# Shifts in Volatility Driven by Large Stock Market Shocks

Yiannis Dendramis[†] George Kapetanios[*] Elias Tzavalis[†]

**Abstract**

This paper presents a new stochastic volatility model which allows for persistent shifts in volatility of stock market returns, referred to as structural breaks. These shifts are endogenously driven by large return shocks (innovations), reflecting large pieces of market news. These shocks are identified from the data as being bigger than the values of two threshold parameters of the model: one for the negative shocks and one for the positive shocks. The model can be employed to investigate economic (or market) sources of volatility shifts, without relying on exogenous information from the sample. In addition to this, it has a number of interesting features which enables us to study the dynamic or changing in magnitude effects of large return shocks on future levels of market volatility. The above properties of the model are shown based on a study for the US stock market volatility. For this market, the model identifies from the data as large negative return shocks these which are smaller than -2.05% on weekly basis, while as large positive return shocks those which are bigger than 2.33%.

*Keywords*: Stochastic volatility, structural breaks
*JEL Classification*: C22, C15

[*]Department of Economics, Queen Mary University of London, Mile End Road, London E1 4NS, UK. email: G.Kapetanios@qmul.ac.uk. Tel: +44(0)20 78825097.

[†]Department of Economics, Athens University of Economics and Business, Athens 104 34, Greece. email: etzavalis@aueb.gr.

# 1 Introduction

There is recently considerable evidence indicating the existence of discrete-time and persistent shifts in the conditional variance process of asset (stock) returns. These shifts, referred to as structural breaks in the empirical finance literature, are related to large negative, or positive, stock market return shocks, which reflect substantial pieces of market news (see Diebold and Pauly (1987), Lamourex and Lastrapes (1990), Tzavalis and Wickens (1995), Diebold and Inoue (2001), Andreou and Ghysels (2002), Mikosch and Starica (2004), Morana and Beltratti (2004), and Kramer and Tameze (2007), Rapach and Strauss (2008), *inter alia*). There are many economic reasons for which larger, in magnitude, return shocks can cause persistent shifts in the level of market volatility, which can change over time. These can be attributed to financial leverage effects or feedback volatility (risk premium) effects (see, e.g., French et al (1987), Schwert(1990), Campbell and Hentschel (1992), Bekaert and Wu (2000), or more recently, Mele (2007) and Ozdagli (2012)). If the above shifts in volatility are not accounted for, they will overstate evidence of very high persistence in volatility (see, e.g., Psaradakis and Tzavalis (1999) and, for a survey, Andreou and Ghysels (2009)). As more recently noted by Malik (2011), a smaller degree of persistency can dampen the feedback volatility effects faster, and thus positive (good) return shocks will result in significant drops of volatility.[1]

The empirical literature mentioned above treats large stock market return shocks as exogenous. To capture their effects on volatility, it relies on the intervention-dummy variable analysis of Box and Tiao (1975), based on out-of-sample information to determine the time points that the breaks driven by large shocks occur. Of course, more sophisticated multibreak testing procedures can be applied to find out from the data the timing of the breaks, like those employed for breaks in the mean of

[1] Campbell and Hentschel (1992) provide an analytic relashionships which show how the volatility persistency parameter determines the volatility feedback effects.

series (see, e.g., Bai and Perron (2003)). But, as in the intervention analysis, these methods do not treat the break process as an endogenous process, specified as a part of volatility model which is the focus of this paper. By doing this, volatility models can allow for richer dynamics which can help to identify different economic sources (or market events) of volatility shifts from the data and to study the dynamic effects of market return shocks on volatility functions. Separating the impact of these return shocks on volatility from those of ordinary return shocks can also have important implications for long-term portfolio management and hedging, as it will bring more focus on controlling important sources of risks caused by long-term shifts in volatility leaving aside its short-term ones. As shown by many studies based on intervention analysis, these type of shifts in volatility or stock prices comovements tend to be mainly driven by large market shocks (see Karolyi and Stulz (1996), or Chen et al (2003)).

In this paper we suggest a parametric model of breaks in volatility which are endogenously driven by large stock return shocks. This model allows us to capture persistent or cyclical shifts of large stock market pieces of news on volatility, without relying on any exogenous (out-of-sample) information. Depending on their sign, these shocks can be identified though our model by being larger (or smaller) than a positive (or negative) value of a threshold parameter. These values can be estimated from the data based on a search procedure. The different sign of the values of the threshold parameters that our model considers will enable us to investigate asymmetric effects of large return shocks on volatility. To study the dynamic effects of large shocks on volatility, the paper estimates generalized impulse response functions (GIRFs) of volatility with respect to these large shocks. These functions will allow us to study the above dynamic effects by integrating out any possible future or history effects of return shocks on volatility, which can affect the sample path of volatility.

Another interesting feature of our model is that it allows for shifts in volatility which are stochastic in both time and magnitude. The stochastic magnitude of volatility shifts considered by our model is consistent with evidence of clustering of stock returns of different magnitude, over time. This also distinguishes our model from other parametric threshold volatility models allowing for shifts in the volatility function parameters of fixed magnitude, see, e.g., Glosten's et al (1993) threshold GARCH model. Note that the last model also assumes known values of threshold parameters.

The econometric framework employed to build up our volatility model is that of the discrete-time SV model with leverage effects (see, e.g., Taylor (1986), Harvey et al (1994), and Harvey and Shephard (1994), *inter alia*). This model is extended to allow for a stochastic break process with the properties mentioned above. Since our model is nonlinear, to estimate its parameters and retrieve from the data its state variables, namely the stochastic volatility, the break process and a time-varying coefficient capturing stochastic changes in the magnitude of breaks, we use a Bayesian Markov Chain Monte Carlo (MCMC) method often employed in the literature to estimate SV models with leverage effects (see, e.g., Omori et al (2006)). This method has been extended to estimate threshold parameters endogenously from the data, through a grid search procedure. The paper shows that the above estimation procedure of the state variables and parameters of the model leads to consistent estimates of them. To evaluate the performance of this estimation procedure, the paper conducts a Monte Carlo exercise.

