International Journal of Forecasting 34 (2018) 733–747

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting risk with Markov-switching GARCH models: A large-scale performance study

David Ardia[a,b,*], Keven Bluteau[a,c], Kris Boudt[c,d], Leopoldo Catania[e]

[a] Institute of Financial Analysis, University of Neuchâtel, Neuchâtel, Switzerland
[b] Department of Finance, Insurance and Real Estate, Laval University, Québec City, Canada
[c] Solvay Business School, Vrije Universiteit Brussel, Belgium
[d] School of Business and Economics, Vrije Universiteit Amsterdam, The Netherlands
[e] Department of Economics and Business Economics, Aarhus University and CREATES, Denmark

ARTICLE INFO

Keywords:
GARCH
MSGARCH
Forecasting performance
Large-scale study
Value-at-risk
Expected shortfall
Risk management

ABSTRACT

We perform a large-scale empirical study in order to compare the forecasting performances of single-regime and Markov-switching GARCH (MSGARCH) models from a risk management perspective. We find that MSGARCH models yield more accurate Value-at-Risk, expected shortfall, and left-tail distribution forecasts than their single-regime counterparts for daily, weekly, and ten-day equity log-returns. Also, our results indicate that accounting for parameter uncertainty improves the left-tail predictions, independently of the inclusion of the Markov-switching mechanism.

1. Introduction

Under the regulation of the Basel Accords, risk managers of financial institutions must make use of state-of-the-art methodologies for monitoring financial risks (Board of Governors of the Federal Reserve Systems, 2012). Clearly, regime-switching time-varying volatility models and Bayesian estimation methods can be considered strong candidates for being classified as state-of-the-art methodologies. However, many academics and practitioners also consider the single-regime volatility model and the use of frequentist estimation via maximum likelihood (ML) as state-of-the-art. Risk managers disagree as to whether the computational complexity of a regime-switching model and the Bayesian estimation method pay off in terms of a higher accuracy of their financial risk monitoring system. We study this question in the context of monitoring the individual risks of a large number of financial assets.

The specification of the conditional volatility process is key among the various building-blocks of any risk management system, especially for short horizons (McNeil, Frey, & Embrechts, 2015). Research on the use of time series models for modeling the volatility has proliferated since the creation of the original ARCH model by Engle (1982) and its generalization by Bollerslev (1986), and multiple extensions of the GARCH scedastic function have been proposed for capturing additional stylized facts that are observed in financial markets, such as nonlinearities, asymmetries, and long-memory properties; see Engle (2004) for a review. These so-called GARCH-type models are essential tools for risk managers today.

An appropriate risk model should be able to accommodate the properties of financial returns. Recent academic studies have shown that many financial assets exhibit structural breaks in their volatility dynamics, and that ignoring this feature can have a big effect on the precision

* Correspondence to: University of Neuchâtel, Rue A.-L. Breguet 2, CH 2000, Neuchâtel, Switzerland.
E-mail addresses: david.ardia@unine.ch (D. Ardia), keven.bluteau@unine.ch (K. Bluteau), kris.boudt@vub.be (K. Boudt), leopoldo.catania@econ.au.dk (L. Catania).

https://doi.org/10.1016/j.ijforecast.2018.05.004

of volatility forecasts (see e.g. Bauwens, Dufays, & Rombouts, 2014; Lamoureux & Lastrapes, 1990). As was noted by Danielsson (2011), this shortcoming in individual forecasting systems can have systemic consequences. Indeed, he refers to these single-regime volatility models as one of the culprits of the great financial crisis: *"(...) the stochastic process governing market prices is very different during times of stress compared to normal times. We need different models during crisis and non-crisis and need to be careful in drawing conclusions from non-crisis data about what happens in crises and vice versa".*

One way to address the *switch* in the return process is provided by Markov-switching GARCH models (MSGARCH), whose parameters can change over time according to a discrete latent (i.e., unobservable) variable. These models can adapt quickly to variations in the unconditional volatility level, which improves risk predictions (see e.g. Ardia, 2008; Marcucci, 2005).

The initial studies of Markov-switching autoregressive heteroscedastic models applied to financial time series focused on ARCH specifications, and thus omitted a lagged value of the conditional variance in the variance equation (Cai, 1994; Hamilton & Susmel, 1994). The use of ARCH rather than GARCH dynamics leads to computational tractability in the likelihood calculation. Indeed, Gray (1996) shows that, given a Markov chain with K regimes and T observations, the evaluation of the likelihood of a Markov-switching model with general GARCH dynamics requires integration over all K^T possible paths, rendering the estimation infeasible. While this difficulty is not present in ARCH specifications, the use of lower-order GARCH models tends to offer a more parsimonious representation than higher-order ARCH models.

