Estimating Stochastic Volatility: The Rough Side to Equity Returns

Abstract

This Project evaluates the forecasting performance of a Brownian Semi-Stationary (BSS) process in modelling the volatility of 21 equity indices. We implement a sophisticated Hybrid Scheme to simulate BSS processes with high efficiency and precision. These simulations are useful to price derivatives, accounting for rough volatility. Then we calibrate the BSS parameters for the realised kernel of 21 equity indices, using data from the Oxford-Man Institute. We conduct one- and ten-step ahead forecasts on six indices and find that the BSS outperforms our benchmarks, including a Log-HAR specification, in the majority of cases.

Authors: Reviewers:
Lukas Grimm   Prof. Christian Brownlees
Jonathan Haynes   Prof. Eulalia Nualart
Daniel Schmitt

Barcelona, 31 May 2017

Contents

1   Introduction

Accurate models of asset volatility are useful in asset pricing and risk management. Forecasts of volatility are used in derivatives pricing and hedging, market making, market time, portfolio selection, and many other financial activities. In each case it is the predictability of volatility that is of importance. An options trader wants to know how the volatility of the asset that he is trying to price will evolve over the lifetime of the contract. A risk manager wants to know the probability that his portfolio will decline in the future. A market-maker may want to widen her spread when volatility is expected to rise. She may also want to know the forecasted volatility of her position to hedge against large price swings of her inventory. A portfolio manager may want to sell a stock before it becomes too volatile. Other portfolio managers may want to take a position on volatility as an asset class in itself, either as a macro hedge or as a speculative play. In corporate finance, volatility is used as an input for pricing Real Option Values when valuing prospective projects. Finally, in monetary policy, volatility of asset prices are used as a measure of risk and uncertainty. These examples give just a flavour of the practical motivation behind being able to accurately forecast asset volatility.

A good model of asset volatility must be able to forecast volatility. This is not straightforward since volatility itself is not observed. We need to find a proxy for it. Typically volatility models are used to forecast absolute or squared returns, but they may also be used to predict quantiles or even possibly the complete density function. For example, a good model specification for volatility is required to price and hedge certain types of financial derivatives or to forecast value-at risk (VaR) quantiles. Better proxies have been developed more recently in the continuous time finance setting by exploiting quadratic variation to obtain realised volatility. This estimator may not provide an accurate estimate if prices are measured in the presence of market micro structure noise. A wide range of robust realised measures of volatility, such as the realised kernel estimator, have been further developed to address this. This paper uses the realised kernel estimator as our robust estimator of volatility.

A well-specified volatility model should perform well over time. The accuracy of a volatility model in forecasting out-of-sample can be evaluated by comparing the results using a loss function, such as the Quasi-likelihood (QL) or Mean-Square Error (MSE), against benchmark models.

The main purpose of this Project is to evaluate the forecasting performance of the BSS model against a series of benchmark volatility models that preceded it. The benchmarks are a rolling variance, Exponential Weighted Moving Average (EWMA) and Log-HAR specification. We download daily data on the volatility of 21 stock indices from around the world. After conducting a one-step and ten-step ahead forecasts of the log volatility of these 21 assets, we then compare the MSE and QL against our benchmark models.

The paper is structured as follows: Section 2 reviews the literature on estimating volatility; Section 3 sets out the theoretical background; Section 4 introduces the data for the empirical analysis described in Section 5 and Section 6 concludes.

2   Literature Review

Volatility modelling is a topic that has received considerable interest. The extensive literature is probably a reflection of the importance of volatility in investment, valuation, risk management and monetary policy making.

A number of stylised facts about the volatility of financial asset prices have emerged over the years and been confirmed in numerous studies. Mandelbrot (1963) and Fama (1965) both reported evidence of persistence in volatility. They found that large changes in the price of an asset are often followed by subsequent large changes, and small changes are often followed by subsequent small changes. The long memory property of the volatility of a financial asset has been widely accepted as a stylised fact since the seminal work of Ding et al. (1993), Andersen and Bollerslev (1998) and Andersen et al. (2003). Initially long memory referred to the slow decay of the autocorrelation function (ACF), that is anything slower than an exponential decay. More recently, long memory has been formalized as non-integrability of the autocorrelation function. Gatheral et al. (2014) found evidence of 'roughness' in volatility by analysing high frequency price data on DAX and Bund futures contracts and US equity indices. Bennedsen et al. (2016) find a similar pattern in the volatility of E-mini S&P 500 futures contracts at intraday time scales and 29 individual US equities at daily frequency. A good volatility model should be able to capture these stylised facts of persistence and long memory, and roughness.

Over time modellers have tried to incorporate these stylised facts through different functional forms. Early models were based on simple historical values. Then following the seminal work of Engle (1982) the vast majority of subsequent studies on modelling volatility relied on his Autoregressive Conditional Heteroskedastic (ARCH) framework. There is now a large and diverse time-series literature on volatility modelling. Poon and Granger (2003) provide a good overview of how early models of asset volatiltiy tried to progressively improve to incorporate the stylised facts of persistence and long memory. Roughness is well covered in Gatheral et al. (2014).

The simplest historical model is the random walk, which simply uses the previous period of volatility,(σ_t) to forecast volatility in the next period, σ_{t+1}. Next, there are a set of (deterministic) models based on historical average methods. These include the Moving Average (MA), Exponential Moving Average (EMA) and the Exponentially Weighted Moving Average (EWMA) methods. In contrast to the other two, the EWMA places greater weight on more recent volatility estimates. Autoregressive (AR) models predict future values based on past values of order p. By including past volatility errors we arrive at the ARMA model specification. Finally, if we introduce a differencing of order I(d), we get to the AutoRegressive Integrated Moving Average (ARIMA) models when d=1 and the AutoRegressive Fractional Integrated Moving Average (ARFIMA) when d<1. All these models are similar in that the model specifications generate predictions based on historical estimates of volatility.

The next more sophisticated group of volatility models is the ARCH family of conditional volatility models. These predict future volatility based on the conditional variance of returns via maximum likelihood estimation. The first example is the ARCH(q) model of Engle (1982) where σ_t is a function of q past squared returns. In the GARCH(p,q) version of Bollerslev (1986) additional dependencies are permitted on p lags of past σ^2 and q lags of past square returns. Numerous extensions of have since been proposed, including: the Threshold GARCH (TARCH)

that allows for asymmetries from leverage effects, the Exponential GARCH (EGARCH) by Nelson (1991) that relaxes the non-negative parameter restrictions and the fractionally integrated version (FIGARCH) of Baillie et al. (1996) with $d \geq 0$.[1]

In the stochastic volatility modelling framework, volatility is subject to a source of innovation that may or may not be related to the factors that drive returns. Poon and Granger (2003) explains how modelling volatility as a stochastic variable immediately leads to fat tail distributions for returns. The autoregressive term in the volatility process introduces persistence, and correlation between the two innovation terms in the volatility process and the returns process produces the volatility asymmetry (see Hull and White (1987)and Hull and White (1988)). Heynen and Kat (1994), Heynen (1995) and Yu (2002) found that stochastic volatility forecasts performed best for stock indices, but Heynen and Kat (1994) concluded that the EGARCH and GARCH produced better forecasts for exchange rates. Long memory stochastic volatility models have also been proposed by allowing volatility to have a fractional integrated order (see Harvey (1998)). The noise term makes the stochastic volatility model more flexible, but the cost is that they do not have a closed form and, as a result, can not be estimated directly, e.g. by maximum likelihood. Rather they can be estimated via simulation (e.g. Duffie and Singleton (1993)) or numerical integration.

More recent market practice is to use local-stochastic-volatility (LSV) models, where $\sigma_t = \sigma_t(S_t, t)$, which both fit the market exactly and generate reasonable dynamics Gatheral et al. (2014). In recent years there has been increased interest in rough models of volatility. This is due to both theoretical developments in implied volatility modelling El Euch et al. (2016) as well as empirical evidence based on realised volatility (see Gatheral et al. (2014)). Researchers have found that standard Brownian motion is not 'rough' enough and is non-stationary, so more recent works have taken inspiration from fractional Brownian Motion (fBM), which is a stationary process that is able to exhibit roughness. Bennedsen et al. (2016), inspired by the fractional stochastic volatility (FSV) model of Comte and Renault (1996), propose using a Brownian Semi-Stationary (BSS) process in order to allow a decoupling of the persistence and roughness of the volatility in the simulating and forecasting of volatility.

In the previous section we highlighted how the correct specification of volatility is of fundamental importance for option pricing. As volatility models described above have become more sophisticated, this has lead to more advanced models of option pricing.

It is now well-known that the first seminal Black-Scholes formula Black and Scholes (1973) for option pricing fails to explain the implied volatility in out-of-the-money option contracts. By relying on a standard geometric Brownian motion with the simplified assumptions of a constant interest rate and a constant volatility, the Black-Scholes model has consistently failed to explain that the real-life 'implied volatility' of option contracts heavily depends on the the time to maturity of the contract and the extent to which the contract is currently in or out of the money. Practitioners refer to this phenomenon as the 'volatility smile.' This led to subsequent improved model specifications for volatility in option pricing. For example, Hull and White (1987) propose an option pricing model with volatility of the underlying asset as not only time-varying but

---

[1]As Hwang and Satchell (1998) and Granger (2000) point out that a major weakness of adopting the FIGARCH is that a positive I(d) process has a positive drift term or a time trend at the volatility level which is unobserved in practice. This makes it a poor model for our purpose.

also subject to a specific risk, coined the 'stochastic volatility paradigm'. This adjustment helps explain some of the stylised facts in the volatility smile. Comte and Renault (1998) propose a model to capture volatility persistence and particularly the occurrence of fairly pronounced smile effects even for rather long maturity options. This takes advantage of the Comte and Renault (1996) FSV model above. This was one of the first option pricing models to account for volatility persistence. More recent models have tried to also better account for the roughness, e.g. Bayer et al. (2016).

