOMC usually meets 8 times per year.[3]

Turning now to our three factor models, the findings in Panel A of Table 2 indicate that the estimated value of κ_θ increases significantly, relative to the value obtained under the SM model. In other words, $\theta(t)$ reverts to its long-run mean, $\overline{\theta}$, at a much faster rate than the SM model. The other estimated parameters in the CHEN and CHENJ models are as expected. The most notable caveat is for CHENJ, where the jump up and jump down sizes both decrease when compared with the estimates from the SVJ model, which suggests that the SVJ model accommodates time-varying mean variation in the jump specification. This in turn suggests that model misspecification imparted by not including enough time varying components in short rate models may lead to substantially

[2] As starting values in our optimization, we use a jump intensity of 3 up and 3 down jumps per year (*i.e.* $\lambda_u, \lambda_d = 3$). Further, the jump sizes are initially set to 25 basis points, following Piazzesi (2005).

[3] There are some emergency meetings beyond the 8-meeting agenda if needed.

biased estimates of the remaining parameters in the model. It appears that increasing model complexity substantively changes other parameter estimates.

Panel B of Table 2 summarizes estimation results for the "Pre-1990s" sample period. Of note is that all models that we discussed above do capture the relatively higher interest rate mean and volatility in the "Pre-1990s", as expected. The CIR model has $\overline{r} = 9.24\%$, and a much higher mean reversion speed κ_r. Similarly, the long term mean, $\overline{\theta}$, is estimated to be 8.32% in the SM model. Furthermore, our estimates of the speed of mean reversion (i.e. κ_r κ_θ and κ_v) are higher in Panel B. Interestingly, the estimates in Panel B are not particularly supportive of our stochastic volatility and three-factor models, in the sense that the parameters are poorly identified with standard errors that increase relatively to Panel A. However, we shall see that these models do fare well against the competitors when comparing simulated distributions using our BCS type tests. Of final note is that we have more jumps up and jumps down in this period, and jump sizes are higher than those reported in Panel A. These results are sensible, and reflect the high volatility of interest rate during the "Pre-1990s" period.

Panel C of Table 2 summarizes estimation results for the "Stable 1990s" period. The stable interest rates during this period ensure that the long run mean estimators (i.e. $\overline{r}$ and $\overline{\theta}$), the volatility estimators (i.e. σ_r, σ_θ and σ_v), and the speed of mean reversion estimators (i.e. κ_r, κ_θ and κ_v) are amongst the lowest across all sample periods. In early 1990s, the Federal Reserve Bank began to target interest rates quite vigorously and meet 8 times per year, since 1994. Not surprisingly, the market short rate tended to fluctuate in a narrow band around the interest rate target until the next target was announced. Thus, rates between successive FOMC meetings were quite stable. One possibility here is that the change in targets might be captured by up and down jumps, particularly given that our SVJ and CHENJ models estimate the total jumps during this period to be around 7. This is consistent with the fact that the Fed changed interest rate targets infrequently during this period.[4] Moreover, jump down sizes are much bigger than the jump up sizes during this period, in line with the empirical evidence.

Finally, "Post-1990s" estimation results are provided in Panel D of Table 2. In this period, the Federal Reserve Bank decreased interest rates from 6% in 2001 to 1% in 2003; and then increased rates to 5.25% by 2007, and decreased them to 2% by April 2008. The sharp up and down swings in the target rate induce an apparent bimodality in the distribution of the Eurodollar rate. The

[4] The Fed changed interest rate target 10 times in 1991, 3 times in 1992, 6 times in 1994, 3 times in 1995, once in 1997, 3 times in 1998, 1999, and 2000, and twice before March 2001. The Fed didn't change interest rate in 1993 and 1996. (From the Federal Reserve Board)

estimates for this period are in some sense "between" those reported for the "Pre-1990s" and "Stable 1990s" periods. The most interesting finding in this period is that both the jump up size and jump down size are bigger than any other period, and the total jump intensity is around 11 times per year.

5.3 Specification test results

Tables 6-9 report our specification results for the 4 sample periods. Tests are carried out using $\tau-$step ahead confidence intervals. We set $\tau = \{1, 2, 4, 12\}$, corresponding to one week, two weeks, one month, and one quarter ahead conditional distribution evaluation periods. The confidence intervals that we use in test statistic construction are chosen based on the properties of our historical data. In particular, we set $\underline{u}$ and $\overline{u}$ equal to $\overline{X} \pm \sigma_X$ and $\overline{X} \pm 0.5\sigma_X$, respectively. Additionally, we set our simulation sample length as $S = 10T$,where T is the historical sample length. The simulation sample length for latent variables is set at $N = 10T$. In our implementation of the bootstrap, we set block length to be 20, and carry out 100 bootstrap replications. In the tables, test statistics (denoted by V_T) and 5%, 10%, 15% and 20% bootstrap critical values are given. Single starred entries denote rejection at the 10% level.

5.3.1 "Post Bretton-Woods"

Results are presented in Table 3. Not surprisingly, the CIR model is rejected for all τ, with any confidence interval. Moreover, the SV model performs the best amongst the CIR, SV and SVJ models. These results are in line with those reported in BCS (2008). Interestingly, though, we find that the three-factor model (CHEN) not considered in BCS (2008) performs at least as well as the SV model. moreover, the failure to reject the CHEN model at the 20% level indicates that the CHEN model is in some crude sense "better" than the SV model. Increased model complexity may help capture the spot rate dynamics when long samples of data across many historical regimes are used to calibrate models. Of further note is that the CHENJ model performs better than the SVJ model.

To further illustrate the findings of Table 3, we plot kernel densities of our simulated data and actual data in the Panel A of Figure 2, for selected models. Specifically, we choose points that represent the left tail, middle points and right tail of our historical data as evaluation points, and construct kernel density estimates. Figure 2a contains plots of the simulated density at $x = 0.03$.

Compared with the CIR, the SV and CHEN - simulated data are more concentrated around the actual data point. Moreover, the 3-factor CHEN model has higher kurtosis than the SV model. These findings are in line with the specification test result suggesting that CHEN is superior to the other candidate models. Similar results are reported in Figures 2b-d for other evaluation points.

5.3.2 "Pre 1990s"

Table 4 presents specification results for the "Pre-1990s" period. Although some models perform relatively better than others, the extremely high rejection frequency indicates that none of the six models adequately capture interest rate dynamics for this period. This finding is consistent with Hong and Li (2005), where their univariate model is rejected without exception, multifactor models provide only a little improvement, and even their most sophisticated model does not capture the dynamics of short rates. Panel B of Figure 2 contains plots of simulated densities for the CIR, SV and CHEN models. Figures 2e-h exhibit that none of the models yield simulated densities that are centered around the actual data points. The CIR model yields the worst results, amongst the three models. Note also that as the evaluation point is increased from $x = 0.08$, to $x = 0.20$, the simulated densities for the three models move further away from the actual points, suggesting, as expected, the difficulties in mimicking tail behavior.

