Journal of Empirical Finance 18 (2011) 743–764

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

In- and out-of-sample specification analysis of spot rate models: Further evidence for the period 1982–2008☆

Lili Cai [a], Norman R. Swanson [b,*]

[a] Department of Finance, Shanghai Jiao Tong University, 535 Fahuazhen Road, Shanghai, 200052, China
[b] Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA

ARTICLE INFO

Article history:
Received 15 April 2009
Received in revised form 12 May 2011
Accepted 13 May 2011
Available online 23 May 2011

JEL classification:
C1
C5
G0

Keywords:
Interest rate
Multi-factor diffusion process
Specification test
Out-of-sample forecasts
Conditional distribution
Model selection

ABSTRACT

We review and construct consistent in-sample specification and out-of-sample model selection tests on conditional distributions and predictive densities associated with continuous multifactor (possibly with jumps) and (non)linear discrete models of the short term interest rate. The results of our empirical analysis are used to carry out a “horse-race” comparing discrete and continuous models across multiple sample periods, forecast horizons, and evaluation intervals. Our evaluation involves comparing models during two distinct historical periods, as well as across our entire weekly sample of Eurodollar deposit rates from 1982 to 2008. Interestingly, when our entire sample of data is used to estimate competing models, the “best” performer in terms of distributional “fit” as well as predictive density accuracy, both in-sample and out-of-sample, is the three factor Chen (Chen, 1996) model examined by Andersen, Benzoni and Lund (2004). Just as interestingly, a logistic type discrete smooth transition autoregression (STAR) model is preferred to the “best” continuous model (i.e. the one factor Cox, Ingersoll, and Ross (CIR: 1985) model) when comparing predictive accuracy for the “Stable 1990s” period that we examine. Moreover, an analogous result holds for the “Post 1990s” period that we examine, where the STAR model is preferred to a two factor stochastic mean model. Thus, when the STAR model is parameterized using only data corresponding to a particular sub-sample, it outperforms the “best” continuous alternative during that period. However, when models are estimated using the entire dataset, the continuous CHEN model is preferred, regardless of the variety of model specification (selection) test that is carried out. Given that it is very difficult to ascertain the particular future regime that will ensue when constructing ex ante predictions, thus, the CHEN model is our overall “winning” model, regardless of sample period.

1. Introduction

Diffusion processes are used in virtually all aspects of continuous time finance from yield curve to exchange rate modeling, and 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 an ongoing effort to specify models that adequately capture the dynamics of financial variables across

☆ The authors owe many thanks to the editor, Franz Palm, and to three anonymous referees, all of whom provided numerous useful comments and suggestions on earlier versions of this paper. The authors are also grateful to Torben Andersen, Valentina Corradi, Frank Diebold, Walter Distaso, Elena Goldman, Roger Klein, John Landon Lane, and George Tauchen for helpful comments on earlier version of the paper. The authors also own many thanks to seminar participants at the University of Chicago and Rutgers University for numerous useful suggestions. Swanson gratefully acknowledges financial support from a Rutgers University Research Council grant.

* Corresponding author. Tel.: +1 908 415 0638.
E-mail addresses: lilycc@gmail.com (L. Cai), nswanson@econ.rutgers.edu (N.R. Swanson).

0927-5398/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2011.05.008

reasonable spans of time, rather than across specific historical episodes. In this paper we first review recent methodological advances in the area of specification and predictive accuracy testing, and subsequently undertake a specification search of alternative short rate models, thereby adding to the rich literature begun by the key research of Chan et al. (1992). Our search focuses on a variety of multi factor continuous models both with and without jumps as well as simple and nonlinear discrete models.