The added value of this Project is to give a broad overview to the new class of fractal processes for modelling volatility. We validate the BSS model of Bennedsen et al. (2016), identify potential areas for improvement and apply it to a wider range of equity indices.

3 Theory

Let us consider an asset whose price at time t, S_t, follows a Geometric Brownian motion, such that the dynamics of S_t in continuous time are given by the stochastic differential equation

$$dS_t = S_t(\mu_t dt + \sigma_t dB_t) \tag{1}$$

where $B_t = (B_t)_{t \geq 0}$ is standard Brownian motion, $\mu = (\mu_t)_{t \geq 0}$ a drift process and $\sigma_t = (\sigma_t)_{t \geq 0}$ a spot volatility process. As we are interested in volatility, our focus is with respect to the process $\sigma_t = (\sigma_t)_{t \geq 0}$. In particular, we adopt the following model outlined by Bennedsen et al. (2016)

$$\sigma_t = \xi e^{X_t} \tag{2}$$

where ξ originally denotes the variance swap forward curve. However, given the complexity of estimating this curve it is left as a free positive parameter hence the model is primarily driven by the process X_t.

This section begins by exploring how best to measure past realisations of σ_t. It then introduces some stylised facts of volatility and shows how they may be incorporated through X_t.

3.1 Observing and Measuring Volatility

True volatility of a price process is unobservable and hence must be proxied for. High frequency data allows for very accurate proxies to be estimated however issues arise when market microstructure noise is introduced. In the following we present a brief theoretical overview of how estimate volatility under these conditions.

Specifying a step size $\Delta > 0$ such that $T = n\Delta$ for some large $n \in \mathbb{N}$, one can define the *integrated variance*

$$IV_t^\Delta := \int_{t-\Delta}^{t} \sigma_s^2 ds, \qquad t = \Delta, 2\Delta, \ldots, n\Delta \tag{3}$$

where σ_s^2 is the spot variance. Choosing Δ sufficiently small, the integrated variance provides an estimate of the spot variance

$$\hat{\sigma}_t^2 = \Delta^{-1} \hat{IV}_t^\Delta \qquad t = \Delta, 2\Delta, \ldots, n\Delta \tag{4}$$

Theoretically,

$$\lim_{\Delta \to 0} \hat{\sigma}_t^2 = \sigma_t^2 \tag{5}$$

One can approximate $\hat{IV}_t^{\Delta}$ through the *realised variance* which is defined by discretising the integral in equation (2) as follow

$$RV_t = \sum_{i=1}^{N} (S_i - S_{i-1})^2 \tag{6}$$

where S_i and S_{i-1} are the observed prices at the beginning and end of each discretisation cell. Barndorff-Nielsen (2002) show that the realised variance converges to $\hat{IV}_t^{\Delta}$ in probability. However, at high frequencies prices are subject to market microstructure noise, in effect masking the true price and thus spot volatility of the asset. As shown by Masoliver et al. (2014), if we define the observed price as

$$S_t = S_t^* + \epsilon_t \tag{7}$$

where S_t^* is the true price and ϵ_t is a noise term, the realised variance can be decomposed as follows

$$RV_t = RV_t^* + 2\sum_{i=1}^{N} (S_t - S_{t-1})(\epsilon_t - \epsilon_{t-1}) + \sum_{i=1}^{N} (\epsilon_t - \epsilon_{t-1})^2 \tag{8}$$

resulting in a biased estimator of $\hat{IV}_t^{\Delta}$. Given this bias, alternative methods have been proposed to estimate $\hat{IV}_t^{\Delta}$, such as realised kernels (Andersen et al. (2001); Barndorff-Nielsen (2002)), two-scale estimators (Zhang et al., 2006), and pre-averaging methods (Jacod et al., 2009). Masoliver et al. (2014) find that the realised kernel outperforms realised variance at high frequencies, hence we choose it as our estimator of $\hat{IV}_t^{\Delta}$.

3.2 Stylised Facts of Volatility

While it has been well documented that volatility in financial markets is *persistent*, more recently it has also been empirically shown to be *rough* (Gatheral et al., 2014). That is, standard Brownian motion fails to accurately simulate empirical volatility of financial assets because the process is smoother than the realised prices.

3.2.1 Fractal Geometry

In order to model more realistic processes, researchers can use fractal geometry, e.g. see Comte and Renault (1996), Mandelbrot (1963). Fractal geometry is distinct from Euclidean geometry, specifically in the sense that it allows for non-integer or *fractal* dimensions. In order to quantify the roughness, one can count the number of circles of radius r required to cover the time series. Increasing r and repeating the process leads to the following relationship

$$N(2r)^d = 1, \qquad d \in (1,2) \tag{9}$$

where N is the number of circles required, r is the radius and d is the fractal dimension of a line. For a straight line, that is a deterministic process, d is 1 while for a random walk it is 1.5, given that the process has an equal probability of going up or down. For $1 < d < 1.5$ the process is somewhere between deterministic and random, that is, it is somewhat trending. On the other hand, $1.5 < d < 2$ implies it is rougher than a random walk, i.e. it has more reversals (see Peters (1994)).

3.2.2 Volatility is Rough

Since prices are thought to be rough, hence more erratic than a standard Brownian motion, they may be modelled by a process with fractal dimension $1.5 < d < 2$. Directly related to d is the Hurst exponent H

$$d = 2 - H, \qquad H \in (0,1) \tag{10}$$

and the fractal index α

$$d = 1.5 - \alpha, \qquad \alpha \in \left(-\frac{1}{2}, \frac{1}{2}\right) \tag{11}$$

which Bennedsen et al. (2016) dub the *roughness index*. The Hurst exponent, proposed in Hurst et al. (1965) is a measure of dependence as it captures the scaling behaviour of correlations within a time series with respect to the observation period and time resolution (Di Matteo et al., 2003). It is straight forward to define H in terms of the fractal index

$$H = \frac{1}{2} + \alpha \tag{12}$$

Considering a covariance stationary time series $X(t)$, Di Matteo et al. (2003) show H can be generalised in terms of the q-order moments of the distribution of $X(t)$'s increments,

$$K_q(\tau) = \frac{E[|X(t+\tau) - X(t)|^q]}{E[|X(t)|^q]} \sim \left(\frac{\tau}{v}\right)^{qH(q)} \tag{13}$$

where K_q is the q-order moment, τ is the time interval between increments, X_t is the underlying time series, q is some order greater 0, v is the time resolution and $H(q)$ is the generalized Hurst parameter of order q. Here $\sim$ means that the ratio of the left and right hand side tends to a constant as τ goes to infinity. Of particular interest is the second moment which allows Bennedsen et al. (2016) to find the following expression

$$1 - \rho(h) = 1 - Corr(\log \sigma_t, \log \sigma_{t+h}) \sim |h|^{2\alpha+1}, \qquad |h| \to 0 \tag{14}$$

where ρ denotes the autocorrelation function of log volatility, $|h|$ is the absolute lag and v is set to 1. Taking logs, Bennedsen et al. (2016) show that the variable of interest α can then be estimated from

$$\log(1 - \hat{\rho}(h)) = c + a \log|h| + \epsilon_h, \qquad h = \Delta, 2\Delta, \ldots, m\Delta \tag{15}$$

6

via OLS regression, where m denotes the bandwidth and Δ the step size. Note that we obtain alpha from $\alpha = \frac{\hat{a}-1}{2}$.

3.2.3 Volatility is Persistent

Volatility in financial markets has been shown to exhibit strong persistence as documented by Andersen and Bollerslev (1998). This phenomenon is often referred to as long range dependence. Formally, a time series is said to exhibit long range dependence if the following holds

$$\lim_{h \to \infty} \rho(h) = c_\rho h^{-\beta} \tag{16}$$

where c_ρ is a constant greater 0 and β is the rate of decay of the autocorrelation function. Beran (1994). Similar to the previous case of α, the expression can be found by taking K_q to the limit. When $\beta \in (0, 1)$ we see that

$$\int_0^\infty |\rho(h)| dh = \infty \tag{17}$$

In other words the autocorrelation will never be zero, even for an infinite lag, hence the autocorrelation function is not integrable.

Using this relationship Bennedsen et al. (2016) again take logs in order to estimate β using OLS as follows

$$\log(\hat{\rho}(h)) = c + b \log|h| + \epsilon_h, \qquad h = M\Delta, 2\Delta, \ldots, M'\Delta \tag{18}$$

Where, $M < M'$ sets the estimation window and Δ is again the step size. One can recover β since $\beta = -\hat{b}$. Further, Beran (1994) shows that β can be expressed in terms of the Hurst exponent as $H = 1 - \frac{\beta}{2}$ and thus as $d = 1 + \frac{\beta}{2}$.

3.3 Stochastic Processes

As alluded to earlier, in order to model persistence in volatility a process with fractal dimension $1 < d < 1.5$ is required as this implies non-randomness or dependence. However, modelling the roughness requires $d > 1.5$ hence the two properties are in direct conflict with one another when described by a single parameter such as d or H. Thus we require a stationary process which not only incorporates both the *roughness* as well as the *persistence* but does so using separate parameters. In the following we introduce three different stochastic processes. We evaluate these as candidates for X_t in the context of the stylised facts. Simulations of these processes can be found in Figure 3.