5.3.3 "The Stable 1990s"

The most interesting result we find in Table 5 is that the univariable CIR model beats all multifactor models in the "Stable 1990s" period. The CIR model fails to be rejected for all values of τ, and for all confidence intervals, at any confidence level. This result is in stark contrast to our findings for the whole sample and the "Pre-1990s" periods. As discussed in Section 5.1, this is perhaps not surprising given the stable monetary regime of the 1990s; and reminds us that our different models perform very differently depending upon the particular regime from whence the data we fit are taken. Figures 3a-d display plots of the simulated densities for the CIR, SM and CHENJ models. We choose the SM model instead of the SV model in the Panel A of Figure 3 because the SM model is the "second best" from amongst the six candidate models. The CHENJ model is depicted in order to illustrate how the jump process distorts the simulated densities in this stable period. As evidenced upon inspection of Figure 3a, the CIR model is superior to the SM model in that CIR-simulated data is more concentrated around the actual evaluation point. As expected, the simulated density for the CHENJ model is far from the evaluation point. Similar conclusions emerge upon examination of figures 3c-d.

5.3.4 "Post-1990s"

The bimodality associated with the "Post-1990s" period makes our specification test results quite different from those reported for our other sample periods. The SM model "outperforms" other models. The rejection frequency for the CIR model is low as well. Surprisingly, we fail to reject the null of correct specification for the complex three-factor with jumps CHENJ model. These results are sensible given the underlying characteristics in this period. The SM model captures the changing mean in this period. However, since the Federal Reserve Bank changed the target rate with higher frequency and larger magnitudes during this period, the CHENJ model which combines stochastic mean and jump components also performs well.[5] Moreover, the CHENJ model appears to be able to capture possible outliers in this period via its inclusion of two latent variables. Figures 3e-f contain plots of empirical densities for the SM, CIR and SV models at the two modal points ($x = 0.02$ and $x = 0.05$). The densities associated with the SM model have higher peak and narrower tails than those for the CIR model, at both evaluation points. As expected, the SV model has densities with very fat tails; but these are centered far from the evaluation point. Of final note is that there are evaluation points for which other models do beat the SM model, but the SM performs best in overall. This is a feature which should be expected, and underscores the importance of using not only the portmanteau type specification tests but also individual densities for evaluating alternative models.

6 Concluding Remarks

This paper implements simulation based specification tests in order to study the performance of different affine multifactor diffusion processes across various historical sample periods. We begin with outlining the simulation testing framework used to carry out our specification analysis. thereafter, we examine Eurodollar interest rate data using (i) a basic statistical analysis; (ii) specification tests; and (iii) simulation based kernel density estimates of distributions around particular evaluation points on the support of our historical dataset. we find that the CIR model, which is often rejected in the literature, performs best among the candidate models in "The Stable 1990s", while there is little to choose between one and two factor models when considering the "Post 1990s" pe-

[5] The Fed changed interest rate target 11 times in 2001, once in 2002 and 2003, 5 times in 2004, 8 times in 2005, 4 times in 2006 and 2008, 3 times in 2007.

riod. Examination of "Pre 1990s" data, on the other hand, suggests there is little to choose between 2 and 3 factor models, and the one factor CIR model performs poorly. Interestingly, when our entire sample of weekly data from 1970 to 2008 is used to estimate the competing models, the "best" performer is the three factor CHEN (1996) model examined by Andersen, Benzoni and Lund (2004). We conclude that the choice of model for simulating the future distribution of short rates is highly sample dependent; and when long samples of data are used which likely span different economic regimes, more complicated models are likely to be selected as "best" than is actually warranted, particularly if the current regime is expected to continue throughout the length of the simulation period of interest. Put differently, structural breaks appear to be an important component of simple diffusion models of the short rate.

Many topics for further research remain. For example, from a theoretical perspective, it remains to construct specification tests that do not integrate out the effects of latent factors. It also remains to construct truly "ex ante" model selection type predictive accuracy tests using the techniques discussed in this paper as a starting point. From an empirical perspective, it remains to determine whether it is in some sense "optimal" to fit models to shorter data samples when simulating future scenarios, and if so, exactly how "short" should samples be?

7 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, 2006, Nonparametric Transition-Based Tests for Diffusions, Working paper.

Aït-Sahalia, Y. and R. Kimmel, 2007, Maximum Likelihood Estimation of Stochastic Volatility Models, Journal of Financial Economics, 2007, 83, 413-452.

Aït-Sahalia, Y., 2008, Closed-Form Likelihood Expansions for Multivariate Diffusions, Annals of Statistics, 2008, 36, 906-937.

Andersen, T.G., and J. Lund, 1997, Estimating Continuous Time Stochastic Volatility Models of Short Term Interest Rates, Journal of Econometrics 72, 343-377.

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

Balduzzi, P., S.R. Das, and S. Foresi, 1998, The Central Tendency: A second Factor in bond Yields, Review of Economic and Statistics 80, 62-72.

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

Bhardwaj, G., V. Corradi and N.R. Swanson, 2008, A Simulation Based Specification Test for Diffusion Processes, Journal of Business and Economic Statistics, 26, 176-93.

Black, F. and M. Scholes, 1973. The pricing of options and corporate liabilities, Journal of Political Economy, 81, 637–654.

Chacko, G., and S. Das, 2002, Pricing Interest Rate Derivatives: A General Approach, Review of Financial Studies 15, 195-241.

Chacko, G., and L. Viceira, 2003, Spectral GMM Estimation of Continuous-Time Processes, Journal of Econometrics 116, 259-292.

Chan, K.C., G. A. Karolyi, F. A. Longstaff and A. B. Sanders, 1992, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, Journal of Finance 47, 1209-1227.

Chen, L, 1996, Stochastic Mean and Stochastic Volatility - A Three-Factor Model of Term Structure of Interest Rates and its Application to the Pricing of Interest Rate Derivatives, Blackwell Publishers, Oxford, UK.

Chen, B. and Y. Hong, 2005, Diagnostics of Multivariate Continuous-Time Models with Application to Affine Term Structure Models, Working Paper, Cornell University.

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

Corradi, V. and N.R. Swanson, 2007, Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes, International Economic Review, February 2007, 48, 67-109.

Corradi, V. and N.R. Swanson, 2008, Predictive Density Accuracy Tests for Diffusion Processes, Working paper, University of Warwick.

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. and C. Li, 2006, Forecasting the Term Structure of Government Bond Yields, Journal of Econometrics, 130, 337-364.

Diebold, F.X., C. Li, and V.Z. Yue, 2007, Global Yield Curve Dynamics and Interactions: A Generalized Nelson-Siegel Approach, NBER working paper.

Duffee, G. R., 2002, Term Premia and Interest Rate Forecasts in Affine Models, Journal of Finance 57, 405-443.

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

Durham, G. B., 2003, Likelihood-Based Specification Analysis of Continuous-Time Models of the Short-Term Interest Rate, Journal of Financial Economics 70, 463-487.

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

Gallant, A.R., and G. Tauchen 1997, Estimation of Continuous Time Models for Stock Returns and Interest Rates, Macroeconomic Dynamics 1, 135-168.