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. Black and Scholes, 1973; Cox et al., 1985; Hull and White, 1990; Vasicek model, 1977). Indeed, 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 et al. (2004)). This issue has implications not only for pricing formulae derived from these 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, Aït-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 and other approximations (as well as general work on conditional Kolmogorov testing — see e.g. Andrews (1997) and Corradi and Swanson (2005a)) have led to the development of numerous consistent specification tests for evaluating individual models. Some of the earliest key papers on "goodness of fit" testing of continuous time models include those by Stanton (1997), Conley et al. (1997), Jiang (1998), and Jones(2003). Many specification tests for continuous models fall within one of two different categories. One category focuses on nonparametric tests. For example, tests characterized by comparing model implied transition densities with their nonparametric estimated (e.g. using kernels) counterparts see (e.g. Aït-Sahalia, 1996, 2002; Aït-Sahalia et al., 2009); and tests involving the examination of generalized cross spectra see (e.g. Chen and Hong, 2008; Hong and Li, 2005). Another category that includes papers by Gallant and Tauchen (1997), Andersen and Lund (1997), Dai and Singleton (2000), Ahn et al. (2002), Andersen et al. (2004), Thompson (2008), Aït-Sahalia and Kimmel (2007), and Corradi and Swanson (2005a), to name but a few, who use parametric methods to examine the "goodness of fit" of models. The testing approaches reviewed and used in this paper fall within this category. Namely, we review, extend and implement the simulation based test for the correct specification of a diffusion process due to Bhardwaj et al. (2008). This test is in the spirit of the conditional Kolmogorov test of Andrews (1997). In addition, we discuss a simple extension to the test of Corradi and Swanson (2011) for comparing the accuracy of predictive densities derived from (possibly misspecified) diffusion models. These tests are continuous time generalizations of 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 et al., 1999, 2001).

It should be noted that the tests used in this paper are also closely related to the interesting nonparametric specification tests of Hong (2002), Hong et al. (2004, 2007), and Chen and Hong (2008), some of which are based upon the use of the conditional characteristic function (ccf) in conjunction with the generalized cross spectrum. Our in-sample specification test is in the same spirit as these tests. Both, for example, are motivated by the classical Kolmogorov–Smirnov test, and our test along with many of their tests does not require a closed form solution for the transition density. However, our tests converge at a parametric rate while theirs converge at nonparametric rates. Moreover, our out-of-sample predictive density type model selection tests have the added feature that estimation is recursive, parameter estimation error does not vanish asymptotically and is explicitly accounted for, and multiple (possibly misspecified) models are jointly compared.

Of further note is that the difference between our approaches to in-sample (and out-of-sample) specification testing (and predictive density type model selection) and that taken elsewhere can be easily motivated within the framework used by Diebold et al. (1998a), Bai (2003), Hong (2002) and Hong et al. (2004). In their paper, DGT use the probability integral transform (see e.g. Rosenblatt, 1952) to show that $F_t(y_t|\Im_{t-1},\theta_0)$, is identically and independently distributed as a uniform random variable on [0, 1], where $F_t(\cdot|\Im_{t-1},\theta_0)$ is a parametric distribution with underlying parameter θ_0, y_t is the random variable of interest, and $\Im_{t-1}$ is the information set containing all "relevant" past information (see below for further discussion). They thus suggest using the difference between the empirical distribution of $F_t\left(y_t|\Im_{t-1},\hat{\theta}_T\right)$ and the 45° — degree line as a measure of "goodness of fit", where $\hat{\theta}_T$ is some estimator of θ_0. This approach has been shown to be very useful for financial risk management (see e.g. Diebold et al., 1999), as well as for macroeconomic forecasting (see e.g. Clements and Smith, 2000, 2002; Diebold et al., 1998b). Likewise, Bai (2003) proposes a Kolmogorov type test of $F_t(u|\Im_{t-1},\theta_0)$ based on the comparison of $F_t\left(y_t|\Im_{t-1},\hat{\theta}_T\right)$ with the CDF of a uniform on [0, 1]. As a consequence of using estimated parameters, the limiting distribution of his test reflects the contribution of parameter estimation error and is not nuisance parameter free. To overcome this problem, Bai (2003) uses a novel approach based on a martingalization argument to construct a modified Kolmogorov test which has a nuisance parameter free limiting distribution. This test has power against violations of uniformity but not against violations of independence. Two features differentiate our approach from that taken in the above papers. First, we assume strict stationarity, while they do not. Second, we allow for dynamic misspecification under the null hypothesis, while they do not. While our approach is clearly less general because of the first feature, the second feature allows us to obtain asymptotically valid critical values even when the conditioning information set does not contain all of the relevant past history. More precisely, we are interested in testing for correct specification, given a particular information set which may or may not contain all of the relevant past information. This is relevant when a Kolmogorov test is constructed, as one is generally faced with the problem of defining $\Im_{t-1}$. If enough history is not included, then there may be dynamic misspecification. Additionally, finding out how much information (e.g. how many lags) to include may involve pre-testing, hence leading to a form of sequential test bias. By allowing for dynamic misspecification, we do not require such pre-testing. Another key feature of our approach concerns the fact that the limiting distribution of Kolmogorov type tests is affected by