3.3.1 Brownian Motion

The baseline for a continuous stochastic process is a standard Brownian motion. Let Δ be a constant time increment, then one can define Brownian motion as a process such that

$$B_0 = 0, \qquad B_t = \sum_{i=1\Delta}^{t=n\Delta} \epsilon_i, \qquad \epsilon_i \sim i.i.d. \mathcal{N}(0, \Delta) \tag{19}$$

Given that each ϵ is independent of the others, it is not possible to model persistence through this process. Further, given that there are equal probabilities of the next increment going up or down, there are too few reversals to obtain a rough enough trajectory. In other words, its fractal dimension d, is exactly equal to 1.5.

3.3.2 Fractional Brownian Motion

Similar to B_t, fractional Brownian motion, $B_H(t)$ starts at zero and is a continuous-time Gaussian process. Mandelbrot and Van Ness (1968) defined it as

$$B_H(t) := \frac{1}{\Gamma(H + \frac{1}{2})} \left(\int_{-\infty}^{0} [(t-s)^{H-1/2} - (-s)^{H-1/2}] dB(s) + \int_{0}^{t} (t-s)^{H-1/2} dB(s) \right) \tag{20}$$

where $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1} e^{-x} dx$ and H denotes the Hurst index, see Dicker (2004). The main difference between $B_H(t)$ and B_t is that the increments of the former are not necessarily independent. Thus $B_H(t)$ is characterised by the following covariance function

$$\mathbb{E}[B_H(t)B_H(s)] = \tfrac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}) \tag{21}$$

Setting the Hurst index to $\frac{1}{2}$, the covariance function will be equal to s, and we have B_t. This is in line with our previous definition of $d = 1.5$ for B_t and shows that $B_H(t)$ is a generalisation of B_t. While $B_H(t)$ is a promising candidate, allowing $d \in (1,2)$ and thus for rougher or more persistent processes than B_t, it does not allow for both as H is the sole parameter.

3.3.3 Brownian Semi-Stationary Motion

In order to model volatility and take into account both α and β, Bennedsen et al. (2016) define specific kernels for a *Brownian Semistationary motion*. BSS was originally introduced by Barndorff-Nielsen and Schmiegel (2009) as

$$BSS_t = \int_{-\infty}^{t} g(t-s) v_s dB_s \tag{22}$$

where $g(t-s)$ is the square integrable kernel function, and v_s is an adapted covariance stationary volatility term (see Barndorff-Nielsen et al. (2009)). It is worth noting that when v_s is constant or deterministic BSS_t remains Gaussian. However when v_s is taken to be a stochastic process in itself, BSS_t becomes non-Gaussian.

Bennedsen et al. (2016) introduce the *gamma* and *power law* kernel where the former models short range dependence and the latter long range dependence. The gamma kernel is given by

$$g(t-s) = (t-s)^\alpha e^{-\lambda(t-s)}, \qquad (t-s) > 0, \qquad \alpha \in \left(-\frac{1}{2}, \frac{1}{2} \right), \qquad \lambda \in (0, \infty) \tag{23}$$

where α is the roughness index and λ the memory parameter. The theoretical ACF of the gamma-kernel is defined as

$$\rho_{gamma}(h) = \frac{\left(\frac{|h|}{2\lambda}\right)^{\alpha+1/2} K_{\alpha+1/2}(\lambda|h|)}{(2\lambda)^{-2\alpha-1}\pi} \tag{24}$$

where $K_v(x)$ is the modified Bessel function of the third kind. This gamma BSS process is said to have short term memory as its ACF decays exponentially fast in the limit,

$$\rho(h)\sim e^{-\lambda h}h^{\alpha}, \qquad h \to \infty \tag{25}$$

The power law kernel is given by

$$g(t-s) = (t-s)^{\alpha}(1+(t-s))^{-\gamma-\alpha}, \quad (t-s)>0, \quad \alpha \in \left(-\frac{1}{2},\frac{1}{2}\right), \quad \gamma \in \left(\frac{1}{2},\infty\right) \tag{26}$$

where γ is the memory parameter. The ACF of the power law kernel is given by

$$\rho_{power}(h) = \frac{\int_0^{\infty} g(x)g(x+|h|)dx}{B(2\alpha+1,2\gamma-1)} \tag{27}$$

where $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt$. The autocorrelation functions are included here as they are required in order to forecast volatility later on.

It is now apparent why the BSS process under a gamma or power law kernel should provide significantly better forecasts. The kernels allow both roughness and persistence to enter separately into a covariance stationary stochastic process. Therefore, we should expect models driven by a BSS process to perform better than if they were driven by $B_H(t)$ or B_t.

3.4 Simulation

The accurate and efficient simulation of stochastic processes is vital, especially when pricing derivatives for which no closed form solution is available. In this case one has to resort to Monte Carlo simulation to generate a large number of possible price paths, compute the path's exercise value and average these out in order to get the expected payoff of the derivative. Intuitively, the more price paths the better the approximation will be. Hence, efficiency is a sought after attribute in simulation methods.

Regarding the three processes introduced in the previous section, one can differentiate between B_t which has independent increments and $B_H(t)$ as well as BSS which have correlated increments. When simulating dependent processes such as the latter two, we have to take into account the correlation between increments. This can be done via either exact or approximate estimation methods.

An example of an exact simulation method is the Cholesky decomposition, which entails computing the process's covariance matrix for the entire sample path using the theoretical autocovariance function. This is computationally intensive since a sample path of length N requires a covariance matrix of dimension $N \times N$. A much more efficient method would be to

discretise the process in question and use Riemann sums to approximate each realization. In the case of the $B_H(t)$ and BSS process, efficiency can be further improved by realizing that these are convolutions. Hence, using the Fast Fourier Transform (FFT) one can convert the discretized process into the frequency domain, where a convolution in the time domain becomes a multiplication, and once solved convert it back into the time domain using the inverse FFT. Bennedsen et al. (2015) further improved the accuracy of the Riemann approach by combining it with exact simulation, which he calls the hybrid scheme. Bennedsen et al. (2016) notice that in the particular case that the kernel $g(x)$ behaves as a power-law function near zero, as $x \to 0$, the Riemann sum fails to approximate the kernel. Hence using exact simulation for the first couple of steps before switching to the Riemann sum approximation significantly improves the accuracy while remaining much more efficient than the Cholesky decomposition.

Figure 3 shows some simulations of BM, fBM with different values of α and BSS compared to the realised kernel of the FTSE MIB, as an example of the historical pattern of asset volatility. It is easy to see that the volatility of the equity index is poorly modelled by standard Brownian motion. By altering the Hurst parameter (reflected by the corresponding α value), fBM can produce rougher simulations that are more realistic to the volatility of equity indices. The BSS simulation is also able to replicate this roughness. In the Results section we will discuss in detail how to calibrate these processes to a given time series, such as these equity indices.

4 Data Set

We downloaded daily realised variances for 21 equity indices from around the world from the period January 2000 to May 2017. The full sample includes: S&P 500, FTSE 100, Nikkei 225, DAX, Russel 2000, All Ordinaries, DJIA, Nasdaq 100, CAC 40, Hang Seng, KOSPI Composite, AEX Index, Swiss Market, IBEX 35, S&P CNX Nifty, IPC Mexico, Bovespa, S&P/TSX Composite, Euro STOXX 50, FT Straits Times and FTSE MIB. Seventeen are large cap indices, one small cap (Russell 2000) and three indices are a mixture of large and small-caps. Eight indices are based on companies from Europe, four from the US, four from Asia, and the rest are based on companies from Brazil, Canada, Australia and Mexico. In terms of methodology, 80% are calculated based on a market capitalisation weighting, while the others are computed using a metric based on either price or trading volume.

The realised volatilities are based on 5-minute prices and were downloaded from the Oxford-Man Institute's Realized Library data, available at http://realized.oxford-man.ox.ac.uk/. For each of these 21 indices the Oxford Man Institute's Realised library records daily returns, daily realised variances and daily realised kernels. We use the daily returns and the daily realised kernel, introduced by Barndorff-Nielsen et al. (2008), as it has some robustness to the effects of market micro-structure effects. See Heber et al. (2009) for more detail on how these measures are computed. See Shephard and Sheppard (2010) and Barndorff-Nielsen et al. (2009) for more background on the realised measures.

If the market was closed or the data was regarded as being of too low quality for that index the database shows a missing value (NA), except for days when all the markets are simultaneously closed, in which case the day is not recorded in the database. As a result, for example, Saturdays are never present in the library. The full time period left us with 4521 observations, excluding a

number of missing values which varied across index. All missing values were dropped before conducting our analysis. We also extracted index return data from the same source.

Figure 1 reports the basic summary statistics for the 21 stock indices in our sample. The range of realised volatility varies quite considerably across indices over the full time period. For example, the S&P 500 had a minimum daily volatility of 0.0002% and a maximum of 0.9313% with a mean of 0.0111%, standard deviation of 0.0255%, a positive skewness of 15.05, and a kurtosis of 426.84. The FTSE 100 volatility in the sample period ranged from 0.0002% to 0.3256%, with a mean of 0.0084%. We note that the skewness and the kurtosis of the series imply some deviation from a standard Gaussian distribution. Some deviations from normality also become apparent when we plot QQ plots of the series. Figure 2 shows the QQ Plot for the S & P 500 series as an illustration.

5 Results

5.1 Estimation of the BSS Model

To fit the BSS model to the data described in Section 4 we must estimate the model parameters introduced in section 3, namely the roughness parameters α and β and the memory parameter λ. In doing so we closely follow the procedures outlined in Bennedsen et al. (2016).