Heston, S., 1993, A closed-form solution for options with stochastic volatility with applications to bonds and currency options, Review of Financial Studies 6, 327-343.

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.T. Li, and F. Zhao, 2007, Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates, Journal of Econometrics, 141, 736-776.

Hull, J., and A. White, 1987, The Pricing of Options on Assets with Stochastic Volatility, Journal of Finance, XLII, 281-300.

Hull, J., and A. White, 1990, Pricing interest-rate derivative securities, The Review of Financial Studies, Vol 3, No. 4, 573-592.

Jiang, G., and J. L. Knight, 2002. Estimation of Continuous Time Processes via Empirical Characteristic Function, Journal of Business and Economic Statistics 20, 198–212.

Piazzesi, M., 2001, Macroeconomic jump effects and the yield curve, working paper, UCLA.

Piazzesi, M., 2004, Affine Term Structure Models, Manuscript prepared for the Handbook of Financial Econometrics, University of California at Los Angeles.

Piazzesi, M., 2005, Bond Yields and the Federal Reserve, Journal of Political Economy 113, 311-344.

Sullivan R., A. Timmermann, and H. White, 1999, Data-Snooping, Technical Trading Rule Performance, and the Bootstrap, The Journal of Finance, vol. 54, issue 5, 1647-1691.

Sullivan, R., A. Timmermann, and H. White, 2001. Dangers of data-driven inference: The case of calendar effects in stock returns. Journal of Econometrics, 249– 286.

Tauchen, G., 1997, New Minimum Chi-Square Methods in Empirical Finance, in Advances in Econometrics, Cambridge University Press, 279-317.

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

White, H., 2000, A Reality Check for Data Snooping, Econometrica, 68, 1097-1126.

Vasicek, O. A., 1977, An Equilibrium Characterization of the Term Structure, Journal of Financial Economics 5, 177-188.

Table 2: Parameter estimates for various spot interest rate models

	Panel A: Sample period 01/1971-04/2008						Panel B: Sample period 01/1971-02/1991					
Parameter	CIR	SM	SV	SVJ	CHEN	CHENJ	CIR	SM	SV	SVJ	CHEN	CHENJ
κ_r	0.3068	0.2169	0.2104	0.2415	0.1261	0.4037	0.6707	0.2680	0.2633	0.3012	0.2410	0.4311
	(0.1872)	(0.2892)	(0.0196)	(0.0084)	(0.5749)	(0.0414)	(0.2815)	(1.3272)	(0.0217)	(0.0104)	(0.5612)	(0.0541)
$\overline{r}$	0.0657		0.0555	0.0739			0.0924		0.0831	0.0719		
	(0.0383)		(0.0293)	(0.0930)			(0.0583)		(0.0332)	(0.0876)		
σ_r	0.1180	0.0168					0.1304	0.0139				
	(0.0281)	(0.0181)					(0.0618)	(0.0241)				
κ_θ		0.2926			0.3373	0.2763		0.4996			0.3012	0.2248
		(1.8639)			(0.5592)	(2.05324)		(1.3286)			(0.6714)	(3.0004)
$\overline{\theta}$		0.0558			0.0504	0.0659		0.0832			0.0844	0.0939
		(0.0163)			(0.0174)	(0.15981)		(0.2507)			(0.1402)	(0.1785)
σ_θ		0.0217			0.0021	0.1907		0.0015			0.0001	0.1845
		(0.6261)			(0.5669)	(9.26265)		(1.2231)			(0.6112)	(6.4524)
κ_v			2.8066	2.2931	2.8979	2.1159			3.5006	2.5922	3.004	2.1204
			(0.2488)	(0.0036)	(0.1150)	(1.90270)			(1.0124)	(1.0132)	(0.9802)	(1.8557)
$\overline{v}$			0.0003	0.0005	0.0002	0.0006			0.0002	0.0007	0.0002	0.0006
			(0.8117)	(0.0003)	(0.0003)	(0.00127)			(0.9278)	(0.0004)	(0.0012)	(0.0009)
σ_v			0.0191	0.0149	0.0054	0.0034			0.0124	0.0102	0.0094	0.0059
			(0.0319)	(0.0045)	(0.0006)	(0.02581)			(0.0571)	(0.0158)	(0.0007)	(0.0412)
ρ			-0.1921	-0.1050					-0.3379	-0.1105		
			(0.00966)	(0.0471)					(0.0012)	(0.0612)		
λ_u				5.4919		8.1042				7.8910		9.1260
				(10.3636)		(2.22153)				(9.2389)		(4.1125)
ζ_u				0.0007		0.0005				0.0028		0.0001
				(0.0035)		(0.00455)				(0.0052)		(0.0047)
λ_d				3.1210		2.8334				3.5981		3.8258
				(5.2057)		(1.20637)				(6.0245)		(1.3107)
ζ_d				0.0011		0.0007				0.0027		0.0000
				(0.0114)		(0.0100)				(0.0187)		(0.0100)
	Panel C: Sample period 03/1991-01/2001						Panel D: Sample period 02/2001-04/2008					
Parameter	CIR	SM	SV	SVJ	CHEN	CHENJ	CIR	SM	SV	SVJ	CHEN	CHENJ
κ_r	0.2590	0.0256	0.0380	0.3018	0.0611	0.2707	0.0661	1.5631	0.0501	0.2959	0.0494	0.3638
	(0.1872)	(0.2892)	(0.0196)	(0.0084)	(0.5749)	(0.0414)	(0.1872)	(0.2892)	(0.0196)	(0.0084)	(0.5749)	(0.0414)
$\overline{r}$	0.0493		0.0306	0.0538			0.0066		0.0297	0.0341		
	(0.0383)		(0.0293)	(0.0930)			(0.0383)		(0.0293)	(0.0930)		
σ_r	0.0325	0.0011					0.0270	0.0072				
	(0.0281)	(0.0181)					(0.0281)	(0.0181)				
κ_θ		0.3700			0.2792	0.3072		0.0948			0.3175	0.2870
		(0.8639)			(0.5592)	(2.0534)		(0.8639)			(0.5592)	(2.0534)
$\overline{\theta}$		0.0262			0.0464	0.0635		0.0369			0.0395	0.0501
		(0.0163)			(0.0174)	(0.1591)		(0.0163)			(0.0174)	(0.1591)
σ_θ		0.1393			0.0000	0.1175		0.0559			0.0032	0.1844
		(0.6261)			(0.5669)	(9.2625)		(0.6261)			(0.5669)	(9.2625)
κ_v			3.4392	2.5901	3.1010	2.1182			2.8131	2.4921	2.1006	2.1192
			(0.2488)	(0.0036)	(0.1150)	(1.9027)			(0.2488)	(0.0036)	(0.1150)	(1.9027)
$\overline{v}$			0.0000	0.0000	0.0000	0.0000			0.0000	0.0001	0.0001	0.0005
			(0.8117)	(0.0003)	(0.0003)	(0.0012)			(0.8117)	(0.0003)	(0.0003)	(0.0012)
σ_v			0.0014	0.0098	0.0056	0.0001			0.0001	0.0053	0.0011	0.0056
			(0.0319)	(0.0045)	(0.0006)	(0.0258)			(0.0319)	(0.0045)	(0.0006)	(0.0258)
ρ			-0.4783	-0.1027					-0.0080	-0.1039		
			(0.00966)	(0.0471)					(0.00966)	(0.0471)		
λ_u				5.8901		5.1601				9.8901		8.1663
				(10.3636)		(2.2215)				(10.3636)		(2.2215)
ζ_u				0.0006		0.0006				0.0008		0.0007
				(0.0035)		(0.00455)				(0.0035)		(0.00455)
λ_d				1.6102		2.8361				2.5109		2.8284
				(5.2057)		(1.2063)				(5.2057)		(1.2063)
ζ_d				0.0030		0.0016				0.0014		0.0022
				(0.0114)		(0.0100)				(0.0114)		(0.0100)

