International Journal of Forecasting 33 (2017) 770–785

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Realised variance forecasting under Box-Cox transformations

Nick Taylor

School of Economics, Finance and Management, University of Bristol, Bristol, BS8 1TN, UK

ARTICLE INFO

Keywords:
Volatility
Risk
Forecasting competitions
Loss function
Reality check

ABSTRACT

This paper assesses the benefits of modeling Box-Cox transformed realised variance data. In particular, it examines the quality of realised variance forecasts with and without this transformation applied in an out-of-sample forecasting competition. Using various realised variance measures, data transformations, volatility models and assessment methods, and controlling for data mining issues, the results indicate that data transformations can be economically and statistically significant. Moreover, the quartic root transformation appears to be the most effective in this regard. The conditions under which the use of transformed data is effective are identified.

1. Introduction

Power transformations, and Box-Cox (BC) transformations more generally, have long been recognised as an effective way of obtaining well-specified models with symmetric errors and stable error variances; see Box and Cox (1964) and Tukey (1957). More recently, the literature has focused on assessing the out-of-sample performances of time series models applied to BC transformed data. For instance, Bårdsen and Lütkepohl (2011), Lütkepohl and Xu (2012) Mayr and Ulbricht (2015) and Proietti and Lütkepohl (2013) demonstrate that the out-of-sample forecasts from models that use BC transformed macroeconomic series can be more accurate than those from models that use the *original* (non-transformed) series (cf. Nelson & Granger, 1979). Inspired by these results, we consider whether BC transformations are useful within the context of forecasting future realised variances.

The use of transformations in the context of the realised variance is not new. Indeed, the application of models to the log transformed realised variance is common practice; see, e.g., Andersen, Bollerslev, and Diebold (2007); Andersen, Bollerslev, Diebold, and Labys (2003), Corsi (2009), Hansen, Huang, and Shek (2012) and Koopman and Scharth (2013).[1] More recently, BC transformations have been considered in this context; see, e.g., Gonçalves and Meddahi (2011), Nugroho and Morimoto (2016), Weigand (2014) and Zheng and Song (2014).[2] We add to this body of literature by examining the relative out-of-sample performances of a range of contemporary models applied to various BC transformed (and original) realised variance measures. In doing so, we look at previously considered transformation parameters (the square root and log transformations), those that have not been used widely to date (the quartic root transformation), and those based on the nature of the data used (that is, an estimated transformation parameter).

The studies conducted by Weigand (2014) and Zheng and Song (2014) are the most similar to the current paper, in that both consider the out-of-sample costs/benefits of

E-mail address: nick.taylor@bristol.ac.uk.

[1] The use of the log transformed realised variance is based on previous findings that have shown these data to be Gaussian distributed; see, e.g., Andersen, Bollerslev, Diebold, and Ebens (2001) and Andersen, Bollerslev, Diebold, and Labys (2001).

[2] Transformations are not always applied. For instance, Bollerslev, Patton, and Quaedvlieg (2016) augmented the popular long-memory heterogenous autoregressive (HAR) model of Corsi (2009), but decided not to apply the log transformation as per the Corsi-proposed HAR model.

http://dx.doi.org/10.1016/j.ijforecast.2017.04.001
0169-2070/

applying BC transformations within the context of volatility models. In the former study, Weigand (2014) proposes two BC transformed models of the multivariate realised variance (viz. the 'matrix Box-Cox model of realized covariances' and the 'Box-Cox dynamic correlation model'). In the latter study, the framework builds on the stochastic volatility model proposed by Koopman and Scharth (2013), such that BC transformed realised variance is a linear function of the unobserved underlying volatility. Both of these studies indicate that BC transformations (similar to the log transformation) are beneficial in this context. We complement and build on these studies in three ways. First, we consider a wide range of univariate realised volatility models, all of which are popular and/or have been proposed only recently. Second, we conduct hypothesis tests that examine not only the relative performances of BC transformed models and the original model, but also the relative performances of models with different transformations (for instance, log versus quartic root transformations). Third, we analyse whether the relative performance is uniform over different realised variance measures, both within a particular market and across different market indices. Furthermore, the probable determinants of this variation are investigated.

