Journal of Econometrics 185 (2015) 60–81

Contents lists available at ScienceDirect

Journal of Econometrics

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

Estimation of affine term structure models with spanned or unspanned stochastic volatility[☆]

Drew D. Creal[a,*], Jing Cynthia Wu[a,b]

[a] Chicago Booth, United States

[b] NBER, United States

ARTICLE INFO

Article history:
Received 3 October 2013
Received in revised form
2 October 2014
Accepted 5 October 2014
Available online 4 November 2014

JEL classification:
C32
E32
E43

Keywords:
Affine term structure models
Unspanned stochastic volatility
Estimation

ABSTRACT

We develop new procedures for maximum likelihood estimation of affine term structure models with spanned or unspanned stochastic volatility. Our approach uses linear regression to reduce the dimension of the numerical optimization problem yet it produces the same estimator as maximizing the likelihood. It improves the numerical behavior of estimation by eliminating parameters from the objective function that cause problems for conventional methods. We find that spanned models capture the cross-section of yields well but not volatility while unspanned models fit volatility at the expense of fitting the cross-section.

1. Introduction

We propose new estimation procedures for affine term structure models (ATSMs) with spanned or unspanned stochastic volatility that use linear regression to simplify and stabilize estimation. For spanned models, our procedure recovers the maximum likelihood estimator but only requires numerically optimizing over a lower dimensional parameter space. The stability of our method makes it possible for us to study local maxima, explain why they exist, and their economic implications. We show how our insights from spanned models can be extended to estimate unspanned stochastic volatility (USV) models despite the fact that for USV models the likelihood function is not known in closed-form. Estimating a range of popular models, we find that models with spanned volatility fit the cross section of the yield curve better, while those with unspanned volatility fit the volatility better.

ATSMs are popular among policy makers, practitioners, and academic researchers for studying bond prices, monetary policy, and the macroeconomic determinants of discount rates; for overviews, see Piazzesi (2010), Duffee (2012), Gürkaynak and Wright (2012), and Diebold and Rudebusch (2013). As the literature on ATSMs has developed over the last decade, there is a consensus that estimation can be challenging; see, e.g. Duffee (2002), Ang and Piazzesi (2003), Kim and Orphanides (2005), and Hamilton and Wu (2012). New procedures for Gaussian ATSMs have made them easier to estimate, further increasing their popularity; see, Joslin et al. (2011), Christensen et al. (2011), Hamilton and Wu (2012), Adrian et al. (2012) and Diez de Los Rios (2013). However, these procedures do not address models with stochastic volatility. Moreover, in USV models as proposed by Collin-Dufresne and Goldstein (2002) and Collin-Dufresne et al. (2009), the likelihood function is not known in closed-form. Potential solutions to this problem are the closed-form expansions of the likelihood for continuous-time models developed by Aït-Sahalia (2008) and Aït-Sahalia and Kimmel (2010) and the expectation maximization (EM) algorithm. Combining our approach with these could

☆ We thank Yacine Ait-Sahalia, Boragan Aruoba, Michael Bauer, Alan Bester, John Cochrane, Frank Diebold, Rob Engle, Jim Hamilton, Chris Hansen, Guido Kuersteiner, Ken Singleton, two anonymous referees, and seminar and conference participants at Chicago Booth, NYU Stern, NBER Summer Institute, Maryland, Bank of Canada, Kansas, UMass, and Chicago Booth Junior Finance Symposium for helpful comments. Drew Creal thanks the William Ladany Faculty Scholar Fund at the University of Chicago Booth School of Business for financial support. Cynthia Wu also gratefully acknowledges financial support from the IBM Faculty Research Fund at the University of Chicago Booth School of Business. This paper was formerly titled "Estimation of non-Gaussian affine term structure models".

* Corresponding author.
E-mail addresses: dcreal@chicagobooth.edu (D.D. Creal), cynthia.wu@chicagobooth.edu (J.C. Wu).

http://dx.doi.org/10.1016/j.jeconom.2014.10.003
0304-4076/

potentially improve estimation; we demonstrate this for the EM algorithm explicitly.

Our main contribution are new procedures for estimating ATSMs with spanned or unspanned stochastic volatility. For models with spanned factors where volatility factors price bonds, we propose to maximize a concentrated likelihood that when optimized gives exactly the same estimator as maximizing the original likelihood function. However, it only requires numerically optimizing over a subset of the parameters. The concentrated likelihood function is simple to construct from linear regressions. Using this approach, estimation of spanned models only takes a fraction of a second to several minutes compared to hours when optimizing the original likelihood.

For USV models where the volatility factors do not price bonds, the log-likelihood function is not known in closed-form adding another layer of difficulty. Nevertheless, we show how the intuition behind the concentrated likelihood for spanned models can be extended to estimate USV models using the EM algorithm of Dempster et al. (1977). The maximization step of the EM algorithm solves a similar problem as optimizing the likelihood function of a spanned Gaussian ATSM. Consequently, we can construct a concentrated objective function for the EM algorithm using linear regressions just as we did for spanned models.

Our method outperforms conventional approaches both in terms of stability of convergence and speed. A study for a 3-factor model with one spanned volatility factor shows that our method guarantees convergence as long as it is locally identified, and it converges to a number of local maxima repeatedly. Aside from being able to find the global maximum, our method helps us to locate and understand the economic implications of different local maxima. Conversely, the conventional method of directly maximizing the original likelihood never converges fully to any of the local maxima, nor does it converge to the same point twice in repeated trials even when it is initialized under the same local mode. This makes it difficult for researchers to differentiate between points near a well-behaved local maximum having the same economic meaning and locations corresponding to local maxima that are economically different. The median time it takes for our new procedure is less than 2 minutes for this model, whereas the conventional approach takes over 2 hours.

