International Journal of Forecasting 30 (2014) 449–463

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting return volatility: Level shifts with varying jump probability and mean reversion

Jiawen Xu [a,b,*], Pierre Perron [c]

[a] School of Economics, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, 200433, China

[b] Key Laboratory of Mathematical Economics (SUFE), Ministry of Education, Shanghai, 200433, China

[c] Department of Economics, Boston University, 270 Bay State Rd., Boston MA 02215, United States

ARTICLE INFO

Keywords:
Structural change
State space model
Regime switching
Long-memory

ABSTRACT

We extend the random level shift (RLS) model of Lu and Perron (2010) to the volatility of asset prices, which consists of a short memory process and a random level shift component. Motivated by empirical features, (a) we specify a time-varying probability of shifts as a function of large negative lagged returns; and (b) we incorporate a mean reverting mechanism so that the sign and magnitude of the jump component change according to the deviations of past jumps from their long run mean. This allows the possibility of forecasting the sign and magnitude of the jumps. We estimate the model using twelve different series, and compare its forecasting performance with those of a variety of competing models at various horizons. A striking feature is that the modified RLS model has the smallest mean square forecast errors in 64 of the 72 cases, while it is a close second for the other 8 cases. The improvement in forecast accuracy is often substantial, especially for medium- to long-horizon forecasts. This is strong evidence that our modified RLS model offers important gains in forecasting performance.

1. Introduction

Recently, there has been an upsurge of interest in the possibility of confusing long-memory with structural changes in levels. This idea extends that expounded by Perron (1989, 1990), who showed that structural change and unit roots are easily confused. When a stationary process is contaminated by structural changes in the mean, the estimate of the sum of its autoregressive coefficients is biased towards one, and tests of the null hypothesis of a unit root are biased toward non-rejection. This phenomenon has been shown to apply to the long-memory context as well. That is, when a stationary short-memory process is contaminated by structural changes in level, the estimate of the long-memory parameter is biased away from zero and the autocovariance function of the process exhibits a slow rate of decay. Relevant references on this issue include the studies by Diebold and Inoue (2001), Engle and Smith (1999), Gourieroux and Jasiak (2001), Granger and Ding (1996), Granger and Hyung (2004), Lobato and Savin (1998), Mikosch and Stărică (2004), Parke (1999), and Teverovsky and Taqqu (1997).

The literature on modeling and forecasting the stock return volatility is voluminous. Two approaches that have proven useful are the GARCH and stochastic volatility (SV) models. In their standard forms, the ensuing volatility processes are stationary and weakly dependent, with autocorrelations that decrease exponentially. This in is contrast to the empirical findings obtained using various proxies for volatility (e.g., daily absolute returns) which indicate autocorrelations that decay very slowly at long lags. In light of this, several long-memory models have been proposed. For example, Baillie, Bollerslev, and Mikkelsen (1996) and Bollerslev and Mikkelsen (1996) considered

* Corresponding author at: School of Economics, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, 200433, China.
E-mail addresses: xu.jiawen@mail.shufe.edu.cn (J. Xu), perron@bu.edu (P. Perron).

0169-2070/$ – see front matter © 2014 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijforecast.2013.12.012

fractionally integrated GARCH and EGARCH models, while Breidt, Crato, and de Lima (1998) and Harvey (1998) proposed long memory SV (LSV) models, where the log of volatility is modeled as a fractionally integrated process.