Implementation of our model to investigate if level shifts in the volatility of the US stock market aggregate return can be attributed to large stock market pieces of news (shocks), reflected in stock returns, leads to a number of interesting conclusions. First, it shows that, indeed, shifts in volatility can be triggered by large stock market return shocks. Most of these shifts are quite persistent and are due to negative large return

shocks, associated with financial crises. Our model identifies as large negative return shocks these with values less than -2.05% of weekly returns, while as large positive shocks those whose values are bigger than 2.33%. This asymmetry of the estimates of the threshold parameters reveals that market participants consider as large negative pieces of news return innovations of smaller magnitude than those corresponding to positive pieces of news. It can explain shapes of news impact functions based on implied values of volatility which are asymmetric towards large negative return innovations, as is observed in practice (see, e.g., Engle and Ng (1993) or , more recently, Ederington and Guan (2010)). Finally, the estimated GIRFs by the model clearly indicate that large positive return shocks imply drops in future stock market volatility, which are substantial. As those of large negative return shocks, which increase volatility, these effects are found to be quite persistent. This is in contrast to ordinary (small) return shocks, which are found to be smaller and to die out very fast.

The paper is organized as follows. Section 2 presents our model and discusses some of its main features. Section 3 presents the estimation method of the model. In Section 4, we report the results of a small Monte Carlo study, assessing the performance of the estimation method of the model to provide accurate estimates of its structural parameters and state variables. Section 5 presents the results of the empirical application of the model to the US stock market data. Apart from estimating the model, this section involves calculating its news impact and generalized impulse response functions, analyzing the dynamic effects of large return shocks on volatility. Finally, Section 6 concludes the paper.

# 2 Model specification

Consider the following stochastic volatility model of a stock return series at time $t$, $r_t$:

$$r_t - \mu = \exp\left(\frac{h_t}{2}\right)\varepsilon_t, \qquad t = 1, 2, ..., n \tag{1}$$

with stochastic volatility process

$$h_{t+1} = b_{t+1} + \phi h_t + \eta_t \tag{2}$$

and a break process, given as

$$b_{t+1} = b_t + I(\mathcal{A}_t)\gamma_t, \tag{3}$$

where $h_t$ is the logarithm of the conditional variance of return $r_t$, referred to as volatility, $\varepsilon_t$ and $\eta_t$ are the stock return and volatility innovations, respectively, distributed as

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim NID(\mathbf{0}, \mathbf{\Sigma}), \text{ with } \quad \mathbf{\Sigma} = \begin{pmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2 \end{pmatrix}, \tag{4}$$

where $\rho$ is the correlation coefficient between innovations $\varepsilon_t$ and $\eta_t$ capturing leverage effects, and $\gamma_t$ is a time-varying coefficient distributed as $\gamma_t \sim N\left(0, \sigma_\gamma^2\right)$.

Model (1)-(3) extends the standard stochastic volatility (SV) model with leverage effects to allow for discontinuous shifts in the level of volatility process $h_t$ of unknown time. These shifts, which are modelled through process (3), are driven by large return innovations $\varepsilon_t$, referred to as large return shocks (or news). They are identified by indicator function $\mathcal{I}(\mathcal{A}_t)$ taking the value 1 if the event $\mathcal{A}_t = \{\varepsilon_t > r^R \text{ or } \varepsilon_t > r^L\}$ occurs, and zero otherwise, where $\left(r^L, r^R\right)$ is a pair of threshold parameters which can be estimated based on sample information.[2] The events captured by set $\mathcal{A}_t$ can be thought of as reflecting large pieces of positive (or negative) stock market news when $\varepsilon_t > r^R$ (or $\varepsilon_t < r^L$), affecting stock market returns at time $t$. Since the breaks

[2] In the literature, these news are sometimes recognised as outliers in the level of series $y_t$ (see, e.g., Huang (2007)).

captured by the above SV model, defined by equations (1)-(3), are endogenously driven by stock return innovations $\varepsilon_t$, this model will be henceforth denoted with the acronym SVEB, which stands for stochastic volatility (SV) with endogenous breaks.

One interesting feature of the SVEB model is that the specification of break process $b_{t+1}$, governing level shifts in volatility $h_{t+1}$, allows for both the timing and the magnitude of these shifts to be stochastic in nature. The timing of a possible shift in $h_{t+1}$ is controlled by innovations $\varepsilon_t$ through indicator function $\mathcal{I}(\mathcal{A}_t)$, while its magnitude is determined by the time varying coefficient $\gamma_t$, which is a random variable distributed as $N\left(0, \sigma_\gamma^2\right)$. The stochastic nature of this coefficient enables for a more flexible approach of modelling cyclical shifts in volatility, relaxing the assumption that these shifts are of constant magnitude over time. The last assumption is made by existing threshold volatility models like that of Glosten et al (1993), or its extensions suggested by So et al (2003) and Smith (2009). Note that these threshold models have been introduced in the literature to capture possible asymmetries in volatility function $h_{t+1}$ (i.e., threshold effects) which can explain leverage effects. The SVEB model can be thought of as an extension of the stochastic permanent break model of Engle and Smith (1999) for the conditional mean of economic series, as was extended by Kapetanios and Tzavalis (2010) to allow for stochastic in magnitude shifts, to model structural breaks in volatility.

In addition to the above properties, the SVEB model has a number of other interesting features, which can be proven very useful in practice. First, by allowing the threshold parameters $r^L$ and $r^R$ to differ to each other, i.e. $r^L \neq r^R$, it can be employed to unveil from the data values of stock return innovations that are considered by market participants as large shocks. The threshold volatility models mentioned above treat these values of $r^L$ and $r^R$ as known, and set them to zero, i.e. $r^L = r^R = 0$. By allowing for different values of $r^L$ and $r^R$ both in terms of sign and magnitude,

the SVEB model can capture asymmetries of stock market news on volatility function beyond those implied by leverage effects of standard stochastic volatility models. As shown in the empirical section of the paper, these asymmetries can produce patterns of stock market news impact functions (NIFs) which are more close to reality. Finally, our model can be employed to study the dynamic effects of large negative, or positive, return shocks on the future path of volatility through generalized impulse response functions (GIRFs), which are net of the effects of possible future (or past) return shocks on volatility. The latter can obscure the true effects of large return shocks on volatility function. The same can be done for ordinary return shocks, defined as $r^R \leq \varepsilon_t \leq r^L$.