Recall that the roughness parameter α introduced in Equation (14) can be estimated by taking the natural logarithm on both sides, leading to the OLS regression Equation (15):

$$\log(1 - \hat{\rho}(h)) = c + a\log|h| + \varepsilon_h, \quad h = \Delta, 2\Delta, \dots, m\Delta \tag{28}$$

where $\hat{\rho}(h)$ is the empirical autocorrelation function at lag h. For the step size we chose $\Delta = 1$ day such that each lag corresponds to one day for a bandwidth of $m = \lceil n^{1/3} \rceil \approx 17$ days, where n corresponds to the number of observations for each vector (ticker) under consideration and $\lceil \cdot \rceil$ the ceiling.

We run the regression for each of the 21 tickers in our dataset and recover alpha using the relationship $\hat{\alpha} = (\hat{a} - 1)/2$. Figure 4 a) shows the OLS regression results for the S&P 500 as an illustration. The full set of alpha estimates are in Figure 4 b). We find that there is some substantial variation in α across the 21 indices. The smoothest ticker is the Swiss Market Index, with an estimated alpha of -0.3268. The roughest series is the Australian All Ordinaries with an estimated alpha of -0.4228, close to the roughest possible value of -0.5 (corresponding to a Hurst index of 0). The average alpha estimate across the full cross section is -0.3706, indicating that the volatility of equity indices is indeed rough.

We also investigated patterns across region of stock index, by type of index (i.e. large-cap, small-cap or all), and by methodology of index. Figure 5 illustrates the estimated parameter values by geographic region. There were no strong patterns by type or methodology. We did notice, perhaps unsurprisingly, that stock indices from similar regions, particularly all those indices covering US companies, reported similar estimate levels of roughness and persistence in volatility over the 2000-2017 sample period analysed.

A natural question is whether the estimates of alpha change over time. In order to answer this question we estimated alpha on a rolling basis for each security using a window of 60 consecutive

trading days. Figure 6 shows the trailing 60 trading days alpha estimates for the S&P 500. We find that alpha does substantially vary over time. The estimates broadly fall within the bound $(-0.5, -0.1)$. Recall that a value of alpha close to -0.5 implies roughness and -0.1 smoother. In line with the findings in Bennedsen et al. (2016), we observe several peaks of smoothness that coincide with periods of market turmoil. For example there is a clear peak of alpha around the Lehman collapse in 2008. We estimated the rolling variance for all the 21 equity indices and find that the same pattern that volatility exhibits less roughness during market turmoil. This is again in line with the findings of Bennedsen et al. (2016) and Gatheral et al. (2014). Of course the periods of market turmoil can be different across markets. For example we noted a higher alpha peak corresponding to the dotcom bubble of the early 2000s in the NASDAQ 100 when compared to the other U.S. indices which have less focus on the technology sector.

Next we estimated the beta parameter over the entire sample to gauge the persistence of the underlying series, see Figure 4 c). Recall that Equation (18) allows us to estimate beta using simple OLS. The findings reported in Figure 4 c) are estimated for $M = m$ to $M' = m + 9$ (note that m in our sample corresponds to 17). We obtain the highest beta of approximately 0.5 for the Nikkei. We obtain the lowest reading for the TSX Composite index, which compares to an average beta of 0.1996 for all 21 indices. These results differ from the findings in Bennedsen et al. (2016). The reason behind this difference can be found in the very large sensitivity of Beta estimates with respect to the estimation window defined by M and M', and of course the underlying data. To illustrate the sensitivity of beta estimates we have calculated betas for all 21 indices fixing M to 17 and varying M' from 26 to 278 representing an autocorrelation window ranging from 10 days to one year. Figure 7 shows our results. Indeed we find the beta estimates to vary substantially and we confirm that beta estimates are highly sensitive to the choice of estimation range M to M'. Theory suggests the choice of large M. We repeated the exercise for a window of $M = 1$ to $M = m \approx 17$ and find results closer to the ones reported in Bennedsen et al. (2016). Here, the KOSPI Composite index exhibits the highest persistence with an estimated beta of 0.0636. On the other extreme is the Bovespa index with a beta of 0.1942. The overall sample beta mean is 0.1079, indicating a relatively high persistence overall.

Lastly we estimate the memory parameter, λ for the gamma (and power) kernel BSS models. Lambda can be estimated by a methods-of-moments (MoM) procedure. Following Bennedsen et al. (2016) we use the previously fitted alpha estimates from above to plug them into the theoretical autocorrelation function for the gamma kernel in Equation (24). The next step is to match the theoretical and the empirical ACF and minimize the squared differences over lambda.[2] The results are shown in Figure 4 d). The results range from the lowest reading for the Swiss Market Index, 0.0041, to a maximum of 0.0230 for the Bovespa, with a mean of 0.0101. Note that a small lambda would imply a model with long memory. We find that the estimates of λ are highly sensitive to the choice of the lag of the ACF, $\hat{\rho}(h)$, in equation (15), which impacts on the functionality of the BSS model (discussed in more detail in the next section).

Having estimated all the relevant parameters, we are now able to construct the model calibrations for the fBM and gamma-kernel BSS specifications. For example Figure 8 shows our results for the estimated values of the S&P 500 using 500 i.i.d. drafts of a multivariate Gaussian vector with mean zero and covariance matrix as described in Bennedsen et al. (2015).

---

[2]This has been done using the autocorr and fmincon commands in matlab

The Truncated BSS (TBSS) model and the BSS trajectories shown in Figure 8 make use of a modelling technique, referred to as hybrid scheme introduced in Bennedsen et al. (2015). The hybrid scheme discretises the stochastic integral in the time domain and approximates the first kappa steps by a power function near zero and a standard Riemann approximation after. The TBSS process is an extension to the BSS process with applications in rough Bergomi models. The TBSS process allows us to directly compare the trajectories generated with the Hybrid Scheme with those of the Riemann approximations. The red lines in the TBSS and BSS Hybrid panels represent approximations setting kappa to 3. To compare the persistence fit of the BSS process we compared the empirical autocorrelation function with the autocorrelation function from the fitted processes. Figure 9 shows the results for the S&P 500. The BSS model satisfyingly replicates the autocorrelation pattern of the empirical ACF.

More important than simulating the processes with fitted parameters, we are now able to conduct forecasts using the BSS model.

5.2 Forecasting with BSS

In this section we use the BSS process to forecast realised variance and then evaluate the forecast accuracy of the model and compare it against a series of benchmarks. We do this in a few steps. First, we conduct an in-sample forecast. We use the entire in-sample dataset to estimate the parameter for the models. Second, we use those parameters to forecast the realised variance h steps ahead. Finally, we compare the Mean Square Error (MSE) and Quasi-likelihood (QL) against a set of benchmark models. Let us discuss these steps in turn before reporting the results.

Firstly, to forecast with the BSS process we assume Gaussianity of the processes and follow the steps taken in Bennedsen et al. (2016). In particular we rely on the fact that for a zero-mean Gaussian vector $(x_{t+h}, x_t, x_{t-1}, ..., x_{t-m})^T$ the distribution of x_{t+h} conditionally on $(x_{t+h}, x_t, x_{t-1}, ..., x_{t-m})^T = a \in \mathbb{R}^{m+1}$ is

$$X_{t+h}|[x_t, x_{t-1}, ..., x_{t-m})^T = a] \sim \mathcal{N}(\mu, \Xi^2) \tag{29}$$

where $\mu = \Gamma_{12} \cdot \Gamma_{22}^{-1} \cdot a = Z \cdot a$, and Γ_{22} is the correlation matrix of the vector a and $\Gamma_{12} = [\rho_{x_{t+h},x_t}, \rho_{x_{t+h},x_{t-1}}, ..., \rho_{x_{t+h},x_{t-m}}]$, where the correlations are obtained from the theoretical correlation functions given in equation (24) and (27). The variance term is given by $\Xi^2 = \text{Var}(x_t) * (1 - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21})$ where, $\Gamma_{21} = \Gamma_{12}^T$. [3]

In the forecasting exercise we need a window of m previous data points on which we can then compute an h step ahead forecast. The m is a rolling window and the parameters (e.g. $\hat{\alpha}$, $\hat{\lambda}$) for the model are computed from the full data in ex-ante.

In order to retrieve the BSS process from our dataset we first take the square root to arrive at volatility. Next we drop empty observations to remain consistent with the estimation of model parameters, recall that we construct a series of consecutive trading days. For any index, this leaves us with the vector for variances, σ. Recall the relation between volatility and BSS-process as stated in Equation (2). Thus, to recover the realised BSS-process, $\{X\}_{t \in \{1,...,T\}}$, we first take

---

[3]To implement this matrix multiplication in practice we use the 'fliplr' command in Matlab on the Γ)12 matrix to ensure the correct time order of correlation lags.

the natural logarithm and then de-mean the vector as shown below:

$$\ln \boldsymbol{\sigma} = \begin{bmatrix} \ln \sigma_1 \\ \vdots \\ \ln \sigma_T \end{bmatrix}, \quad \boldsymbol{X} = \left[\boldsymbol{I} - \frac{1}{T} \boldsymbol{i}\boldsymbol{i}' \right] \ln \boldsymbol{\sigma}$$

where $\boldsymbol{i}$ denotes the unit vector of size T, $\boldsymbol{i}'$ its transpose and $\boldsymbol{I}$ the identity matrix. The result follows from $1/T \sum_{t=1}^{T} \ln \sigma_t = 1/T \sum_{t=1}^{T} \ln \xi + X_t = \ln \xi + 1/T \sum_{t=1}^{T} X_t = ln\xi$ and zero mean property of the BSS process. [4]

From this new vector $\boldsymbol{X}$ we feed the rolling window a. Note that the functional form of volatility given in Equation (2), requires us to forecast a log-normal distributed process. To account for this we make use of the moment generating function and obtain our forecasts according to:

$$\mathbb{E}[\exp(X_{t+\Delta}|\mathcal{F}_t)] = \exp\left(\mathbb{E}[X_{t+\Delta}|\mathcal{F}] + \frac{1}{2}\text{Var}[X_{t+\Delta}|\mathcal{F}] \right) \tag{30}$$

This term is then approximated by inserting μ and Ξ^2 from above. Pre-multiplying the h-step ahead forecast vector obtained in this fashion by ξ from the de-meaning step, we get the forecast vector for realised volatility $\hat{\boldsymbol{\sigma}}_{t+h}$.