Notes: Results are based on exactly identified GMM estimation, using conditional moments. Numerical values in parentheses are standard errors. Parameter estimates are constructed using weekly interest rate data expressed in decimal form on a yearly basis. The dataset includes 1996 weekly observations on the Eurodollar deposit Rate for the period 01/1971 to 04/2008.

Table 3: BCS Specification Test Results - "Post Bretton Woods" (01/1971-04/2008)

τ	$(\underline{u},\overline{u})$	Z_T	5% CV	10% CV	15% CV	20% CV
		Panel A: CIR model				
1	$\overline{X} \pm 0.5\sigma_X$	7.6849*	3.1545	2.547	2.4623	2.3259
	$\overline{X} \pm \sigma_X$	3.2358*	2.5774	2.1852	1.8792	1.7112
2	$\overline{X} \pm 0.5\sigma_X$	7.8948*	3.4789	2.8263	2.6949	2.5459
	$\overline{X} \pm \sigma_X$	3.5375*	3.4202	2.7714	2.2988	2.1371
4	$\overline{X} \pm 0.5\sigma_X$	8.1198*	4.0356	3.6690	3.2178	3.0321
	$\overline{X} \pm \sigma_X$	4.6621*	4.1774	3.4463	2.7442	2.5116
12	$\overline{X} \pm 0.5\sigma_X$	8.2578*	4.9334	4.6311	4.1175	3.8153
	$\overline{X} \pm \sigma_X$	4.3236*	4.3271	4.0430	3.2378	2.9430
		Panel B: SM model				
1	$\overline{X} \pm 0.5\sigma_X$	2.5413*	2.3687	2.0954	1.9418	1.7709
	$\overline{X} \pm \sigma_X$	1.3541	1.9753	1.5800	1.5287	1.3832
2	$\overline{X} \pm 0.5\sigma_X$	4.9409*	2.865	2.5337	2.3589	2.1716
	$\overline{X} \pm \sigma_X$	1.9275	2.7766	2.1624	2.0454	1.8669
4	$\overline{X} \pm 0.5\sigma_X$	4.458*	3.6752	3.0572	2.8474	2.6978
	$\overline{X} \pm \sigma_X$	2.6682	3.6691	2.8754	2.5688	2.4361
12	$\overline{X} \pm 0.5\sigma_X$	4.6565*	4.9105	4.5935	4.1004	3.7776
	$\overline{X} \pm \sigma_X$	3.6026	4.2818	3.9715	3.1955	2.896
		Panel C: SV model				
1	$\overline{X} \pm 0.5\sigma_X$	2.7497*	2.1241	1.8347	1.6676	1.5071
	$\overline{X} \pm \sigma_X$	1.588	2.0788	1.6446	1.5606	1.3061
2	$\overline{X} \pm 0.5\sigma_X$	4.5939*	3.0177	2.4339	2.3535	2.1769
	$\overline{X} \pm \sigma_X$	2.1179	2.8281	2.2565	1.9379	1.8781
4	$\overline{X} \pm 0.5\sigma_X$	4.7983*	3.7316	3.0996	2.9495	2.7172
	$\overline{X} \pm \sigma_X$	2.7067	3.6797	2.986	2.5461	2.4111
12	$\overline{X} \pm 0.5\sigma_X$	4.5106	4.9098	4.6402	4.1072	3.7795
	$\overline{X} \pm \sigma_X$	3.3907	4.309	3.9615	3.2079	2.9145
		Panel D: SVJ model				
1	$\overline{X} \pm 0.5\sigma_X$	7.6238*	3.3036	2.7993	2.6599	2.4747
	$\overline{X} \pm \sigma_X$	3.6849*	3.1811	2.4833	2.3179	2.0802
2	$\overline{X} \pm 0.5\sigma_X$	9.3633*	3.5341	3.0816	2.8441	2.7297
	$\overline{X} \pm \sigma_X$	6.8586*	3.8487	3.09	2.7141	2.5736
4	$\overline{X} \pm 0.5\sigma_X$	9.4501*	3.9981	3.5636	3.1653	2.9555
	$\overline{X} \pm \sigma_X$	8.6409*	4.0904	3.3501	2.882	2.6444
12	$\overline{X} \pm 0.5\sigma_X$	9.5099*	4.9325	4.6329	4.1214	3.8286
	$\overline{X} \pm \sigma_X$	7.5193*	4.3004	4.0379	3.2431	2.9499
		Panel E: CHEN model				
1	$\overline{X} \pm 0.5\sigma_X$	2.3669*	2.0425	1.9033	1.7238	1.5467
	$\overline{X} \pm \sigma_X$	0.9058	1.6132	1.3218	1.1854	1.1292
2	$\overline{X} \pm 0.5\sigma_X$	2.2719	2.5855	2.3067	2.1109	1.9748
	$\overline{X} \pm \sigma_X$	1.2835	2.0783	1.8050	1.6640	1.5379
4	$\overline{X} \pm 0.5\sigma_X$	4.6785*	3.3399	2.8554	2.5715	2.4243
	$\overline{X} \pm \sigma_X$	1.9738	3.2468	2.4091	2.3106	2.0996
12	$\overline{X} \pm 0.5\sigma_X$	5.1255*	4.8303	4.5125	4.0266	3.7428
	$\overline{X} \pm \sigma_X$	2.7088	4.2981	3.7180	3.1320	2.8495
		Panel F: CHENJ model				
1	$\overline{X} \pm 0.5\sigma_X$	2.0889	3.1863	2.6303	2.4718	2.3248
	$\overline{X} \pm \sigma_X$	2.2357	2.8803	2.3542	1.9214	1.8177
2	$\overline{X} \pm 0.5\sigma_X$	2.3842	3.582	3.0588	2.8634	2.6814
	$\overline{X} \pm \sigma_X$	3.4545*	3.7212	2.9444	2.5755	2.3482
4	$\overline{X} \pm 0.5\sigma_X$	4.4168*	4.0294	3.6211	3.1599	2.9629
	$\overline{X} \pm \sigma_X$	7.5231*	4.2088	3.4633	2.9612	2.5637
12	$\overline{X} \pm 0.5\sigma_X$	4.5989	4.9292	4.6377	4.1277	3.8382
	$\overline{X} \pm \sigma_X$	10.2708*	4.291	4.0231	3.2393	2.9466

Notes: Numerical entries in the table are specification test statistics (Z_T) and 5% &, 10%,15% & 20% nominal level critical values, for tests constructed using intervals given in the first column of the table, and for τ =1, 2, 4, 12 (see discussion in Section 5.3 for complete details). Single starred entries denote rejection at the 10% level. The simulation periods considered is $10T$, and T denotes the number of observations in the sample. The block length is set equal to 20 observations, and empirical bootstrap distributions are constructed using 100 bootstrap replications. See Section 2 for further details.