The SVEB model can nest different stochastic volatility models, which may be employed in practice. When $r^L = r^R = 0$, the model reduces to a version of SV model with time-varying coefficient (TVC) effects. This corresponds to the TVC model introduced by Harvey (see Harvey (1989)) for the mean of economic or financial series. This model assumes shifts in volatility at every period, which do not conform with the notion of structural breaks observed in stock return volatility series. When $\sigma_\gamma^2 = 0$, then volatility $h_{t+1}$ is driven by ordinary shocks $\eta_t$ and thus, the SVEB model reduces to the standard SV model with leverage effects, often used in practice. As it stands, the SVEB model can generate a non-stationary pattern for volatility process $h_{t+1}$, given that the variance of the process governing breaks $b_{t+1}$ grows with the time-interval of the data. If stationarity of volatility process $h_{t+1}$ is a desirable property or required by the data, then stationarity of break process $b_{t+1}$ would be required for this. There are a number of restrictions which can be imposed on $b_{t+1}$ to make this process stationary (see Cogley and Sargent (2002)). A straightforward one is the following:

$$b_{t+1} = \delta_t b_t + I(A_t)\gamma_t, \tag{5}$$

where

$$\delta_t = \begin{cases} 1 & \text{if } I(|b_t| < b) \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

This condition implies that $b_{t+1}$ is bounded by $b$ and, hence, it renders $h_{t+1}$ stationary, too.[3] In the next theorem, we prove that restriction (5) implies strict stationarity of $h_{t+1}$ provided that $|\phi| < 1$.

**Theorem 1** *If $|\phi| < 1$ and condition (5) hold, then $h_t$ is strictly stationary.*

The proof of the theorem is given in the Appendix.

# 3 Model Estimation

In this section we present the estimation procedure that we will employ to obtain sample estimates of the parameters of the volatility function of the SVEB model, collected in vector $\theta = (\phi, \sigma_\eta, \sigma_\gamma, \rho)'$, and its state variables, collected in vector $\mathbf{a}_t = (h_t, b_t, \gamma_t)'$. To this end, we will rely on the Bayesian MCMC estimation method suggested by Omori et al (2006). This method is extended to provide estimates of threshold parameters $r^L$ and $r^R$ endogenously from the data, based on a grid search procedure.

To implement the above estimation method, we will write the demeaned return process $y_t = r_t - \mu$, implied by equation (1), in terms of the following bivariate set of observations: $\{d_t, y_t^*\}$, where $d_t$ is the sign of $y_t$, and $y_t^* = \log y_t^2$, i.e.

$$y_t = d_t \exp\left(\frac{y_t^*}{2}\right), \;\; \varepsilon_t^* = \log \varepsilon_t^2 \;\; \text{and} \;\; y_t^* = \log y_t^2 = h_t + \varepsilon_t^*,$$

where $\varepsilon_t^*$ is a transformed $IID$ innovation process, which follows a $\log \chi_1^2$ distribution with one degree of freedom. This transformation of $y_t$ enables us to write the observation and the state equations of the SVEB model as follows:

[3] Further restrictions could be placed on the process $b_{t+1}$ so that, if the bound $b$ is exceeded, the process returns to some prespecified level. We do not advocate a particular mechanism for making the process $b_{t+1}$ stationary. We simply wish to indicate that there exist specifications which give a stationary $b_{t+1}$ process. The exact specification of the process may be left to the empirical researcher depending on their priors on the particular issue at hand.

$$y_t^* = h_t + \varepsilon_t^* \tag{7}$$

and

$$\begin{bmatrix} h_{t+1} \\ b_{t+1} \\ \gamma_{t+1} \end{bmatrix} = \begin{bmatrix} \phi & 1 & I(A_t) \\ 0 & 1 & I(A_t) \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h_t \\ b_t \\ \gamma_t \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \\ w_{t+1} \end{bmatrix}, \tag{8}$$

respectively, where $w_{t+1} \sim NIID\left(0, \sigma_\gamma^2\right)$. The steps taken to estimate the SVEB model given by observation and state equations (7) and (8), respectively, are presented below. These initially assume that the vector of threshold parameters $\mathbf{r} = \left(r^L, r^R\right)'$ is known.

The key feature of the MCMC method employed to estimate the SVEB model is to express the joint density of $\varepsilon_t^*$ and $\eta_t$ as a mixture of $K = 10$ normal distributions, with latent mixture component indicators denoted as $s_t \in \{1, 2, .., 10\}$, for $t = 1, 2, ..., n$. Conditioned on $s_t$, this method produces a model whose state space representation is linear and Gaussian, and thus it enable us to sample the posterior distribution of $\mathbf{s}_n = \{s_t\}_{t=1}^n$, and $\mathbf{a}_n = \{\mathbf{a}_t\}_{t=1}^n$, as well as that of the parameter vector $\theta$ based on the following MCMC scheme. This can be done by initializing vector $\mathbf{s}_n$ and then iterating the following steps to obtain posterior samples:

1. Draw $\mathbf{a}_n, \theta | \, \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{s}_n$ by
   (a) drawing $\theta | \, \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{s}_n$
   (b) drawing $\mathbf{a}_n | \, \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{s}_n, \theta$
2. Draw $\mathbf{s}_n | \, \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{a}_n, \theta$,

until convergence is achieved. To derive the posterior distribution $\theta | \, \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{s}_n$ based on the above algorithm, we will assume a prior distribution of $\theta$, denoted as $\pi(\theta)$, and will calculate the likelihood $g\left(\mathbf{y}_n^* | \, \theta, \mathbf{d}_n, \mathbf{s}_n\right)$ based on the approximation of the bivariate joint density $f\left(\varepsilon_t^*, \eta_t | d_t\right)$ by the following mixture of $K$=10 normal distributions

with mean $m_j$, variance $v_j^2$, denoted as $N\left(\varepsilon_t^*; m_j, v_j^2\right)$, for $j = 1, 2, ..., K = 10$:

$$f\left(\varepsilon_t^*, \eta_t | d_t\right) = f\left(\eta_t | \varepsilon_t^*, d_t\right) f\left(\varepsilon_t^*\right)$$
$$\simeq \sum_{j=1}^{10} p_j N\left(\eta_t; d_t \rho\sigma e^{\frac{m_j}{2}}\left(a_j + b_j\left(\varepsilon_t^* - m_j\right)\right), \sigma^2\left(1-\rho^2\right)\right) N\left(\varepsilon_t^*; m_j, v_j^2\right), \tag{9}$$

where $\{p_j,\ a_j,\ b_j\}_{j=1}^{10}$ constitutes the set of mixing parameters. These parameters are chosen to make the approximation of the true density $f\left(\varepsilon_t^*, \eta_t | d_t\right)$ as tight as possible.[4]