Once we have the model estimated, the second step is to forecast ahead with it. We forecast one-step ahead (i.e. h=1) recursively, re-estimating the model each time based on the new updated information set. We start forecasting after 200 periods of observations.

The final step of the forecasting methodology is to compute the loss functions. Here, we use two Patton class loss functions:[5] the Mean Square Error (MSE), defined as

$$MSE : L(\tilde{\sigma}_t^2, \sigma_{t|t-1}^2) = (\tilde{\sigma}_t^2 - \sigma_{t|t-1}^2)^2; \tag{31}$$

and the Quasi Likelihood (QL) loss function, defined as:

$$QL : L(\tilde{\sigma}_t^2, \sigma_{t|t-1}^2) = \frac{\tilde{\sigma}_t^2}{\sigma_{t|t-1}^2} - log\frac{\tilde{\sigma}_t^2}{\sigma_{t|t-1}^2} - 1. \tag{32}$$

We identified some calibration issues stemming from the Method of Moments (MoM) estimation of $\hat{\lambda}$ that impact on our ability to forecast with the Gamma-BSS model. After thorough exploration of the Bennedsen et al. (2016) methodology we noticed forecast errors exploding for some series and isolated $\hat{\lambda}$ estimation as the driving force. To further evaluate the problem we calculated the MSEs for a range of λ from 0.0005 to 0.05 in steps of 0.0001, see Figure 11. The peak in the first plot shows the critical range around which forecast errors explode. Plot 2 shows the MSE on a logarithmic scale. We find the left hand side of the peak does not contain a global minimum. By zooming in on the right hand side, as shown in plot 3, we find the global minimum (i.e. the point with lowest MSE). Note that this value for λ is far from the minimum identified by MoM. Our findings suggest MoM might not be suited for optimal forecasting results.

---

[4]We ran Jarque–Bera normality tests on the realised $\boldsymbol{X}$ process for the 21 series and reject the null of normality in all cases except one.

[5]Patton (2011) derives a class of loss functions that can be used to rank volatility models.

5.3 Forecasting Performance

After estimating each model from the in-sample data, one-step ahead conditional volatility forecasts are produced for the out-of-sample period. The λ estimation issue, described above, allows us to produce one-step ahead forecasts for six indices, namely: S&P 500; NASDAQ 100; Bovespa; AEX; Swiss Market; FTSE MIB. These six provide a decent geographical coverage of major equity markets. Figure 10 shows the results for the forecasting exercise. The BSS forecast performance appears to be very good for these indices, with one-step ahead forecasted values very close to the realised values. To evaluate this forecast performance in more depth we introduce three benchmark models of volatility.

- Rolling Volatility: calculated as the standard deviation of the returns of the last 200 consecutive trading days

$$\sigma_{rolling,t+1} = std(Returns(t - 200 : t)) \tag{33}$$

- Exponential Weighted Moving Average (EWMA): with a value of $\lambda = 0.94$. We take the square root of σ^2_{EWMA} to obtain volatility.

$$\sigma^2_{EWMA,t+1} = \lambda \cdot \sigma_{EWMA,t} + (1 - \lambda) \cdot (Returns^2_t) \tag{34}$$

- Log HAR model specification: where next period's log volatility is based on a weighted average of the average log volatility over the last day, week and month. We take the exponent of σ_{logHAR} to obtain volatility.

$$\sigma_{LogHAR,t+1} = \beta_0 + \beta_1 \cdot \sigma_{DAY,t-1} + \beta_2 \cdot \sigma_{WEEK,t-1} + \beta_3 \cdot \sigma_{MONTH,t-1} \tag{35}$$

The first two provide more traditional benchmarks. The third is a more recent state-of-the art volatility model that should be harder to beat.

Table 1 reports the MSE and QL losses of the three benchmarks and the BSS model for the one-step and ten-step ahead forecasts of the six indices. Looking first at the MSEs for the one-step ahead forecast, BSS outperforms all three benchmarks for the Swiss Market index. It outperforms the rolling volatility and EWMA, but not the Log-HAR, for the S& P 500, Nasdaq 100 and the Bovespa. The BSS has a lower MSE than the rolling variances, but a higher MSE than the Log-HAR and very similar MSE to the EWMA, for the FTSE MIB and AEX. This pattern is repeated if we consider the QL loss function. In short the BSS is either the best performer or nearly as good as the Log-HAR specification. For the ten-step ahead forecast, the BSS outperforms all three benchmarks under both the MSE and the QL loss function for five of the six indices.

Next, we investigate if this BSS forecast out-performance versus some of benchmarks is statistically significant. Diebold and Mariano (1995) introduce a test to assess if the forecasting ability of two series is statistically different. The test enables us to see if model out-performance against benchmarks is statistically significant. The null hypothesis is that two forecasting strategies have the same predictive ability. Table 2 reports the DM-statistics and the corresponding p-values. For the one-step ahead forecast, the BSS statistically outperforms the rolling volatility

15

model for all six indices at a 5% significance level. BSS statistically outperforms the EWMA for the Nasdaq, Swiss Market Index and Bovespa indices and the Log-HAR in the case of Swiss Market Index. However, the Log-HAR is statistically better than BSS for the other five indices. For the ten-step ahead forecast the BSS outperformance is closer to the Log-HAR benchmark and under the MSE loss function it is significantly better for AEX, Swiss Market Index and FTSE MIB. Under the QL loss function the BSS significantly outperforms for all benchmarks for all series.

6 Conclusion

This project confirms the findings of Gatheral et al. (2014) and Bennedsen et al. (2016) that volatility is indeed both rough and persistent across a wide range of equity indices. We have explored the advantage of using a Brownian Semi-Stationary (BSS) process to model volatility enabling the user to calibrate both stylised facts in contrast to previous generations of fractal processes, like Fractional Brownian Motion. We have successfully implemented simulation methods so that a BSS process can be incorporated within a continuous time asset pricing equation to price options and other exotic derivatives. We then calibrated the parameters for the BSS model using the realised kernel of 21 equity indices. Our parameter estimates confirm the expected roughness and persistence in the series. The parameter for roughness, α, was quite stable across the cross-section of indices, but fluctuated over time. α averaged -0.37 and ranged from -0.33 to -0.42, implying much more roughness than the $\alpha = 0$ implied by Standard Brownian Motion. Estimates of the long memory parameter, λ, were less stable, ranging from 0.0041 to 0.0230. We identify an issue when using MoM estimation that suggests MoM may be sub-optimal for BSS-Gamma forecasting. We forecast with six indices that cover a broad geographical spread and have stable lambda estimates. For the one-step ahead forecast we find that the BSS model outperformed two of our three benchmarks consistently under both MSE and QL loss functions. The BSS beat the Log-HAR benchmark in the case of the index with the longest memory, while it was slightly worse for the other five indices. For the ten-step ahead forecast, under the MSE loss function, the BSS model outperformed all benchmarks consistently for five out of six indices. Under the QL loss function the BSS outperforms all benchmarks, and this outperformance is always statistically significant.

Areas for further research would include investigating the forecasting accuracy of the BSS Power Kernel using a wider range of asset class, such as commodities, real estate funds and foreign exchange rates. Further robustness checks could test the performance of BSS against the family of fractional volatility models. It would also be interesting to further explore the relationship of ξ and its link with the variance swap curve.

7 References

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American statistical association 96*(453), 42–55.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 74*(1), 3–30.

Barndorff-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(2), 253–280.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal 12*(3), C1–C32.

Barndorff-Nielsen, O. E. and J. Schmiegel (2009). Brownian semistationary processes and volatility/intermittency. *Advanced financial modelling 8*, 1–25.

Bayer, C., P. Friz, and J. Gatheral (2016). Pricing under rough volatility. *Quantitative Finance 16*(6), 887–904.

Bennedsen, M., A. Lunde, and M. S. Pakkanen (2015). Hybrid scheme for brownian semistationary processes. *arXiv preprint arXiv:1507.03004*.

Bennedsen, M., A. Lunde, and M. S. Pakkanen (2016). Decoupling the short-and long-term behavior of stochastic volatility. *SSRN*.

Beran, J. (1994). *Statistics for long-memory processes*, Volume 61. CRC press.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of political economy 81*(3), 637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Comte, F. and E. Renault (1996). Long memory continuous time models. *Journal of Econometrics 73*(1), 101–149.

Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance 8*(4), 291–323.

Di Matteo, T., T. Aste, and M. Dacorogna (2003). Scaling behaviors in differently developed markets. *Physica A: Statistical Mechanics and its Applications 324*(1), 183–188.

Dicker, T. (2004). *Simulation of fractional Brownian motion*. Ph. D. thesis, University of Twente.

Diebold, F. X. and R. S. Mariano (1995, July). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics 13*(3), 253–263.

Ding, Z., C. W. Granger, and R. F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance 1*(1), 83–106.

Duffie, D. and K. J. Singleton (1993). Simulated moments estimation of markov models of asset prices. *Econometrica: Journal of the Econometric Society*, 929–952.

El Euch, O., M. Fukasawa, and M. Rosenbaum (2016). The microstructural foundations of leverage effect and rough volatility. arxiv preprint. *arXiv preprint arXiv:1609.05177*.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business 38*(1), 34–105.

Gatheral, J., T. Jaisson, and M. Rosenbaum (2014). Volatility is rough. *working paper*.