Table 4: BCS Specification Test Results - "Pre 1990s" (01/1971-02/1991)

τ	$(\underline{u},\overline{u})$	Z_T	5% CV	10% CV	15% CV	20% CV
		Panel A: CIR model				
1	$\overline{X} \pm 0.5\sigma_X$	3.0621*	3.1776	2.7526	2.5984	2.3811
	$\overline{X} \pm \sigma_X$	4.3167*	2.3715	2.2112	1.9665	1.8351
2	$\overline{X} \pm 0.5\sigma_X$	3.2222*	3.2599	3.0063	2.6997	2.4409
	$\overline{X} \pm \sigma_X$	3.7536*	2.5201	2.3269	2.0755	2.0159
4	$\overline{X} \pm 0.5\sigma_X$	3.2372*	3.2051	3.0603	2.8355	2.5866
	$\overline{X} \pm \sigma_X$	3.9062*	2.6936	2.4758	2.2142	2.0143
12	$\overline{X} \pm 0.5\sigma_X$	3.5181*	3.4822	3.207	2.9923	2.829
	$\overline{X} \pm \sigma_X$	4.8191*	3.3907	2.934	2.5063	2.2456
		Panel B: SM model				
1	$\overline{X} \pm 0.5\sigma_X$	2.1246*	1.894	1.7853	1.5781	1.4595
	$\overline{X} \pm \sigma_X$	1.4402*	1.5117	1.291	1.1216	1.0835
2	$\overline{X} \pm 0.5\sigma_X$	2.9384*	2.3969	2.208	2.0303	1.9181
	$\overline{X} \pm \sigma_X$	2.7169*	1.9704	1.8251	1.6266	1.4405
4	$\overline{X} \pm 0.5\sigma_X$	4.4785*	2.868	2.6531	2.4043	2.3229
	$\overline{X} \pm \sigma_X$	4.1935*	2.4847	2.2121	2.0652	1.7826
12	$\overline{X} \pm 0.5\sigma_X$	5.108*	3.4445	3.2023	2.9759	2.8799
	$\overline{X} \pm \sigma_X$	4.9432*	3.3752	2.9254	2.5031	2.242
		Panel C: SV model				
1	$\overline{X} \pm 0.5\sigma_X$	2.2098*	1.8653	1.711	1.5733	1.4188
	$\overline{X} \pm \sigma_X$	1.2619	1.4679	1.3092	1.1554	1.1051
2	$\overline{X} \pm 0.5\sigma_X$	3.0606*	2.4498	2.2609	2.0829	1.9631
	$\overline{X} \pm \sigma_X$	2.4584*	1.9421	1.7792	1.5924	1.4197
4	$\overline{X} \pm 0.5\sigma_X$	4.1544*	2.946	2.6783	2.4498	2.3623
	$\overline{X} \pm \sigma_X$	3.7409*	2.4776	2.2168	2.0691	1.8145
12	$\overline{X} \pm 0.5\sigma_X$	4.5846*	3.4287	3.1896	2.9691	2.8689
	$\overline{X} \pm \sigma_X$	4.9187*	3.3818	2.9549	2.5059	2.2432
		Panel D: SVJ model				
1	$\overline{X} \pm 0.5\sigma_X$	3.4909*	2.7795	2.6401	2.3864	2.2027
	$\overline{X} \pm \sigma_X$	4.3968*	2.07	1.7367	1.5673	1.3806
2	$\overline{X} \pm 0.5\sigma_X$	4.2639*	3.1524	2.759	2.5538	2.3089
	$\overline{X} \pm \sigma_X$	5.9699*	2.6649	2.1528	1.9113	1.7427
4	$\overline{X} \pm 0.5\sigma_X$	4.9671*	3.1785	3.0284	2.723	2.4721
	$\overline{X} \pm \sigma_X$	7.4484*	2.6825	2.3997	2.1844	1.9203
12	$\overline{X} \pm 0.5\sigma_X$	6.1157*	3.474	3.2245	2.9963	2.8029
	$\overline{X} \pm \sigma_X$	8.2848*	3.3933	2.9431	2.5166	2.2336
		Panel E: CHEN model				
1	$\overline{X} \pm 0.5\sigma_X$	2.0832*	1.8622	1.736	1.562	1.4393
	$\overline{X} \pm \sigma_X$	1.2411	1.3957	1.2458	1.0909	1.0293
2	$\overline{X} \pm 0.5\sigma_X$	2.8921*	2.3902	2.2048	2.0296	1.9082
	$\overline{X} \pm \sigma_X$	2.1367*	1.8986	1.6724	1.5164	1.4043
4	$\overline{X} \pm 0.5\sigma_X$	3.9971*	2.9022	2.6382	2.4548	2.2916
	$\overline{X} \pm \sigma_X$	3.4145*	2.4277	2.1706	2.0134	1.7298
12	$\overline{X} \pm 0.5\sigma_X$	4.4202*	3.4439	3.1872	2.9723	2.8697
	$\overline{X} \pm \sigma_X$	4.4658*	3.3772	2.9859	2.5033	2.2207
		Panel F: CHENJ model				
1	$\overline{X} \pm 0.5\sigma_X$	3.7299*	2.9199	2.6056	2.3829	2.1805
	$\overline{X} \pm \sigma_X$	1.7909	1.9087	1.7949	1.5684	1.447
2	$\overline{X} \pm 0.5\sigma_X$	2.9169	3.1639	2.9512	2.5912	2.4237
	$\overline{X} \pm \sigma_X$	4.2558*	2.393	2.1676	1.9767	1.8339
4	$\overline{X} \pm 0.5\sigma_X$	5.3726*	3.1566	3.0369	2.776	2.5466
	$\overline{X} \pm \sigma_X$	8.8116*	2.6327	2.4196	2.2155	2.0058
12	$\overline{X} \pm 0.5\sigma_X$	7.3999*	3.4854	3.2477	3.0059	2.8021
	$\overline{X} \pm \sigma_X$	11.0151*	3.4029	2.9527	2.5128	2.2434

Notes: See notes to Table 3.