The approximating density given by (9) implies that the vector of innovations $\left(\varepsilon_t^*, \eta_t\right)'$ conditional on the mixture component indicator $s_t = j$ and the sign $d_t$, converges asymptotically to the following random vector:

$$\left\{ \begin{pmatrix} \varepsilon_t^* \\ \eta_t \end{pmatrix} \middle| d_t, s_t = j \right\} \overset{L}{=} \begin{bmatrix} m_j + v_j z_{1t} \\ d_t \rho\sigma\left(a_j + b_j v_j z_{1t}\right) \exp\left(\frac{m_j}{2}\right) + \sigma\sqrt{1-\rho^2} z_{2t} \end{bmatrix}, \tag{10}$$

for $j = 1, 2, ..., K$=10, where $z_{1t}$ and $z_{2t}$ are two independent normally distributed random variables with zero mean and unit variance, and " $\overset{L}{=}$ " signifies convergence in distribution. This result implies that we can write the observation and state equations of the SVEB model (see (7) and (8), respectively) in a conditionally Gaussian state space form, and thus, use the Kalman filter algorithm to compute the likelihood of density $g\left(\mathbf{y}_n^* | \theta, \mathbf{d}_n, \mathbf{s}_n\right)$.[5] In so doing, we will replace the innovations $\varepsilon_t$ of set $\mathcal{A}_t$ by

[4] Optimal values of $a_j$ and $b_j$, as well as of $p_j$ and $m_j, v_j$ for $j = 1, .., 10$, are reported by Omori et al (2006).

[5] In particular, this representation of equations (7) and (8) of the SVEB model are given as follows:

$$y_t^* = \mathbf{e}_t' \mathbf{a}_t + m_j + \mathbf{G}_t' \mathbf{u}_t$$
$$\mathbf{a}_{t+1} = \mathbf{T}_t \mathbf{a}_t + \mathbf{W}_t + \mathbf{H}_t \mathbf{u}_t,$$

where

$$\mathbf{a}_t = \left[h_t, b_t, \gamma_t\right]', \mathbf{e}_t = \left[1, 0, 0\right]',$$
$$\mathbf{T}_t = \begin{bmatrix} \phi & 1 & I\left(\mathcal{A}_t\right) \\ 0 & 1 & I\left(\mathcal{A}_t\right) \\ 0 & 0 & 0 \end{bmatrix}, \ \mathbf{H}_t = \begin{bmatrix} d_t \rho\sigma_1 \exp\left(\frac{m_j}{2}\right) b_j v_j z_{1t} & \sigma_\eta\sqrt{\left(1-\rho^2\right)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\gamma \end{bmatrix}$$
$$\mathbf{G}_t = \left[v_j, 0, 0, 0\right]', \ \mathbf{W}_t = \left[d_t \rho\sigma_\eta \exp\left(\frac{m_j}{2}\right) a_j, 0, 0\right]',$$
$$\mathbf{u}_t \sim NIID\left(\mathbf{0}, \mathbf{I}\right), \ \varepsilon_t = \exp\left(-\frac{\mathbf{e}_t' \mathbf{a}_t}{2}\right) y_t$$

their filtered estimates, denoted as $\varepsilon_{t|t}$, implied by the estimates of volatility state variable $h_t$, denoted as $h_{t|t}$, which are received by the Kalman filter, i.e.

$$\varepsilon_{t|t} = y_t \exp\left(-\frac{h_{t|t}}{2}\right). \tag{11}$$

The large shocks of return process $y_t$ can be easily obtained from filtered estimates $\varepsilon_{t|t}$ based on values of the vector of threshold parameters $\mathbf{r} = \left(r^L, r^R\right)'$.

The above linear Gaussian state space form of the SVEB model allows us to draw posterior samples from the density of $\mathbf{a}_n | \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{s}_n, \theta$ (step 1b of the MCMC algorithm) with the help of a simulation smoother algorithm (see De Jong & Shephard (1995)). For step 1a of the MCMC algorithm, we will draw samples from the density $\pi(\theta | \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{s}_n) \propto g(\mathbf{y}_n^* | \theta, \mathbf{d}_n, \mathbf{s}_n)\pi(\theta)$ using the Metropolis-Hastings algorithm with a proposal density based on the truncated Gaussian approximation of the posterior. To this end, we will first define estimator $\hat{\theta}$ which maximizes $g(\mathbf{y}_n^* | \theta, \mathbf{d}_n, \mathbf{s}_n)\pi(\theta)$. Then, we will generate a candidate $\theta^*$ from the following truncated normal distribution: $TN_R\left(\hat{\theta}, \left(-\frac{\partial^2 \log g(\mathbf{y}_n^* | \theta, \mathbf{d}_n, \mathbf{s}_n)\pi(\theta)}{\partial\theta\partial\theta'}\Big|_{\theta=\hat{\theta}}\right)^{-1}\right)$, where $R = \{|\phi| < 1,\ \sigma_\eta, \sigma_\gamma > 0,\ |\rho| < 1\}$, and we will accept this candidate with the Metropolis-Hastings probability of move.[6]

Regarding the mixture indicator variable $s_t$ (step 2 of the MCMC algorithm), we will use the inverse transform method to sample from the following posterior:

$$\pi(\mathbf{s}_n | \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{a}_n, \theta) \propto \Pr(s_t = j)\, N\left(y_t^* - h_t; m_j, v_j^2\right)$$
$$N\left(h_{t+1} - \phi h_t - b_t - I(\mathcal{A}_t)\gamma_t; d_t\rho\sigma e^{\frac{m_j}{2}}\left\{a_j + b_j\left(y_t^* - h_t - m_j\right)\right\}, \sigma_\eta^2\left(1-\rho^2\right)\right),$$

for all $j$.