Granger, C. W. (2000). *Advances in Statistics, Combinatorics and Related Areas*, Chapter Long Memory Processes-an Economist's Viewpoint, pp. 100. World Scientific Pub Co Inc.

Harvey, A. C. (1998). *Long Memory in Stochastic Volatility*, Chapter 16 in 'Forecasting Volatility in the Financial Markets', pp. 307–320. Butterworth Heinemann.

Heber, G., A. Lunde, N. Shephard, and K. Sheppard (2009). Oxford-man institute's realized library, version 0.1.

Heynen, R. C. (1995). *Essays on derivatives pricing theory*. Ph. D. thesis.

Heynen, R. C. and H. M. Kat (1994). Volatility prediction: a comparison of the stochastic volatility, garch (1, 1) and egarch (1, 1) models. *The Journal of Derivatives 2*(2), 50–65.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The journal of finance 42*(2), 281–300.

Hull, J. and A. White (1988). An analysis of the bias in option pricing caused by a stochastic volatility. *Advances in futures and options research 3*(1), 29–61.

Hurst, H. E., R. P. Black, and Y. Simaika (1965). *Long-term storage: an experimental study*. Constable.

Hwang, S. and S. E. Satchell (1998). Forecasting volatility in the financial markets. pp. 193–225.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic processes and their applications 119*(7), 2249–2276.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business 36*.

Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional brownian motions, fractional noises and applications. *SIAM review 10*(4), 422–437.

Masoliver, M., G. Roig, and S. Singla (2014). Realized volatility estimation.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics 160*(1), 246–256.

Peters, E. E. (1994). *Fractal market analysis: applying chaos theory to investment and economics*, Volume 24. John Wiley & Sons.

Poon, S.-H. and C. W. Granger (2003). Forecasting volatility in financial markets: A review. *Journal of economic literature 41*(2), 478–539.

Shephard, N. and K. Sheppard (2010). Realising the future: forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics 25*(2), 197–231.

Yu, J. (2002). Forecasting volatility in the new zealand stock market. *Applied Financial Economics 12*(3), 193–202.

Zhang, L. et al. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli 12*(6), 1019–1043.

8 Appendix

Figure 1: Descriptive Statistics

| Stock Index	Min	Max	Quartile 1	Quartile 3	Mean	Median	Stdev	Skewness	Kurtosis	Obs	NAs
S&P 500	0.0002%	0.9313%	0.0026%	0.0109%	0.0111%	0.0052%	0.0255%	15.05	426.84	4521	198
FTSE 100	0.0004%	0.3256%	0.0022%	0.0088%	0.0084%	0.0041%	0.0152%	7.92	101.26	4521	175
Nikkei 225	0.0004%	0.3836%	0.0039%	0.0125%	0.0114%	0.0070%	0.0190%	8.66	110.24	4521	328
DAX	0.0003%	0.6432%	0.0049%	0.0180%	0.0177%	0.0093%	0.0314%	8.44	114.73	4521	141
Russel 2000	0.0000%	0.6425%	0.0032%	0.0104%	0.0108%	0.0055%	0.0217%	10.93	213.82	4521	195
All Ordinaries	0.0001%	0.1556%	0.0015%	0.0051%	0.0049%	0.0026%	0.0083%	7.46	84.95	4521	197
DJIA	0.0002%	0.9126%	0.0027%	0.0103%	0.0107%	0.0051%	0.0255%	15.15	407.24	4521	195
Nasdaq 100	0.0000%	0.6669%	0.0029%	0.0137%	0.0135%	0.0057%	0.0262%	8.97	146.32	4521	192
CAC 40	0.0004%	0.4551%	0.0043%	0.0161%	0.0145%	0.0083%	0.0231%	7.67	93.78	4521	114
Hang Seng	0.0004%	0.4461%	0.0032%	0.0092%	0.0088%	0.0052%	0.0162%	14.04	300.46	4521	527
KOSPI Composite	0.0004%	0.6479%	0.0029%	0.0148%	0.0133%	0.0067%	0.0238%	9.78	181.06	4521	261
AEX Index	0.0003%	0.3962%	0.0033%	0.0132%	0.0126%	0.0063%	0.0217%	6.52	66.01	4521	115
Swiss Market	0.0007%	0.2797%	0.0026%	0.0082%	0.0084%	0.0042%	0.0145%	7.16	81.32	4521	189
IBEX 35	0.0004%	0.4766%	0.0048%	0.0177%	0.0150%	0.0100%	0.0213%	8.04	110.57	4521	149
S&P CNX Nifty	0.0003%	1.0108%	0.0041%	0.0143%	0.0148%	0.0073%	0.0328%	13.86	313.12	4521	775
IPC Mexico	0.0003%	0.3010%	0.0024%	0.0069%	0.0066%	0.0038%	0.0112%	10.53	193.04	4521	194
Bovespa	0.0008%	0.8364%	0.0098%	0.0249%	0.0238%	0.0154%	0.0384%	9.44	131.82	4521	289
S&P/TSX Composite	0.0000%	0.3545%	0.0014%	0.0050%	0.0056%	0.0024%	0.0144%	10.99	175.26	4521	781
Euro STOXX 50	0.0000%	1.0922%	0.0047%	0.0174%	0.0165%	0.0089%	0.0320%	13.65	342.42	4521	138
FT Straits Times	0.0006%	0.2773%	0.0021%	0.0064%	0.0057%	0.0035%	0.0087%	11.98	280.22	4521	642
FTSE MIB	0.0003%	0.4304%	0.0037%	0.0148%	0.0129%	0.0074%	0.0199%	7.84	107.97	4521	157

19

Figure 2: QQ Plot for S&P 500
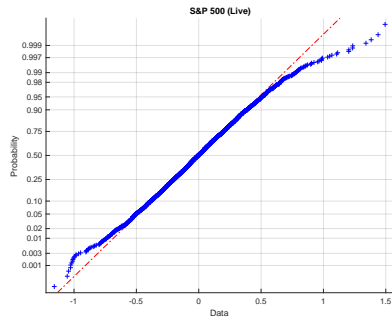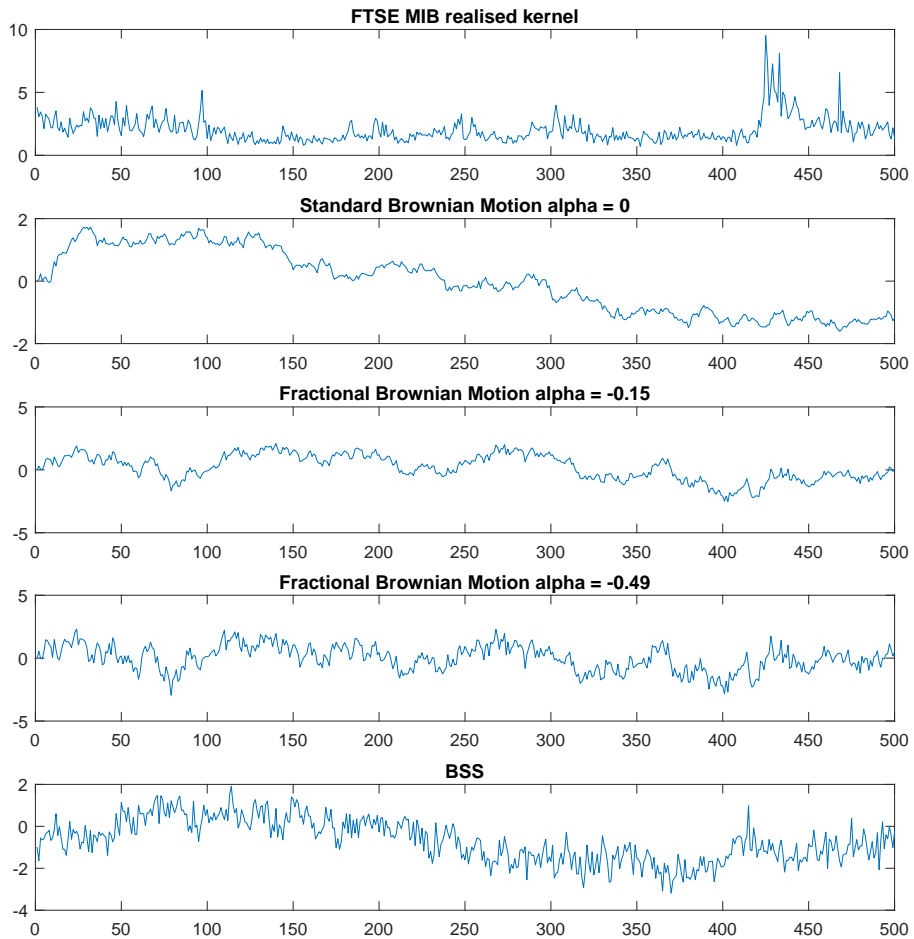