Using our method, we shed light on how local maxima with different economic implications are created in non-Gaussian spanned models. In Gaussian models, different rotations of the factors (such as re-ordering of the factors) result in equivalent global maxima, with identical economic implications. In non-Gaussian models with spanned factors, rotations can have substantial economic impacts. The non-Gaussian state variables must be positive and enter the conditional variance. This creates an asymmetry between the Gaussian and non-Gaussian factors resulting in local maxima that are not economically equivalent.

Another contribution of this paper is to develop a family of discrete-time non-Gaussian ATSMs that encompasses continuous-time models, including both spanned models as in Duffie and Kan (1996), Duffee (2002), Cheridito et al. (2007), and Aït-Sahalia and Kimmel (2010) as well as USV models as proposed by Collin-Dufresne and Goldstein (2002). Gouriéroux et al. (2002) proposed a one factor discrete-time non-Gaussian model and Le et al. (2010) generalized it to have multiple factors. Our model encompasses any admissible rotation of a multivariate discrete-time Cox et al. (1985) process, allowing the factors to be correlated. The model nests the risk-neutral dynamics of other discrete-time ATSMs.[1] In our model, the physical and risk neutral dynamics follow the same stochastic process but with different parameters. The market prices of risk have the extended affine form of Cheridito et al. (2007), which is different than Le et al. (2010). Finally, we also provide the restrictions needed to generate USV in discrete-time versions of the continuous-time models studied by Collin-Dufresne et al. (2009) and Joslin (2010).

We apply our estimation method to a range of popular spanned and unspanned models with three and four factors. Judging by the estimated likelihood, a model with three spanned non-Gaussian factors has the highest likelihood followed by one of the USV models. Gaussian and non-Gaussian models with spanned factors fit the cross-section of yields equally well. However, spanned models do not capture the volatility well at any maturity, even for the best fitting model. This is because the non-Gaussian state variables must simultaneously fit the conditional mean and variance. Maximum likelihood places more weight on the first moment. In order to guarantee unspanned volatility factors, USV models place restrictions on the bond loadings. This causes USV models to sacrifice some cross-sectional fit; their pricing errors are larger than spanned models. On the other hand, USV models fit the dynamics of yield curve volatility well. The USV restrictions are not unique and we show that the choice of which USV restrictions are imposed is not inconsequential.

This paper continues as follows. In Section 2, we specify a general class of discrete-time, non-Gaussian affine term structure models. In Section 3, we describe our new approach to estimation for both spanned and unspanned models. Section 4 describes the data and parameter restrictions of the models. Section 5 studies a three factor spanned model in depth. In Section 6, we study eight three and four factor spanned and unspanned models. In Section 7, we discuss directions for future research and conclude.

2. Model

In this section, we describe a class of discrete-time ATSMs with stochastic volatility that encompass both spanned models, as in Duffie and Kan (1996), Dai and Singleton (2000), Cheridito et al. (2007); and unspanned models, as proposed by Collin-Dufresne and Goldstein (2002).

2.1. Bond prices

The model has a $G \times 1$ vector of conditionally Gaussian state variables g_t, whose volatilities are captured by an $H \times 1$ vector of positive state variables h_t. Under the risk-neutral measure $\mathbb{Q}$, the Gaussian state variables follow a vector autoregression with conditional heteroskedasticity

$$g_{t+1} = \mu_g^{\mathbb{Q}} + \Phi_g^{\mathbb{Q}} g_t + \Phi_{gh}^{\mathbb{Q}} h_t + \Sigma_{gh}\varepsilon_{h,t+1}^{\mathbb{Q}} + \varepsilon_{g,t+1}^{\mathbb{Q}},$$
$$\varepsilon_{g,t+1}^{\mathbb{Q}} \overset{\mathbb{Q}}{\sim} N\left(0, \Sigma_{g,t}\Sigma_{g,t}'\right), \tag{1}$$
$$\Sigma_{g,t}\Sigma_{g,t}' = \Sigma_{0,g}\Sigma_{0,g}' + \sum_{i=1}^{H} \Sigma_{i,g}\Sigma_{i,g}' h_{it},$$
$$\varepsilon_{h,t+1}^{\mathbb{Q}} = h_{t+1} - \mathbb{E}^{\mathbb{Q}}\left(h_{t+1} | \mathcal{I}_t\right)$$

where $\mathcal{I}_t$ captures agents' information set at time t.

The volatility factors h_t are an affine transformation of the exact discrete-time equivalent of a multivariate Cox et al. (1985) process

$$h_{t+1} = \mu_h + \Sigma_h w_{t+1} \tag{2}$$

$$w_{i,t+1} \sim \text{Gamma}\left(\nu_{h,i}^{\mathbb{Q}} + z_{i,t+1}^{\mathbb{Q}}, 1\right), \quad i = 1, \ldots, H \tag{3}$$

$$z_{i,t+1}^{\mathbb{Q}} \sim \text{Poisson}\left(e_i'\Sigma_h^{-1}\Phi_h^{\mathbb{Q}}\Sigma_h w_t\right), \quad i = 1, \ldots, H \tag{4}$$

where e_i denotes the ith column of the identity matrix I_H. We discuss the admissibility restrictions and interpretation of the parameters of the model in Section 2.2.

[1] In this paper, we do not consider the class of non-Gaussian ATSMs built from the non-central Wishart process of Gouriéroux et al. (2009).