Since the MCMC method presented above involves a small approximation error due the sampling variability of $\mathbf{s}_n$, to control this we will reweight the samples of

[6] Another option is to transform the parameter vector $\theta$ to support the $R$ plane. In this case, we do not need to use the truncated normal distribution as the proposal density.

$\left\{h_t^k, b_t^k, \gamma_t^k, \theta^k\right\}_{k=1}^M$ obtained through the MCMC iterations, denoted as $k = 1, 2, ..., M$, using the following weights:

$$w_k^* = \prod_{t=1}^{n} \frac{f\left(\varepsilon_t^{*k}, \eta_t^k \middle| d_t, \theta^k\right)}{\tilde{f}\left(\varepsilon_t^{*k}, \eta_t^k \middle| d_t, \theta^k\right)}, \text{ for } k = 1, 2, .., M$$

where $f(.)$ is the true density of $(\varepsilon_t^*, \eta_t)'$, $\tilde{f}(.)$ is the approximation of $f(.)$, given by equation (9), $\varepsilon_t^{*k} = y_t^* - h_t^k$ and $\eta_t^k = h_{t+1}^k - b_t^k - I\left(A_t^k\right)\gamma_t^k - \phi^k h_t^k$.

As in other threshold models with unknown values of threshold parameters (see, e.g., Chan (1993)), to estimate $\mathbf{r}$ we will adopt a grid search procedure over a range of possible values of it. According to this method, the parameter vector $\theta$, and the marginal likelihood, denoted as $\ln m(y|\mathbf{r})$, are estimated at different values of $\mathbf{r}$. Then, the value of $\mathbf{r}$ that gives the maximum $\ln m(y|\mathbf{r})$ over this grid will be considered as its optimum sample estimate. The estimates of $\theta$ and state vector $\mathbf{a}_t$ corresponding to this value of $\mathbf{r}$ will constitute the maximum likelihood estimates of the model. These estimates will be consistent provided that vector $\mathbf{r}$ is also consistently estimated. The consistency of $\mathbf{r}$ based on grid search estimation method can be proved following analogous arguments to that of Kapetanios and Tzavalis (2010). Below, we present a useful practical remark for the estimation of thresholds' vector $\mathbf{r}$.

**Remark 1**: Since estimation of threshold parameters is problematic in small samples in general (see, e.g., Kapetanios (2000)) and since this problem may be exacerbated by the rarity of breaks, the above grid search can be considerably simplified if we consider values of vector $\mathbf{r}$ which correspond to extreme quantiles of the normalized error of (1), $\varepsilon_t$, such as its 97.5th or 99.0th percentiles.

In the above estimation procedure, to calculate marginal likelihood $m(y|\mathbf{r})$ (or its logarithmic value $\ln m(y|\mathbf{r})$) at the estimate $\theta^*$ we will employ the auxiliary particle filter algorithm (see, e.g., Pitt and Shephard (1999) and Omori et al (2006)) to obtain a value of the likelihood ordinate of the SVEB model, denoted as $g(y|\,\theta^*, \mathbf{r})$, and

Chib's and Jeliaskov (2001) method to calculate the posterior ordinate, $\pi(\theta^*|\,\mathbf{y},\mathbf{r})$.[7] The marginal likelihood $m(y|\mathbf{r})$ can be calculated from $g(\,y|\,\theta,\mathbf{r})$ using Bayes' theorem, i.e.

$$m(y|\mathbf{r}) = \frac{g(\mathbf{y}|\theta^*,\mathbf{r})\,\pi(\theta^*)}{\pi(\theta^*|\,\mathbf{y},\mathbf{r})}, \tag{12}$$

where $\pi(\theta^*|\,y,\mathbf{r})$ and $\pi(\theta^*)$ are the posterior and prior ordinates, respectively. Note that the above expression of marginal density $m(y|\mathbf{r})$ holds for any vector $\theta$, but it is generally considered as being more efficiently estimated when it is calculated at a high mass point like the estimate $\theta^*$.

Apart from calculating marginal likelihood $m(y|\mathbf{r})$, the auxiliary particle filter algorithm can be employed to provide forecasted and filtered values of the state vector $\mathbf{a}_{t+1}$ based on information sets $I_t = \{y_t, y_{t-1}, ...y_1\}$ and $I_{t+1}$, respectively, denoted as $\mathbf{a}_{t+1|t}$ and $\mathbf{a}_{t+1|t+1}$. These can be employed to calculate goodness of fit or forecasting performance measures of the SVEB model. These measures are also useful for model comparison.

# 4 Monte Carlo Study

In this section, we carry out a small scale Monte Carlo study to investigate the performance of the Bayesian MCMC method presented in the previous section to estimate the parameters of the SVEB model and the vector of its state variables. Since the main aim of our Monte Carlo exercise is to assess the performance of the estimation method to filter from the data estimates of the volatility variables $h_t$ and $b_t$, we concentrate on the estimation of the vector of state variables $\mathbf{a}_t = (h_t, b_t, \gamma_t)'$ and the vector of parameters $\theta$, while we treat the threshold vector $\mathbf{r} = (r^L, r^R)$ as known.

[7] This is a simulation based algorithm which has been suggested by Kitigawa (1996) for nonlinear non-Gaussian state space models and it has been successfully applied to evaluate SV models by Berg et al (2004).

In our Monte Carlo experiments, we generate samples of size $n = 1500$ observations, according to model (1)-(3) considering the following values for its structural parameters: $\phi = 0.85$, $\sigma_\eta = 0.28$, $\sigma_\gamma = 0.05$, $\rho = -0.59$. These values correspond to the estimates of the parameters of the SVEB model reported in the empirical section of our paper. To set initial values for state vector $\mathbf{a}_t$, we assume that its initial conditions are given as $\mathbf{a}_0 = (\phi/(1-b_0), b_0, 0)'$, where $b_0 = 0.10$. For the vector of threshold parameters, $\mathbf{r}$, we consider the following values: $\left(r^L, r^R\right) = (-1.96, 2.05)$. These imply 38 negative and 30 positive large shocks for our sample.

