Journal of Financial Economics 91 (2009) 288–318

Contents lists available at ScienceDirect

Journal of Financial Economics

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

Conditional volatility in affine term-structure models: Evidence from Treasury and swap markets ☆

Kris Jacobs [a,b,*], Lotfi Karoui [c]

[a] McGill University, Montreal, Quebec H3A 1G5, Canada
[b] Tilburg University, 5000 LE Tilburg, The Netherlands
[c] Goldman, Sachs & Co., New York, NY 10004, USA

ARTICLE INFO

Article history:
Received 2 August 2006
Received in revised form
1 August 2007
Accepted 17 February 2008
Available online 27 November 2008

JEL classification:
G12

Keywords:
Term-structure model
Affine
Interest rate swap
Treasury market
Conditional volatility
Time series
Cross-section
EGARCH

ABSTRACT

We study the ability of three-factor affine term-structure models to extract conditional volatility using interest rate swap yields for 1991–2005 and Treasury yields for 1970–2003. For the Treasury sample, the correlation between model-implied and EGARCH volatility is between 60% and 75%. For the swap sample, this correlation is rather low or negative. We find that these differences in model performance are primarily due to the timing of the swap sample, and not to institutional differences between swap and Treasury markets. We conclude that the ability of multifactor affine models to extract conditional volatility depends on the sample period, but that overall these models perform better than has been argued in the literature.

1. Introduction

The class of multifactor affine term-structure models (ATSMs) has emerged as the workhorse in the fixed income literature, and a consensus has emerged in the literature that three-factor ATSMs are needed to successfully capture certain stylized facts of the term structure of interest rates.[1] Empirical implementations of these models often find that the term structure can be characterized in terms of the interest rate level, the slope of the term structure, and term-structure curvature. However, recently a number of papers have questioned the ability of multifactor ATSMs, and of three-factor ATSMs in particular, to capture some important aspects of term-structure

☆ Kris Jacobs would like to thank FQRSC, IFM^2, and SSHRC for financial support. Lotfi Karoui was supported by a grant from IFM^2 while he was a Ph.D. student at McGill University. The views expressed here are those of the authors only and not necessarily of the institutions with which they are affiliated. Without implication we would like to thank Caio Almeida, Mike Chernov, Peter Christoffersen, Pierre Collin-Dufresne, Jin Duan, Jan Ericsson, Scott Joslin, Albert Lee Chun, an anonymous referee, and participants in the 2006 NFA Meeting and the 2007 FDIC Annual Derivatives and Risk Management Conference for helpful comments.

* Corresponding author at: McGill University, Montreal, Quebec H3A 1G5, Canada.
E-mail address: kris.jacobs@mcgill.ca (K. Jacobs).

[1] See Chen and Scott (1993) and Balduzzi, Das, Foresi, and Sundaram (1996) for early multifactor models. See Duffie and Kan (1996) for a full characterization and Dai and Singleton (2000) for a useful classification of multifactor ATSMs. Litterman and Scheinkman (1991) establish that three-factor models are able to explain a large part of the variation in bond yields. See Duffee (2002) and Duarte (2004) for recent implementations of three-factor ATSMs.

0304-405X/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jfineco.2008.02.006

dynamics. Part of this recent criticism has been directed at the ability of multifactor ATSMs to model volatility. While these models are able to generate various patterns for unconditional volatility in both swap and Treasury markets (see Dai and Singleton, 2000, 2003), several studies have questioned their ability to model conditional volatility. Using swap data, Collin-Dufresne, Goldstein, and Jones (2006) find that a popular and well-documented three-factor affine model implies volatility paths that are negatively correlated with the GARCH volatility estimates of weekly changes in the six-month rate. Andersen and Benzoni (2005) use intraday Treasury data to show that realized yield volatility is unrelated to principal components extracted from the cross-section, which proxy for model-implied volatility.[2]

These findings question some of the most important building blocks of ATSMs, and more fundamentally question the validity of a large class of arbitrage based term-structure models. It is a key implication of these models that the yields' conditional volatility is a linear combination of the state variables. For example, in the model used by Collin-Dufresne, Goldstein, and Jones (2006), in which volatility is driven by a single state variable, the failure of the model can be explained by the fact that the volatility factor will proxy for the level of the yield curve, which may not be highly correlated with the time series volatility of the six-month rate. This feature is often referred to as unspanned stochastic volatility (USV), and it reflects a tension between the time series and the cross-sectional properties of the model.