Table 5: BCS Specification Test Results - "The Stable 1990s" (03/1991-01/2001)

τ	$(\underline{u},\overline{u})$	Z_T	5% CV	10% CV	15% CV	20% CV
		Panel A: CIR model				
1	$\overline{X} \pm 0.5\sigma_X$	1.743	2.6772	2.519	2.3908	2.2301
	$\overline{X} \pm \sigma_X$	1.4215	1.7076	1.5674	1.4895	1.3517
2	$\overline{X} \pm 0.5\sigma_X$	1.9155	2.9635	2.6364	2.4659	2.3378
	$\overline{X} \pm \sigma_X$	1.4151	2.3901	2.1124	1.9591	1.7915
4	$\overline{X} \pm 0.5\sigma_X$	1.7242	3.2826	2.9931	2.7374	2.501
	$\overline{X} \pm \sigma_X$	2.3068	3.2457	3.0113	2.4816	2.3736
12	$\overline{X} \pm 0.5\sigma_X$	1.7088	3.47	2.9462	2.7378	2.3888
	$\overline{X} \pm \sigma_X$	2.6992	3.557	3.1466	2.9817	2.8276
		Panel B: SM model				
1	$\overline{X} \pm 0.5\sigma_X$	2.2283*	1.8019	1.5585	1.4458	1.3906
	$\overline{X} \pm \sigma_X$	0.2205	0.4275	0.3689	0.3253	0.2926
2	$\overline{X} \pm 0.5\sigma_X$	3.3549*	2.4955	2.2605	1.8627	1.7101
	$\overline{X} \pm \sigma_X$	0.3074	0.6781	0.5977	0.5494	0.493
4	$\overline{X} \pm 0.5\sigma_X$	4.3109*	3.1	2.5784	2.347	2.0711
	$\overline{X} \pm \sigma_X$	0.4181	1.1967	1.0937	0.9361	0.86
12	$\overline{X} \pm 0.5\sigma_X$	3.1586	3.8378	3.2441	3.0625	2.7503
	$\overline{X} \pm \sigma_X$	0.8719	2.6774	2.4825	2.4048	2.3169
		Panel C: SV model				
1	$\overline{X} \pm 0.5\sigma_X$	2.1092	2.4951	2.1144	1.9801	1.8458
	$\overline{X} \pm \sigma_X$	0.2203	0.4064	0.3616	0.3616	0.3169
2	$\overline{X} \pm 0.5\sigma_X$	4.3226*	3.6056	2.8024	2.6646	2.5099
	$\overline{X} \pm \sigma_X$	0.2205	0.6309	0.5861	0.5412	0.4964
4	$\overline{X} \pm 0.5\sigma_X$	5.6127*	3.7268	3.2304	2.6068	2.3773
	$\overline{X} \pm \sigma_X$	0.221	1.171	1.1218	1.0363	1.032
12	$\overline{X} \pm 0.5\sigma_X$	7.2161*	4.8256	4.3276	4.0473	3.6399
	$\overline{X} \pm \sigma_X$	1.1136	3.0511	2.9087	2.8182	2.5466
		Panel D: SVJ model				
1	$\overline{X} \pm 0.5\sigma_X$	3.9202*	2.3198	1.972	1.8152	1.7045
	$\overline{X} \pm \sigma_X$	1.2275*	1.0654	0.9135	0.8631	0.7727
2	$\overline{X} \pm 0.5\sigma_X$	5.6993*	2.91	2.5822	2.2936	2.1378
	$\overline{X} \pm \sigma_X$	3.4687*	2.2945	1.9649	1.9119	1.6689
4	$\overline{X} \pm 0.5\sigma_X$	8.0184*	3.1449	2.8082	2.6843	2.3019
	$\overline{X} \pm \sigma_X$	4.6234*	3.1318	2.9289	2.3785	2.2673
12	$\overline{X} \pm 0.5\sigma_X$	7.1228*	3.5211	2.9906	2.5983	2.4542
	$\overline{X} \pm \sigma_X$	3.2521	3.6154	3.2698	3.0074	2.8243
		Panel E: CHEN model				
1	$\overline{X} \pm 0.5\sigma_X$	1.1898*	1.2982	1.1381	1.0555	0.9659
	$\overline{X} \pm \sigma_X$	0.0441	0.3134	0.2686	0.2686	0.2686
2	$\overline{X} \pm 0.5\sigma_X$	3.7492*	3.4063	2.779	2.5541	2.3748
	$\overline{X} \pm \sigma_X$	0.1764	0.6259	0.5363	0.4915	0.4915
4	$\overline{X} \pm 0.5\sigma_X$	5.3033*	3.7219	3.1929	2.5194	2.3847
	$\overline{X} \pm \sigma_X$	0.8397	1.3842	1.2046	1.1148	0.9984
12	$\overline{X} \pm 0.5\sigma_X$	9.3987*	3.8648	3.4118	3.197	2.9153
	$\overline{X} \pm \sigma_X$	3.3408*	3.1894	2.9631	2.7976	2.6462
		Panel F: CHENJ model				
1	$\overline{X} \pm 0.5\sigma_X$	1.2177	2.661	2.2292	1.9944	1.788
	$\overline{X} \pm \sigma_X$	2.024*	2.249	1.9447	1.8089	1.5967
2	$\overline{X} \pm 0.5\sigma_X$	1.9958	3.5217	3.0749	2.8393	2.6268
	$\overline{X} \pm \sigma_X$	3.5283*	3.2828	2.8685	2.608	2.3076
4	$\overline{X} \pm 0.5\sigma_X$	3.6745*	3.462	3.0426	2.8255	2.5908
	$\overline{X} \pm \sigma_X$	7.8632*	3.403	3.0572	2.7751	2.5688
12	$\overline{X} \pm 0.5\sigma_X$	4.0284*	3.5635	3.0355	2.7559	2.3334
	$\overline{X} \pm \sigma_X$	9.4419*	3.4723	3.0325	2.9193	2.7301

Notes: See notes to Table 3.

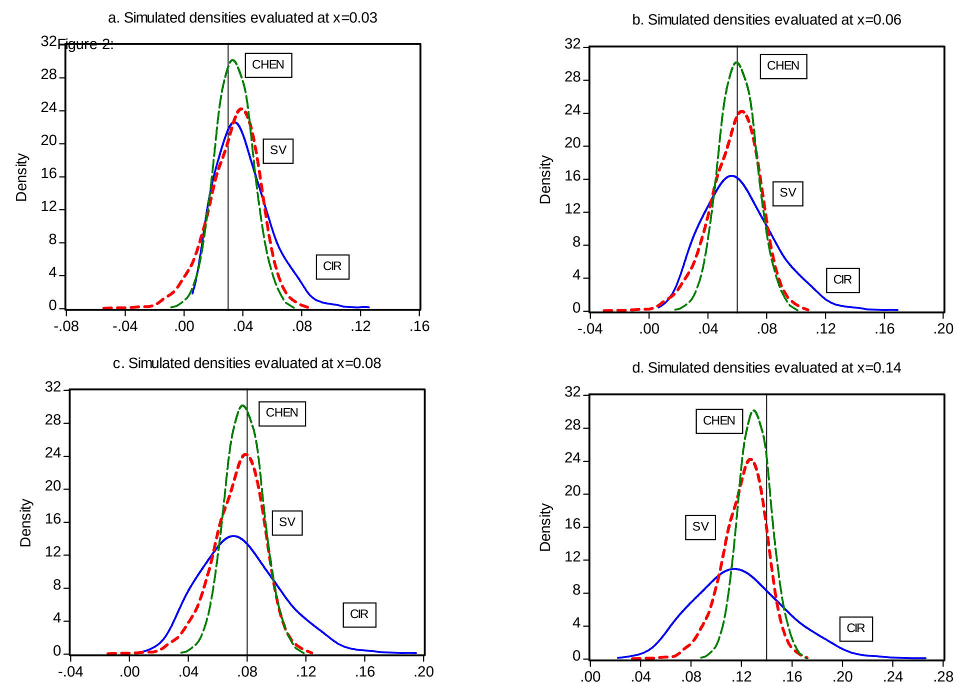
Table 6: BCS Specification Test Results - "Post 1990s" (02/2001-04/2008)