Dueker (1997), Gray (1996) and Klaassen (2002) tackle the *path-dependence problem* of MSGARCH through approximation, by collapsing the past regime-specific conditional variances based on ad hoc schemes. A further solution is to consider alternatives to traditional maximum likelihood estimation. Bauwens et al. (2014) recommended the use of Bayesian estimation methods that are still feasible through so-called data augmentation and particle MCMC techniques. Augustyniak (2014) relied on a Monte Carlo EM algorithm with importance sampling. In our study, we consider the alternative approach provided by Haas, Mittnik, and Paolella (2004), who let the GARCH process of each state evolves independently of those in the other states. In addition to avoiding the path-dependence problem that arises with traditional maximum likelihood estimation, their model also allows for a clear-cut interpretation of the variance dynamics in each regime.

The first contribution of our paper is to test whether MSGARCH models do indeed provide risk managers with useful tools that can improve their volatility forecasts.[1] We answer this question by performing a large-scale empirical analysis in which we compare the risk forecasting performances of single-regime and Markov-switching GARCH models. We take the perspective of a risk manager who is working for a fund manager and conduct our study on the daily, weekly and ten-day log-returns of a large universe of stocks, equity indices, and foreign exchange rates. Thus, in contrast to Hansen and Lunde (2005), who compare a large number of GARCH-type models on a few series, we focus on a few GARCH and MSGARCH models and a large number of series. For single-regime and Markov-switching specifications, the scedastic specifications that we consider account for different reactions of the conditional volatility to past asset returns. More precisely, we consider both the symmetric GARCH model (Bollerslev, 1986) and the asymmetric GJR model (Glosten, Jagannathan, & Runkle, 1993). These scedastic specifications are integrated into the MSGARCH framework using the approach of Haas et al. (2004). For the (regime-dependent) conditional distributions, we use the symmetric and Fernández and Steel (1998) skewed versions of the normal and Student-t distributions. This leads to a total of sixteen models.

Our second contribution is to test the impact of the estimation method on the performance of the volatility forecasting model. Traditionally, GARCH and MSGARCH models are estimated using a frequentist (typically via ML) approach; see Augustyniak (2014), Haas et al. (2004) and Marcucci (2005). However, several recent studies have argued that a Bayesian approach offers some advantages. For instance, Markov chain Monte Carlo (MCMC) procedures can explore the joint posterior distribution of the model parameters, and parameter uncertainty is integrated into the risk forecasts naturally via the predictive distribution (Ardia, 2008; Ardia, Kolly, & Trottier, 2017; Bauwens, De Backer, & Dufays, 2014; Bauwens, Preminger, & Rombouts, 2010; Geweke & Amisano, 2010).

Combining the sixteen model specifications using the frequentist and Bayesian estimation methods, we obtain 32 possible candidates for a state-of-the-art methodology for monitoring the financial risk. We use an out-of-sample evaluation period of 2,000 days, from (approximately) 2005 to 2016, consisting of daily log-returns. We evaluate the accuracy of the risk prediction models in terms of the Value-at-Risk (VaR), the expected shortfall (ES), and the left-tail (i.e., losses) of the conditional distribution of the assets' returns.

Our empirical results suggest a number of practical insights, which can be summarized as follows. First, we

[1] Our study focuses exclusively on GARCH and MSGARCH models. GARCH is the workhorse model in financial econometrics and has been being investigated for decades. It is used widely by practitioners and academics; see for instance Bams, Blanchard, and Lehnert (2017) and Herwartz (2017). MSGARCH is the most natural and straightforward extension to GARCH. Alternative conditional volatility models include stochastic volatility models (Jacquier, Polson, & Rossi, 1994; Taylor, 1994), realized measure-based conditional volatility models such as HEAVY (Shephard & Sheppard, 2010) or realized GARCH (Hansen, Huang, & Shek, 2011), or even combinations of these (Opschoor, van Dijk, & van der Wel, 2017). Finally, note that our study considers only the (1,1)-lag specification for the GARCH and MSGARCH models. While considering higher orders for (MS)GARCH model specifications has a clear computational cost, the payoff in terms of improvements in forecasting precision may be low. In fact, several studies have shown that increasing the orders does not lead to any substantial improvement in the forecasting performance in the case of predicting the conditional variance of asset returns (see e.g. Hansen & Lunde, 2005). We tested whether this result also holds for our sample and investigated the fits of GARCH(p, q) and GJR(p, 1, q) models over the three universes of stocks, indices and foreign exchange rates, for rolling windows of 1500 points, and selected the best in-sample model via the BIC. We found that the (1,1) specification is selected in the vast majority of cases.