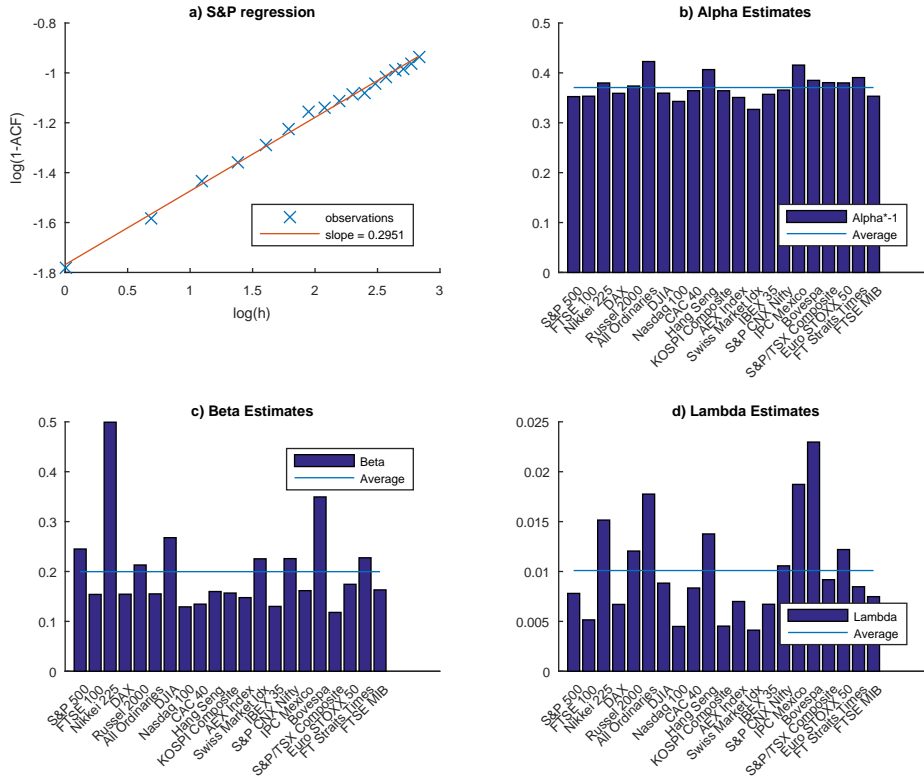Figure 3: Simulated Processes

Figure 4: Estimated Alphas

Figure 5: Estimated Alpha and Betas by geographical region of Stock Index
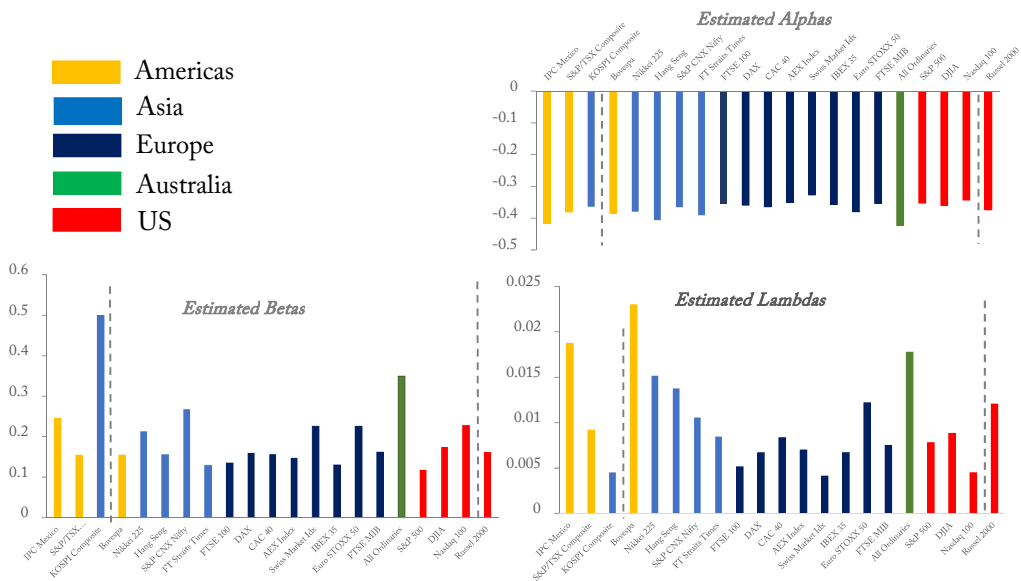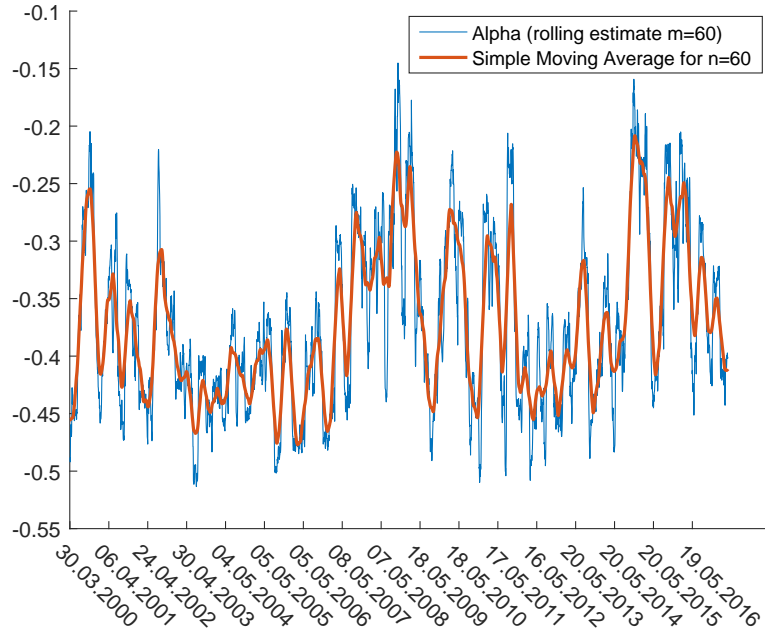
Figure 6: Rolling Alpha S&P 500

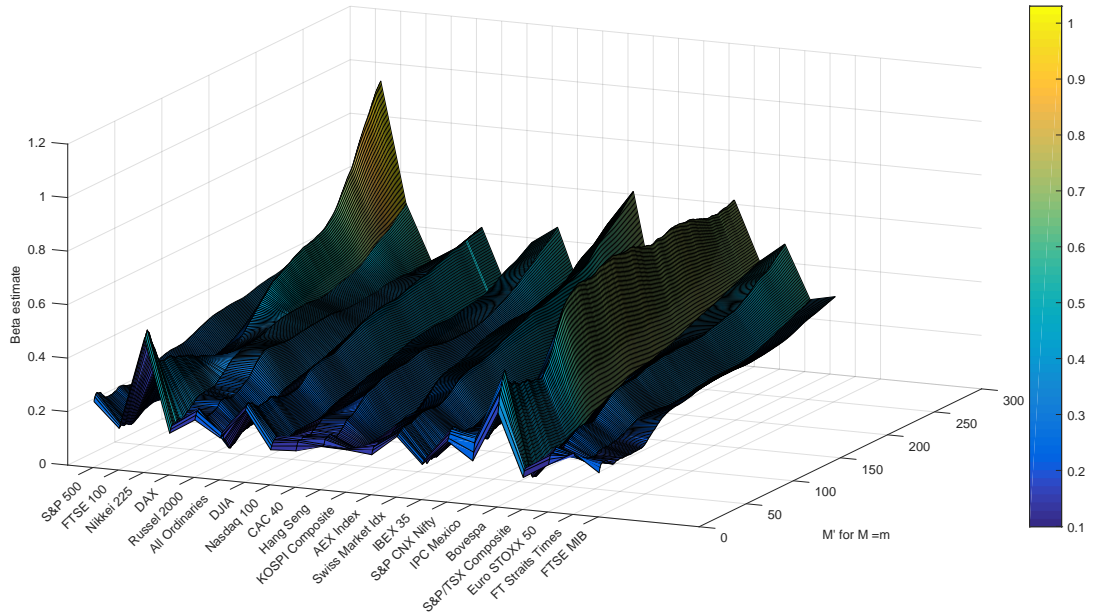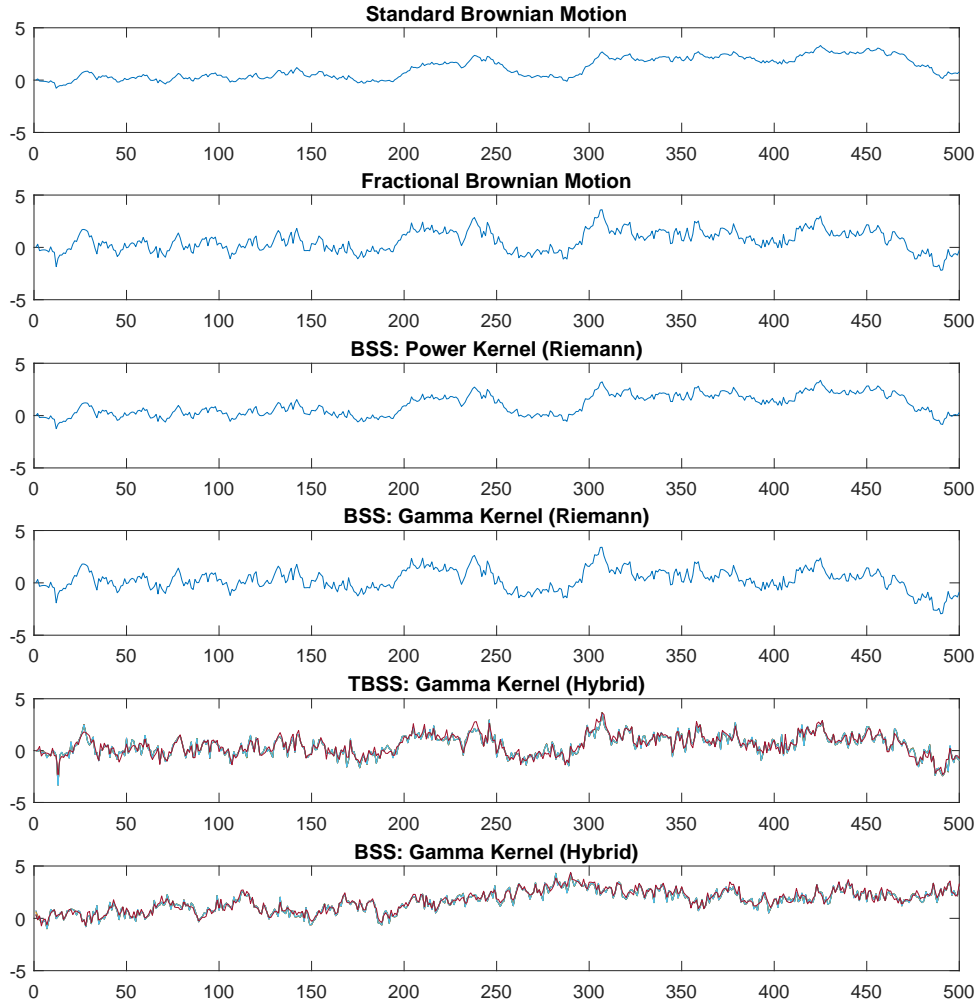Figure 7: Sensitivity Analysis Of the Beta Estimation

Figure 8: Estimated Processes: S&P 500

The red lines for the TBSS and BSS processes show the estimation results for $\kappa = 2$. The standard deviation for the processes has been standardized to one. The deviation between TBSS and BSS paths results from required re-sampling the random vector.