τ	$(\underline{u},\overline{u})$	Z_T	5% CV	10% CV	15% CV	20% CV
		Panel A: CIR model				
1	$\overline{X} \pm 0.5\sigma_X$	0.4175	1.1416	0.969	0.896	0.8026
	$\overline{X} \pm \sigma_X$	1.4744*	1.6624	1.3556	1.2694	1.1859
2	$\overline{X} \pm 0.5\sigma_X$	0.9852	1.9374	1.6573	1.5173	1.2902
	$\overline{X} \pm \sigma_X$	1.7948	2.5406	2.272	1.907	1.8487
4	$\overline{X} \pm 0.5\sigma_X$	1.5983	3.0862	2.5308	2.2108	1.8881
	$\overline{X} \pm \sigma_X$	1.904	3.4499	3.166	2.6652	2.4737
12	$\overline{X} \pm 0.5\sigma_X$	1.3776	3.6922	3.0861	2.7607	2.6046
	$\overline{X} \pm \sigma_X$	4.8384*	4.9892	4.3831	4.113	3.8376
		Panel B: SM model				
1	$\overline{X} \pm 0.5\sigma_X$	1.6069	3.5855	3.3937	3.1795	2.8959
	$\overline{X} \pm \sigma_X$	3.0268	4.1986	3.9257	3.4385	3.3061
2	$\overline{X} \pm 0.5\sigma_X$	1.5331	3.6565	3.3421	3.2023	2.8789
	$\overline{X} \pm \sigma_X$	2.0864	4.2483	3.7961	3.3232	3.2102
4	$\overline{X} \pm 0.5\sigma_X$	1.4989	3.4702	3.1957	3.0996	2.8728
	$\overline{X} \pm \sigma_X$	1.786	4.2643	3.764	3.3165	3.1994
12	$\overline{X} \pm 0.5\sigma_X$	1.0684	3.3503	2.9153	2.6545	2.473
	$\overline{X} \pm \sigma_X$	2.7817	4.4871	3.9436	3.752	3.4919
		Panel C: SV model				
1	$\overline{X} \pm 0.5\sigma_X$	0.3163	0.6696	0.5703	0.4852	0.461
	$\overline{X} \pm \sigma_X$	1.0861*	1.252	1.0647	0.9806	0.8843
2	$\overline{X} \pm 0.5\sigma_X$	0.5631	1.1241	0.9481	0.8816	0.8139
	$\overline{X} \pm \sigma_X$	1.8053*	1.7908	1.6109	1.4518	1.3552
4	$\overline{X} \pm 0.5\sigma_X$	1.3555	1.9716	1.6916	1.3984	1.3023
	$\overline{X} \pm \sigma_X$	2.9899*	2.8014	2.6081	2.2876	2.1157
12	$\overline{X} \pm 0.5\sigma_X$	3.7886*	3.6672	3.2848	2.7219	2.2963
	$\overline{X} \pm \sigma_X$	4.2665*	4.7145	4.0067	3.769	3.5775
		Panel D: SVJ model				
1	$\overline{X} \pm 0.5\sigma_X$	2.4083*	1.8806	1.8207	1.6944	1.6047
	$\overline{X} \pm \sigma_X$	6.1295*	3.5185	3.3276	3.1516	2.7994
2	$\overline{X} \pm 0.5\sigma_X$	5.8102*	3.5004	2.9123	2.7156	2.5503
	$\overline{X} \pm \sigma_X$	9.4311*	4.5093	3.8751	3.5117	3.2937
4	$\overline{X} \pm 0.5\sigma_X$	13.7926*	3.7881	3.4008	3.1998	3.0058
	$\overline{X} \pm \sigma_X$	9.4604*	4.416	3.8425	3.6305	3.3264
12	$\overline{X} \pm 0.5\sigma_X$	15.1804*	3.7861	3.5181	3.2501	3.1429
	$\overline{X} \pm \sigma_X$	9.568*	5.2128	4.1002	3.9262	3.7118
		Panel E: CHEN model				
1	$\overline{X} \pm 0.5\sigma_X$	0.6316	1.1613	1.0557	0.9502	0.8974
	$\overline{X} \pm \sigma_X$	4.4211*	3.8638	3.2832	2.9446	2.7554
2	$\overline{X} \pm 0.5\sigma_X$	3.4429	3.9571	3.4538	3.1609	3.0027
	$\overline{X} \pm \sigma_X$	4.4799*	4.1628	3.4456	3.2871	2.8943
4	$\overline{X} \pm 0.5\sigma_X$	8.6148*	6.2591	5.8883	4.8281	4.7224
	$\overline{X} \pm \sigma_X$	4.4924*	4.4947	3.9868	3.4023	3.2448
12	$\overline{X} \pm 0.5\sigma_X$	3.4744*	3.2064	2.9047	2.5646	2.4222
	$\overline{X} \pm \sigma_X$	4.3134*	6.8154	6.5045	6.0347	5.5766
		Panel F: CHENJ model				
1	$\overline{X} \pm 0.5\sigma_X$	1.8064	3.6712	3.4365	3.0852	2.82
	$\overline{X} \pm \sigma_X$	1.9258	4.1339	3.5039	3.3921	3.1739
2	$\overline{X} \pm 0.5\sigma_X$	1.1563	3.5044	3.363	3.105	2.9053
	$\overline{X} \pm \sigma_X$	3.4039	4.2549	3.6242	3.3535	3.2984
4	$\overline{X} \pm 0.5\sigma_X$	1.6572	3.4857	3.1576	3.0951	2.8673
	$\overline{X} \pm \sigma_X$	4.7218*	4.356	3.5848	3.3104	3.1274
12	$\overline{X} \pm 0.5\sigma_X$	1.6528	3.2081	2.9162	2.6076	2.4327
	$\overline{X} \pm \sigma_X$	5.6215	4.5307	4.0599	3.7381	3.4972