In the MCMC method of sampling the posterior distributions, we draw 6000 samples discarding the initial 500 variates. As priors of the parameters of the SVEB model, we use the following: $\frac{\phi+1}{2} \sim Beta(20, 1.5)$, $\frac{1}{\sigma_\eta^2} \sim Gamma(4.5, 0.15)$, $\frac{1}{\sigma_\gamma^2} \sim Gamma(1.5, 0.005)$, $\rho \sim U(-1, 1)$. These distributions are often used in the literature estimating SV models based on Bayesian methods (see, e.g., Kim et al (1997)). In total, we perform $M$=10 experiments. In Table 1, we report average estimates of the correlation coefficients, denoted $Corr(.,.)$, of the generated state variables $h_t$ and $b_t$ with their smoothed estimates, denoted as $h_{t|n}$ and $b_{t|n}$, respectively, over the above set of experiments. The table also reports average estimates of the mean and variance values of posterior distributions of the structural parameters of the SVEB model.

To better see how closely the suggested Bayesian MCMC method can capture shifts in state variables $h_t$ and $b_t$, in Figure 1 we graphically present the smoothed estimates of $h_{t|n}$ and $b_{t|n}$ against the generated values of them. The estimates of $h_{t|n}$ and $b_{t|n}$ correspond to those with the highest value of correlation coefficient $Corr(h_t, h_{t|n})$, among all experiments conducted. Together with the above graphs, Figure 1 also presents a plot of a return process $r_t$ generated by our SVEB model.

| | True value | Mean | Variance |
|---|---|---|---|
| $\phi$ | 0.85 | 0.8664 | 0.0067 |
| $\sigma_\eta$ | 0.28 | 0.2832 | 0.0128 |
| $\sigma_\gamma$ | 0.05 | 0.0485 | 0.0010 |
| $\rho$ | $-0.59$ | -0.5803 | $0.13\times10^{-2}$ |
| $b_0$ | 0.10 | 0.0224 | $2.15\times10^{-4}$ |
| | | | |
| $Corr(h_t, h_{t\|n})$ | | 0.9453 | $0.15\times10^{-4}$ |
| $Corr(b_t, b_{t\|n})$ | | 0.9638 | $0.17\times10^{-4}$ |

**Table 1:** Monte Carlo results

The results of the table and figure clearly indicate that the suggested Bayesian MCMC method can very efficiently estimate the parameters and state variables of the SVEB model. The patterns of state variables $h_t$ and $b_t$ generated by the SVEB model are cyclical and persistent, while their magnitude change considerably over the sample. These features of $h_t$ and $b_t$ can explain clusters of volatility of asset returns of different size over time, as that implied by the top plot of the figure, which graphically presents a typical return series implied by the SVEB model, and those often observed in practice. See also our empirical analysis, in Section 5. As can be confirmed by Figure 1, this cyclical and stochastic pattern of volatility process $h_t$ can be efficiently captured by break process $b_t$, which can be efficiently estimated by our suggested Bayesian procedure.

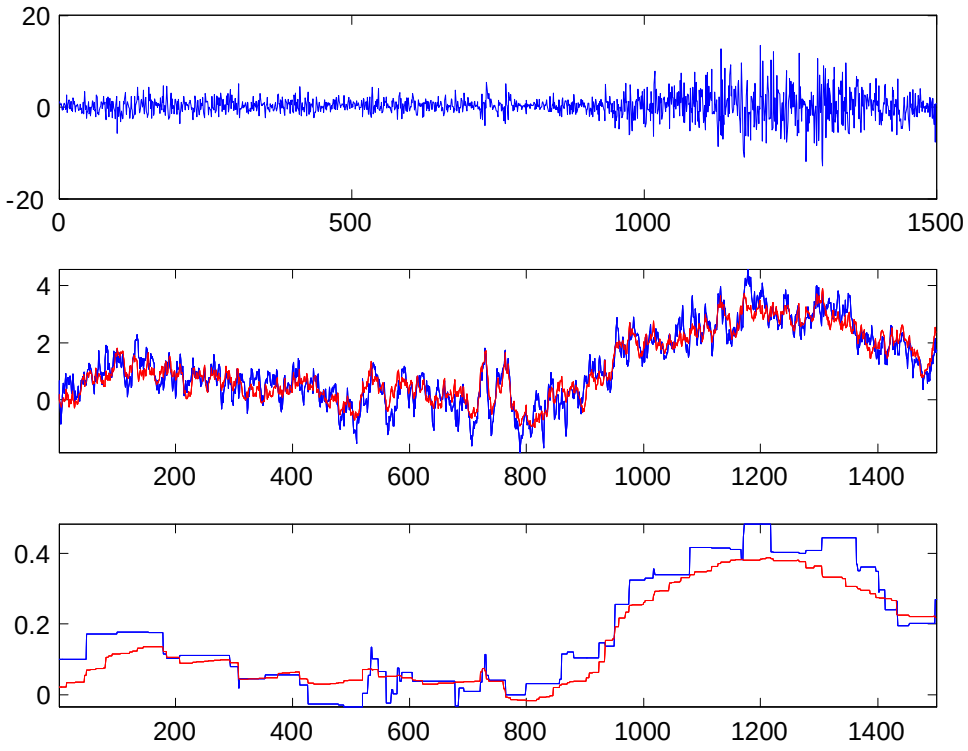


Figure 1: Smoothed estimates of volatility and break processes $h_t|n$ (see middle plot) and $b_t|n$ (see bottom plot), respectively, against their data generated values and stock return $r_t$ (see top plot), implied by the SVEB model.

# 5 Are volatility shifts of the US stock market driven by large shocks?

In this section, we estimate the SVEB model introduced in the previous sections based on US stock market aggregate return to address the following questions: Are shifts in volatility of the US stock market triggered by large return shocks? If yes, what is the magnitude of these large shocks? Are the effects of large return shocks on the volatility of this market asymmetric? How long do they last? Do large positive return shocks decrease, or increase, stock market's future volatility?

The above questions have been of interest to many studies (see, e.g. Engle and Ng (1993), Kane et al (2000) and Li et al (2005), or the related references mentioned in the introduction). Our model can help to provide clear cut answers to these questions since it can identify large negative, or positive, return shocks endogenously from the data. To calculate the US stock market return, we are based on weekly data of the S&P500 index covering the period from the 7th of January 1980 to the 29th of March 2010. Note that, during this period, many extraordinary events occurred in the US stock market which may have caused shifts on its volatility function. Some of them may be associated with financial crises in international stock markets. Examples of such events include the US financial crises of years 1987 and 2008-2009, the Mexican and Argentinian Crises of years 1994-1995, and the East Asian and Russian crises of years 1997-1998.