These empirical findings have far-reaching practical implications, because if the yield curve fails to span volatility, fixed-income volatility risk cannot be hedged by positions in the bond market alone. Consequently, any term-structure model that relates the conditional volatility to the cross-section of yields will fail to capture the time variability of the conditional volatility. To resolve these problems, Collin-Dufresne and Goldstein (2002) and Collin-Dufresne, Goldstein, and Jones (2006) propose a new family of affine models, labeled USV models. However, there is considerable disagreement about the empirical performance of these models (see, for example, Bikbov and Chernov, 2004; Collin-Dufresne, Goldstein, and Jones, 2006; Thompson, 2004).

This paper examines the ability of three-factor ATSMs to simultaneously match the cross-sectional and time series properties of yield volatility. We empirically investigate Treasury yields as well as swap yields. Because reconciling the time series and cross-sectional properties of the model critically depends on the mapping between the physical and risk-neutral model dynamics, we pay particular attention to the market price of risk, and investigate three classes of models: completely affine models, essentially affine models (Duffee, 2002), and extended affine models (Cheridito, Filipović, and Kimmel, 2007). We follow the popular classification by Dai and Singleton (2000), which is very appropriate for volatility modeling. Rather than using the volatility factor as a proxy for the conditional volatility of the short rate, we regress the exact model-implied conditional volatility for yields of different maturities on EGARCH volatility estimates.

For a long sample of monthly Treasury yields from 1970 to 2003, we find a sizeable positive correlation of between 60% and 75% between model-implied conditional volatility and EGARCH estimates of the volatility of yield differences. The specification of the price of risk seems inconsequential. We provide more insight for these findings by showing a large and statistically significant correlation between EGARCH volatility estimates and the level factor.

For swap yields from 1991 to 2005, our findings are very different: correlations are generally negative for long maturities, and positive but rather small for short maturities. Results for swap yields are also less robust, more model-dependent, and less reliable. The differences in model performance between the main Treasury and swap samples are confirmed using other metrics: the slopes of the forecasting regressions indicate more support for the model-implied volatility measures when estimating on Treasury yields, and the relative root mean squared errors (RMSEs) for yields are larger for the swap sample. These findings can be partly explained by the fact that the correlation between EGARCH volatility and the level factor, as well as between model-implied volatility and the level factor, is much weaker in the swap sample.

We investigate the robustness of our findings with respect to the use of a model-free benchmark by using different GARCH models, and by using instantaneous conditional volatility as an alternative to GARCH models, and we find similar results. When we use realized volatility instead of EGARCH volatility as a measure of model-free volatility in the Treasury market, the results change somewhat, but the correlation between model-free and model-implied volatility is still very high. Estimation results obtained using Kalman filtering are also very similar. We repeat the analysis using yield levels instead of differences, and the results are consistent, with somewhat more favorable results for Treasury yields. Finally, our findings for Treasury securities are also robust with respect to the use of interpolation technique.

We next investigate whether the differences in results are due to the sample periods. Results for monthly and weekly Treasury yields for 1991–2003 are more similar to the results for 1991–2005 swap data, in the sense that ATSMs do not perform very well in extracting conditional volatility. Results for monthly Treasury yields for 1952–2003 are very similar to the results for the 1970–2003 sample. Moreover, when we split up the 1970–2003 Treasury data into pre- and post-monetary experiment samples, the models continue to perform rather well. We therefore conclude that the different results for swap and Treasury data are primarily due to the characteristics of the post-1991 data. However, it is not straightforward to determine which feature of the data is causing the models' poor performance during this period.

[2] Collin-Dufresne and Goldstein (2002) show that the conditional volatility implied by term-structure models is unrelated to implied volatility from interest rate options. See also Fan, Gupta, and Ritchken (2003), Jagannathan, Kaplin, and Sun (2003), and Li and Zhao (2006).