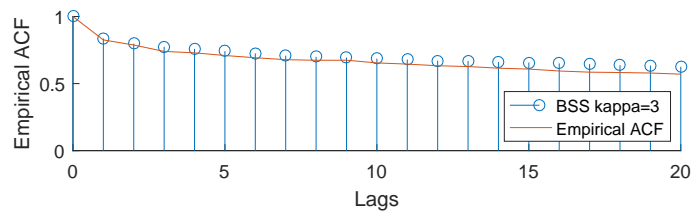Figure 9: Empirical vs. Simulated ACF
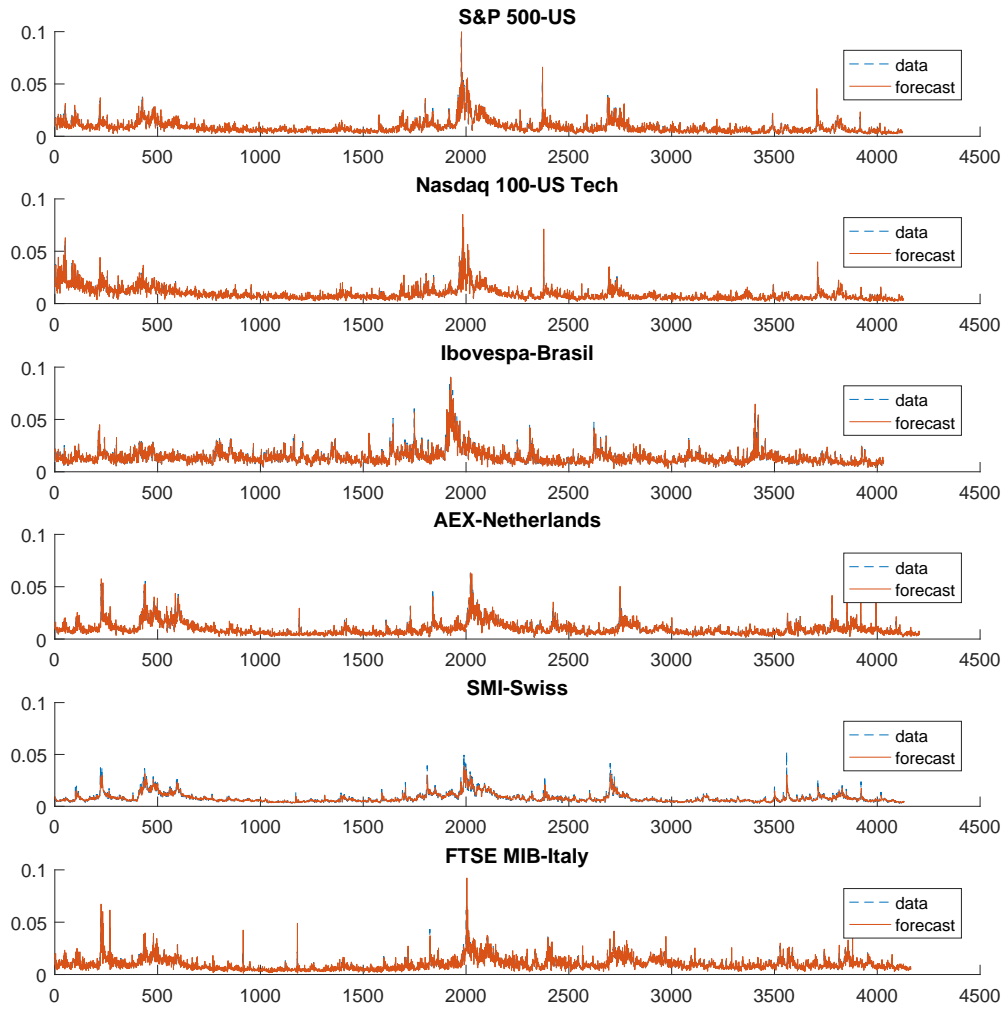
Figure 10: Forecast Performance

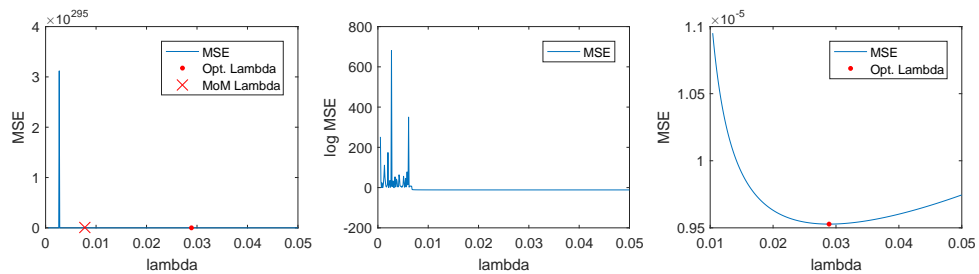Figure 11: Lambda Instability Analysis for the S&P 500

Table 1: Forecast Performance MSE and QL

Index	Metric	Step Ahead = 1				Step Ahead = 10			
		Roll Var	EWMA	Log-HAR	BSS	Roll Var	EWMA	Log-HAR	BSS
S&P 500	MSE 10^5	3.3097	1.6965	1.1210	1.5027	3.6087	2.5157	1.9031	1.8360
	QL	0.1130	0.0647	0.0470	0.0620	2.1791	0.0910	0.0831	0.0760
Nadaq	MSE 10^5	3.2007	1.8984	0.9613	1.5115	3.4306	2.5622	1.5480	1.4946
	QL	0.1048	0.0621	0.0382	0.0532	2.1275	0.0811	0.0666	0.0621
AEX	MSE 10^5	2.8307	1.4024	1.1152	1.4125	3.0980	2.3604	1.9472	1.8069
	QL	0.0929	0.0496	0.0413	0.0563	2.3901	0.0804	0.0724	0.0660
SMI	MSE 10^5	1.7776	0.9667	0.6381	0.5116	1.9616	1.6739	1.2159	1.0790
	QL	0.0781	0.0407	0.0272	0.0221	2.4126	0.0697	0.0533	0.0481
Bovespa	MSE 10^5	5.2725	3.1314	2.1369	2.5650	5.6765	4.4672	3.1737	3.2284
	QL	0.0821	0.0522	0.0415	0.0496	2.3689	0.0723	0.0612	0.0571
FTSE	MSE 10^5	2.9630	1.6479	1.1311	1.8223	3.1910	2.5347	1.9233	1.7963
MIB	QL	0.0961	0.0552	0.0410	0.0659	2.2655	0.0803	0.0700	0.0658

Table 2: Diebold and Mariano (1995) test results Benchmarks vs. BSS forecast errors

Index	Step Ahead = 1						Step Ahead = 10					
	MSE			QL			MSE			QL		
	Roll Var	EWMA	Log-HAR	Roll Var	EWMA	Log-HAR	Roll Var	EWMA	Log-HAR	Roll Var	EWMA	Log-HAR
SP 500	10.6881	1.4007	-2.9181	17.9422	1.3288	-7.9014	19.8982	6.9068	1.0980	22.2359	9.2236	4.5126
	(0.0000)	(0.0807)	(0.0018)	(0.0000)	(0.0920)	(0.0000)	(0.0000)	(0.0000)	(0.1361)	(0.0000)	(0.0000)	(0.0000)
Nasdaq	11.8380	2.8516	-4.4901	21.9667	5.0178	-9.3689	23.1087	10.2427	1.1575	26.0394	12.3021	3.4256
	(0.0000)	(0.0022)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.1235)	(0.0000)	(0.0000)	(0.0003)
AEX	12.5481	-0.1349	-3.7990	13.3000	-3.5122	-8.2258	21.6140	9.0583	3.0580	21.0672	10.0088	5.1308
	(0.0000)	(0.4463)	(0.0001)	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0011)	(0.0000)	(0.0000)	(0.0000)
SMI	19.9699	14.5622	5.4946	28.9355	20.6586	6.6659	24.2836	13.5029	4.9721	25.4787	17.9922	7.0082
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Bovespa	14.0738	4.5146	-3.6657	16.0881	1.7294	-5.5111	19.1422	7.3387	-0.4571	20.4261	11.1119	3.3737
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0419)	(0.0000)	(0.0000)	(0.0000)	(0.3238)	(0.0000)	(0.0000)	(0.0004)
FTSE	7.9916	-1.4069	-5.3118	10.2859	-4.6379	-10.9580	23.8457	12.3160	3.1201	20.3776	10.0739	3.6994
MIB	(0.0000)	(0.0797)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	'(0.0000)	(0.0000)	(0.0009)	(0.0000)	(0.0000)	(0.0001)

Test statistics, p-values in parenthesis. Null hypothesis is no difference between forecast models.