Table 2 presents estimates of the SVEB model based on the Bayesian MCMC estimation procedure suggested in Section 3. This is done for all possible combinations of the following sets of threshold parameters $r^L$ and $r^R$ considered: $r^L = \{-1.65, -1.96, -2.05, -2.33\}$ and $r^R = \{1.65, 1.96, 2.05, 2.33\}$[8]. The above set of $r^L$ and $r^R$ values correspond to negative and positive shock market events $\left(\frac{\varepsilon_t}{\sigma_t} < r^L\right)$ and $\left(\frac{\varepsilon_t}{\sigma_t} > r^R\right)$, where $\sigma_t = \exp\left(\frac{h_t}{2}\right)$, with probabilities to occur 5%, 2.5%, 2% and 1%, respectively. The bottom line of the table presents estimates of the standard SV model, which does not consider structural breaks. The priors of the parameters of the SVEB model considered in the estimation procedure are the same to those assumed by our Monte Carlo study. Analogous priors are used in the literature for the estimation of the SV model. The logarithms of the marginal likelihood values reported in the table, i.e. $\ln m(y|\mathbf{r})$, indicate that, among all possible combinations of $r^L$ and $r^R$ considered, the SVEB model with threshold parameter values $r^L = -2.05$

[8]Note that the estimates of the parameters and state variables of the SVEB model are based on 13000 draws from the posterior discarding the first 1000 and using only the half of them to avoid any problem of serial correlation.

and $r^R = 2.33$ constitutes the best specification of the data, as it gives the highest value of $\log m(y|\mathbf{r})$. This model is also found to describe better the data than the SV model. This can be also confirmed by the values of log-likelihood ordinate $\ln g\left(y|\theta,\mathbf{r}\right)$, reported in the table.

| $r^L$ | $r^R$ | $\phi$ | $\sigma_\eta$ | $\sigma_\gamma$ | $\rho$ | $b_0$ | $\ln g\left(y\|\theta,\mathbf{r}\right)$ | $\ln m(y\|\mathbf{r})$ |
|---|---|---|---|---|---|---|---|---|
| -1.65 | 1.65 | 0.7996 | 0.3000 | 0.0501 | -0.62 | 0.3262 | -3357.7 | -3369.5 |
| -1.65 | 1.96 | 0.8156 | 0.2970 | 0.0510 | -0.616 | 0.3151 | -3358.2 | -3368.7 |
| -1.65 | 2.05 | 0.8207 | 0.2966 | 0.0514 | -0.6144 | 0.3080 | -3360.5 | -3370.3 |
| -1.65 | 2.33 | 0.8286 | 0.299 | 0.0511 | -.6019 | 0.2948 | -3357.6 | -3366.0 |
| -1.96 | 1.65 | 0.8063 | 0.3055 | 0.0563 | -0.6247 | 0.3042 | -3356.1 | -3364.2 |
| -1.96 | 1.96 | 0.8312 | 0.2950 | 0.0575 | -0.6025 | 0.3279 | -3361.6 | -3367.9 |
| -1.96 | 2.05 | 0.8363 | 0.2902 | 0.0581 | -0.6164 | 0.2745 | -3359.4 | -3367.5 |
| -1.96 | 2.33 | 0.8683 | 0.2798 | 0.0504 | -0.5855 | 0.2196 | -3359.4 | -3365.5 |
| -2.05 | 1.65 | 0.8193 | 0.2966 | 0.0548 | -0.6259 | 0.2829 | -3357.6 | -3367.7 |
| -2.05 | 1.96 | 0.8428 | 0.2835 | 0.0572 | -0.6174 | 0.2677 | -3362.1 | -3370.2 |
| -2.05 | 2.05 | 0.8437 | 0.2875 | 0.0578 | -0.6069 | 0.2600 | -3362.5 | -3370.4 |
| -2.05 | 2.33 | 0.8772 | 0.2724 | 0.0506 | -0.5837 | 0.2047 | -3357.1 | -3363.1 |
| -2.33 | 1.65 | 0.8336 | 0.2876 | 0.0546 | -0.6125 | 0.2672 | -3366.6 | -3375.7 |
| -2.33 | 1.96 | 0.8568 | 0.2745 | 0.0600 | -0.6121 | 0.2394 | -3365.0 | -3372.4 |
| -2.33 | 2.05 | 0.8623 | 0.2725 | 0.0611 | -0.6119 | 0.2290 | -3360.0 | -3366.8 |
| -2.33 | 2.33 | 0.8989 | 0.2569 | 0.0518 | -0.5703 | 0.1600 | -3358.3 | -3364.0 |
| SV | | 0.9561 | 0.2230 | | -0.4911 | 0.062 | -3359.1 | -3363.3 |

**Table 2:** Estimates of SVEB model for different values of $\mathbf{r}$=$(r^L,r^R)'$

Table 3 reports more estimation results for the SVEB model with threshold parameters $r^L = -2.05$ and $r^R = 2.33$, while Figure 2 presents smoothed estimates of the volatility and break processes $h_{t|n}$ and $b_{t|n}$, respectively. These estimates are plotted together with return series $r_t$ and ex-post filtered estimates of return innovations, calculated as $\varepsilon_{t|t} = y_t \exp(-\frac{h_{t|t}}{2})$, where $h_{t|t}$ is obtained through the auxiliary particle filter method. This table also presents estimates of the standard SV model and values of the correlation coefficients between future levels of volatility $h_{t+s|n}$, for $s = 1, 2, ...,4, 5$ and return shocks $\varepsilon_{t|t}$, denoted Corr($h_{t+s|n}$,$\varepsilon_{t|t}$), retrieved by the estimates of our SVEB model. This is done for negative large and positive large return shocks, as well for ordinary return shocks, defined $-2.05 \leq \varepsilon_{t|t} \leq 2.33$. Such co-

efficients are frequently used in the literature to examine the dynamic relationship between future levels of volatility and current return shocks (see, e.g., Low (2004) and Denis et al. (2006), Malik (2011)).