More recently, attempts have been made to distinguish between short-memory stationary processes plus level shifts and long-memory models; see, in particular, Granger and Hyung (2004). They documented the fact that, when breaks determined via some pre-tests are accounted for, the evidence for long-memory is weaker. However, this evidence is inconclusive, since structural change tests are severely biased in the presence of long-memory, and log periodogram estimates of the memory parameter are biased downward when sample-selected breaks are introduced. This is an overfitting problem that Granger and Hyung (2004, p. 416) recognized clearly. Stărică and Granger (2005) presented evidence that the log-absolute returns of the S&P 500 index is a white noise series which is affected by occasional shifts in the unconditional variance, and showed that this specification has a better forecasting performance than the more traditional GARCH(1, 1) model and its fractionally integrated counterpart. Mikosch and Stărică (2004) considered the autocorrelation function of the absolute returns of the S&P 500 index for the period 1953–1977. They documented the fact that, for the full period, it resembles that of a long-memory process. However, interestingly, if one omits the last four years of data, the autocorrelation function is very different and looks like one associated with a short-memory process. They explain this finding by arguing that the volatility of the S&P 500 returns increased over the period 1973–1977. Morana and Beltratti (2004) also argue that breaks in the level of volatility partially explain the long-memory features of some exchange rate series. Perron and Qu (2007) analyzed the time and spectral domain properties of a stationary short memory process affected by random level shifts. Perron and Qu (2010) showed that, when applied to daily S&P 500 log absolute returns over the period 1928–2002, the level shift model explains both the shape of the autocorrelations and the path of log periodogram estimates as a function of the number of frequency ordinates used. Qu and Perron (2013) estimated a stochastic volatility model with level shifts by adopting a Bayesian approach using daily data on returns from the S&P 500 and NASDAQ indices over the period 1980.1–2005.12. They showed that the level shifts account for most of the variation in volatility, that their model provides a better in-sample fit than alternative models, and that its forecasting performance is better than standard short or long-memory models without level shifts for the NASDAQ, and just as good for the S&P 500.

Lu and Perron (2010) extended the work of Stărică and Granger (2005) by estimating a structural model directly. They adopted a specification for which the series of interest is the sum of a short-memory process and a jump or level shift component. For the latter, they specified a simple mixture model such that the component is the cumulative sum of a process that is 0 with some probability $(1-\alpha)$, and is a random variable with probability α. To estimate such a model, they transformed it into a linear state space form with innovations having a mixture of two normal distributions, and adopted an algorithm similar to the one used by Perron and Wada (2009) and Wada and Perron (2007). They restricted the variance of one of the two normal distributions to be zero, allowing a simple but efficient algorithm.

Varneskov and Perron (2013) extended the random level shift model further by combining it with a long memory process, modeled as a $ARFIMA(p, d, q)$ process. They provided a forecasting framework for a class of long-memory models with level shifts. Their forecasting experiments using six different data series covering both low frequency and high frequency data showed that the RLS-ARFIMA model outperforms other competing models.

This paper extends that of Lu and Perron (2010) in several directions. First, we let the jump probability depend on some covariates. This allows a more comprehensive and realistic probabilistic structure for the level shift model. The specification adopted is in the spirit of the "news impact curve", as suggested by Engle and Ng (1993). We model the probability of a shift as a function of the occurrence and magnitude of large negative lagged returns. The second modification is to incorporate a mean reverting mechanism in the level shift model, so that the sign and magnitude of the jump component change according to the deviations of past jumps from their long run mean. Apart from being a device that allows a better in-sample description, its advantage is that the sign and magnitude of the jumps can be predicted to some extent. As we will show, this allows much improved forecasts.

We apply the modified level shift model to the following daily return series using absolute returns as a proxy for volatility and a logarithmic transformation in order to have a series which is closer to being normally distributed and also not bounded below by zero: S&P 500 stock market index, Dow Jones Industrial Average (DJIA) index, AMEX index, Nasdaq index, Nikkei 225 index, IBM stock prices, Crude Oil prices, Treasury Bond Futures, and the Trade Weighted US Dollar Index. To assess the sensitivity of our results, we also consider three realized volatility series, also in logarithmic form, constructed from 5-min returns on the S&P 500 and Treasury Bond Futures, as well as a realized volatility series constructed from tick-by-tick trades on the SPY, an exchange-traded fund that tracks the S&P 500. Our point estimate for the average probability of shifts is similar to that of the original model, still a quite small number; but the weight on extreme past negative returns is large enough to result in a significant increase in jump probability when past stock returns are taken into account, thereby inducing a clustering property for the jumps. Also, the estimates indicate that a mean reverting mechanism is present, which changes the sign of the jump. When the past jump component deviates from the long run mean by a large amount, it is brought back towards the long-run mean.

We compare the forecasting performance of our model with those of eight competing models: the original random level shift model (RLS), the popular $ARFIMA(1, d, 1)$ and $ARFIMA(0, d, 0)$ models, a GARCH(1, 1), a fractionally integrated GARCH model (FIGARCH(1, d, 1)), the HAR model, a Multiple Regime Smooth Transition Heterogeneous Autoregressive Model (HARST) and a Markov Regime Switching model. We consider forecast horizons of 1, 5, 10, 20,