The answer to the question of whether the original realised variance measure should be transformed depends on the loss function used to assess the forecasting performance. For instance, if one uses the mean square log (MS-log) error loss function (given by the mean of the squared difference between the log forecast and the log realised value), then modelling the log transformed series will deliver the best results, as the model parameters are optimised with respect to the same loss function that is used to assess the performance.[3] Thus, we avoid favouring a particular BC transformation in this way by following the extant literature and considering the mean square (MS) and quasi-likelihood (QLIK) error loss functions applied to the original realised variance measure. These belong to the Bregman loss function family; see Banerjee, Guo, and Wang (2005), Gneiting (2011) and Patton (2015) for further details.[4] Under these loss functions, the optimal forecast is obtained by minimising any Bregman loss function when using the original data.

However, it is quite possible that the models themselves may not be 'suited' to the original (possibly highly non-Gaussian) data. This leads to the possibility that models applied to transformed data may be superior because they match the true data generating process more closely. For instance, there is considerable evidence of increased (long memory) persistence (and hence predictability) under the log transformation assumption; see Ding, Granger, and Engle (1993), Granger and Ding (1996) and Proietti (2016). As a consequence, the forecasts of the BC transformed model may be superior because of the suitability of the model to the BC transformed data, even though their parameters are not optimised with respect to the loss function that is used to assess the performance. It is this trade-off (parameter optimisation versus model suitability) that we are examining in the current paper by considering whether or not to transform realised variance data.

We use a comprehensive set of realised variance measures to examine whether the use of BC transformations is of value for forecasters. The results indicate that such transformations can improve the forecasts of the future realised variance across a range of models and under both the MS and QLIK loss functions. Moreover, the quality differences between forecasts based on modeling the original and transformed data can be significant even after controlling for data mining by using the reality check statistical tests proposed by White (2000). Of the BC transformations that we consider, it is the quartic root (not the log) transformation that delivers the best results. Finally, we demonstrate that the benefits of BC transformation are not spread evenly over the realised variance measures. Indeed, for some measures no benefits are found, a result that we demonstrate to be driven by the degree of skewness in the original realised variance measure.

The rest of the paper is organised as follows. The next section contains a description of the methodologies employed, and is followed by the empirical results. The final section concludes.

2. Methodologies

This section describes the models and methods used to constructed forecasts, and the means by which the relative forecast quality is assessed.

2.1. Forecast construction: the problem

Let x_t be the original data that we wish to forecast, in our case the realised variance, $x_t > 0$ and $t = 1, 2, \ldots, T$. As x_t is likely to be highly non-Gaussian, we model the BC transformed data, given by

$$y_t = f(x_t; \lambda) = \begin{cases} \dfrac{x_t^{\lambda} - 1}{\lambda}, & \lambda \neq 0, \\ \ln x_t, & \lambda = 0. \end{cases} \quad (1)$$

It follows that $x_t = g(y_t; \lambda) = f^{-1}(y_t; \lambda)$.[5] Suppose that the forecaster models y_t and obtains h-step-ahead forecasts given by the conditional mean of y_{t+h}; that is, $\mathrm{E}[y_{t+h}|\mathcal{F}_t]$, where $\mathcal{F}_t$ is the forecaster's information set. Moreover, suppose that we require the conditional mean of x_{t+h}, that is, $\mathrm{E}[x_{t+h}|\mathcal{F}_t]$, or equivalently $\mathrm{E}[g(y_{t+h}; \lambda)|\mathcal{F}_t]$.[6]

2.2. Forecast construction: the solution(s)

One obvious solution would be to take $g(\mathrm{E}[y_{t+h}|\mathcal{F}_t]; \lambda)$, referred to henceforth as the *naïve adjustment* forecast.

[3] This implicitly assumes that the parameters are estimated by minimising the sum of squared errors.

[4] The Bregman loss function family possess the quality that the use of the MS loss function for ranking the forecast quality leads to a consistent ranking over all members of the Bregman loss function family, for correctly specified models with nested information sets (Patton, 2015).

[5] Note that y_t represents the original realised variance measure when $\lambda = 1$.

[6] Under the Bregman loss function assumption, the conditional mean is the optimal forecast (Banerjee et al., 2005; Gneiting, 2011; Patton, 2015).