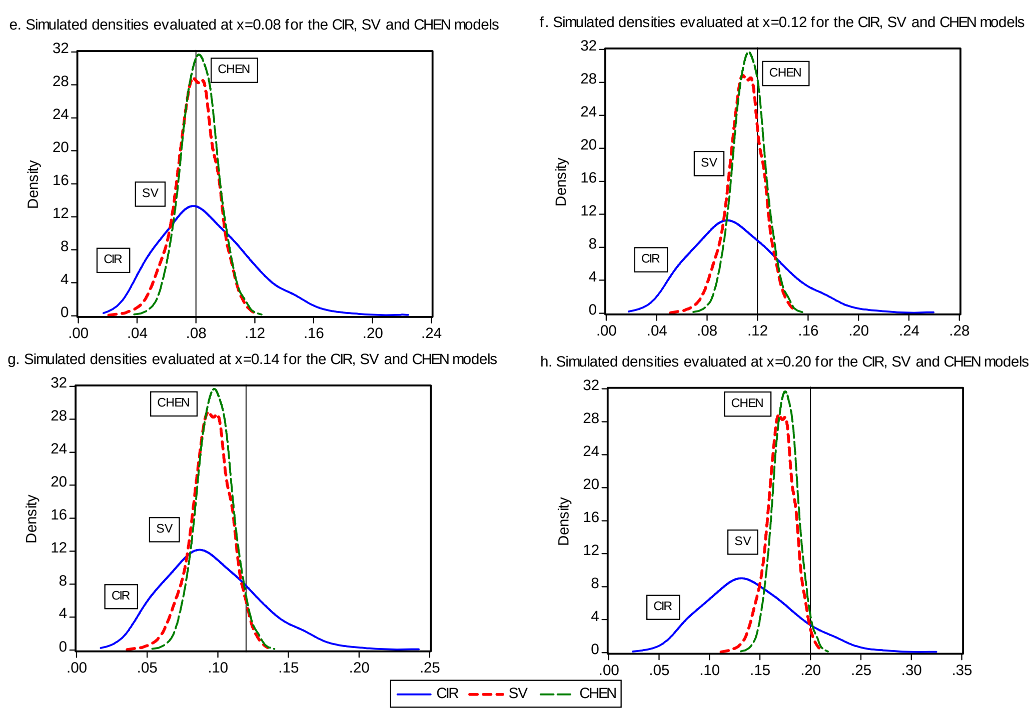
Notes: See notes to Table 3.

Figure 2: Simulated densities for ``Post Bretton-Woods" and ``Pre 1990s"

Panel A: Simulated densities for the CIR, SV and CHEN models - "Post Bretton-Woods" (01/1971-04/2008)



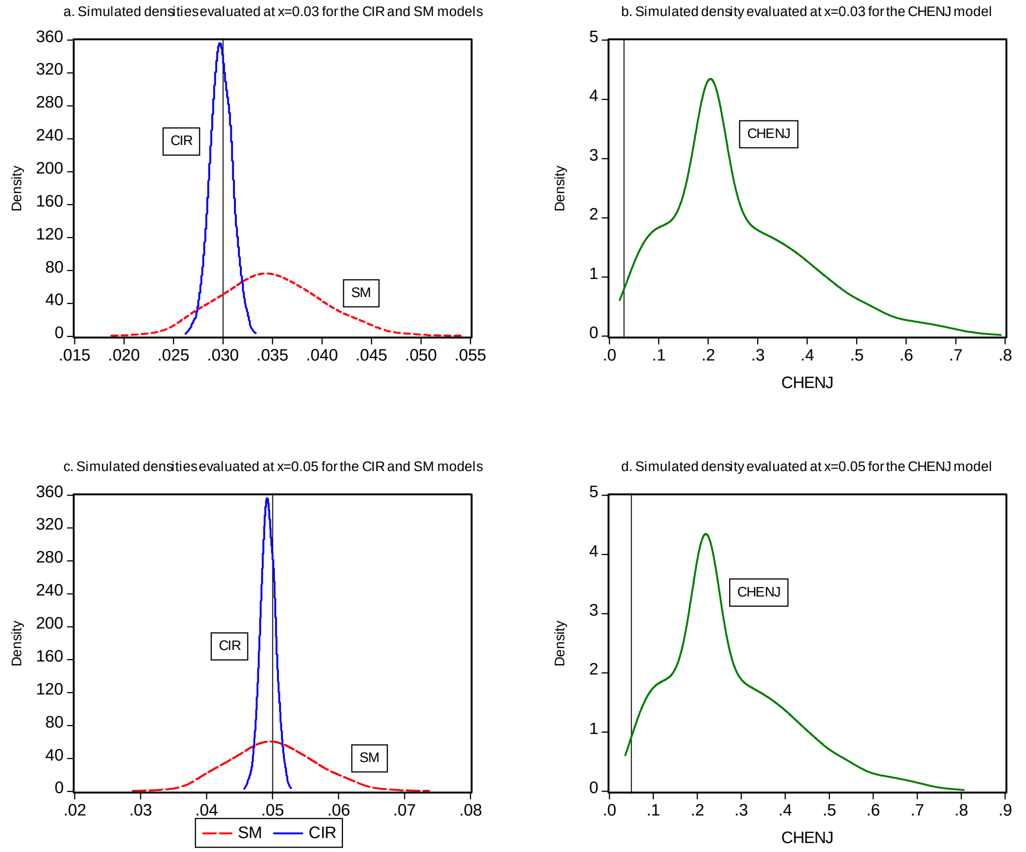
Panel B: Simulated densities for the CIR, SV and CHEN models -"Pre 1990" (01/1971-02/1991)



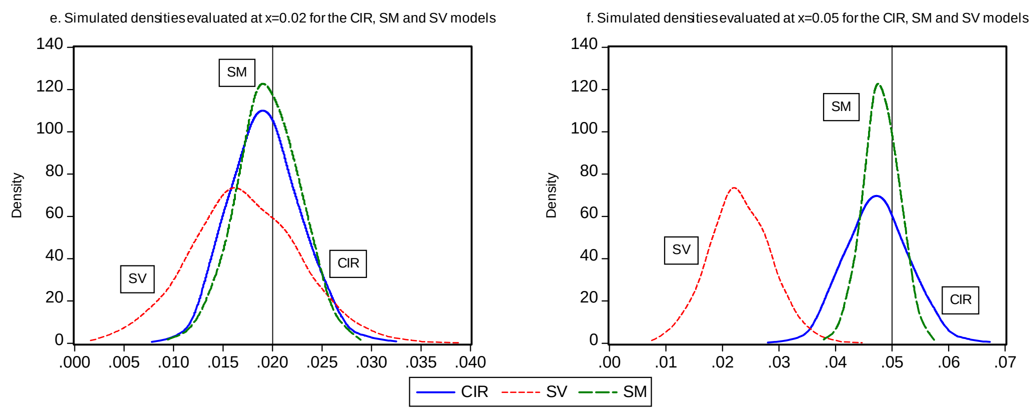
Note: This figure contained kernel density estimates for selected models and selected evaluation points, where evaluation points are taken from the support of the historical data, and correspond roughly to regions of the support sociated with mean or model behavior, as well as tail behavior.

Figure 3: Simulated densities for ``The Stable 1990s'' and ``Post 1990s''

Panel A: Simulated densities for the CIR, SM and CHENJ models - "The Stable 1990s" (03/1991-01/2001)



Panel B: Simulated densities for the CIR, SM and SV models - "Post 1990s" (02/2001-04/2008)



Notes: See Figure 2.