In addition to posterior means of the parameters of the model, note that Table 3 reports posterior standard deviations of the structural parameters and the inefficiency factor (IF), for the estimates of both the SVEB and SV models. The inefficiency factor is defined as $1 + \sum_{s=1}^{\infty} \rho_s$, where $\rho_s$ is the sample autocorrelation at lag $s$. It is calculated from the sampled values of the parameters and it shows how well the Markov chain mixes. The values of IF reported in the table are very small, which means that the MCMC mixes very well. Thus, the structural parameters estimated by the Bayesian method can be thought of as being very accurately estimated. This can be also confirmed by the very small values of the standard deviations reported in the table, for all structural parameters of the SVEB model estimated.

| Estimates of the SVEB model ( $r^L = -2.05$ and $r^R = 2.33$) | | | | | |
|---|---|---|---|---|---|
| | $\phi$ | $\sigma_\eta$ | $\sigma_\gamma$ | $\rho$ | $b_0$ |
| Mean | 0.88 | 0.273 | 0.051 | -0.584 | 0.205 |
| Std. Dev. | 0.001 | 0.002 | 0.0003 | 0.007 | 0.006 |
| IF | 1.254 | 8.307 | 13.679 | 7.638 | 1.254 |
| Corr($h_{t+\tau\|n}$,$\varepsilon_{t\|t}$) | $\tau = 1$ | $\tau = 2$ | $\tau = 3$ | $\tau = 4$ | $\tau = 5$ |
| $\varepsilon_{t\|t} < -2.05$ | -0.40 | -0.28 | -0.34 | -0.28 | -0.38 |
| $\varepsilon_{t\|t} > 2.33$ | -0.02 | -0.09 | -0.0009 | -0.11 | -0.02 |
| $-2.05 \leq \varepsilon_{t\|t} \leq 2.33$ | -0.21 | -0.17 | -0.16 | -0.14 | -0.12 |
| Estimates of the SV model ($\sigma_\gamma^2 = 0$) | | | | | |
| | $\phi$ | $\sigma_\eta$ | $\sigma_\gamma$ | $\rho$ | $b_0$ |
| Mean | 0.95 | 0.223 | | -0.491 | 0.062 |
| Std. Dev. | 0.002 | 0.001 | | 0.006 | $3.27\times10^{-4}$ |
| IF | 1.795 | 4.104 | | 2.866 | 1.795 |

**Table 3:** Estimates of the SVEB and SV models

A number of interesting conclusions can be drawn from the results of Tables 2-3 and Figure 2. First, the estimates of threshold parameters of the SVEB model which are found to better describe the data are $r^L$=−2.05% and $r^R$ = 2.33% on

weekly return basis, which indicate that there are significant asymmetries between the large negative and positive return shocks. These asymmetries mean that market participants consider as large negative pieces of news return innovations of smaller magnitude than those corresponding to positive pieces of news. This result implies that level shifts in volatility can be triggered by negative return innovations which are smaller in magnitude than positive return shocks. Obviously, this should not be taken as a surprise. It can be attributed to stock market participants' attitude to give more weight to negative return news than positive due to their risk aversion behavior. The effects of the above asymmetries of threshold parameters $r^L$and $r^R$ on volatility function can be seen more clearly in a next section, which presents estimates of the impact news function of the SVEB model compared to those implied by the standard SV model.

Second, the allowance for a break process $b_t$ in the volatility function results in a drop of the autoregressive coefficient $\phi$ from 0.95 to 0.88. However, the estimate of $\phi$ is still quite high, which implies that the effects of return shocks on future volatility will last for a substantial number of periods ahead. However, the correlation coefficients between $h_{t+\tau|n}$ and $\varepsilon_{t|t}$, for $\tau = 1, 2, ...5$, reported in the table, are small. For the positive large return shocks, they are almost equal to zero. The latter is consistent with evidence in the literature based also on implied measures of volatility (see, e.g., Fleming et al (1995), Denis et al (2006), Malik (2011)). It may be attributed to the fact that correlation coefficients Corr($h_{t+\tau|n}$,$\varepsilon_{t|t}$) constitute crude measures of the dynamic relationship between $h_{t+\tau|n}$ and $\varepsilon_{t|t}$, which are not net of future (or lasting past) effects of different sign large or ordinary return shocks $\varepsilon_{t|t}$ on future levels of volatility $h_{t+\tau|n}$. Thus, to obtain a clear cut picture of the dynamics effects of $\varepsilon_{t|t}$ on $h_{t+\tau|n}$, in Section 5 we will present the generalized impulse response functions of $h_{t+\tau|n}$ with respect to $\varepsilon_{t|t}$, based on the estimates of our SVEB model.

Regarding the estimates of correlation coefficient $\rho$, which captures leverage effects of stock return innovations on volatility, and variance $\sigma_\gamma^2$, which allows for different over time magnitude effects of large return shocks on volatility, the results of Tables 2-3 reveal that the estimates of these parameters are robust across the different values of the vector of threshold parameters $\mathbf{r}=(r^L,r^R)$ considered. The robust estimate of $\sigma_\gamma^2$ across the different values of $\mathbf{r}$ assumed emphasizes the existence of shifts in volatility of stochastic magnitude over time. The slightly higher estimate of $\rho$ found by the SVEB model, compared to that estimated by the standard SV model, can be attributed to the ability of this model to generate leverage effects through two different sources: the relationship between innovations $\eta_t$ and $\varepsilon_t$, and the level shifts in volatility caused by the large return shocks, as was discussed before. The latter can amplify leverage effects through the state variable $\gamma_t$, and thus can explain higher values of coefficient $\rho$.

Finally, turning into the discussion of the results of Figure 2, one can conclude by inspecting the plots of this figure, that the estimates of volatility $h_{t|n}$ predicted by the SVEB model exhibit persistent and non-linear shifts over time. These shifts clearly have different magnitude over time. As was expected, they can be captured by those implied by the break process $b_t$.[9] The values of return series $r_t$ and the estimates of return shocks $\varepsilon_{t|t}$, presented at the bottom of the figure, indicate that some of the most important changing points of the observed cyclical shifts of both $b_{t|n}$ and $h_{t|n}$ are closely related to sequences of large return shocks associated with financial market crises, like those of US stock markets of years 1987 and 2008-2009, and the Mexican and Argentinian crises of years 1994-1995 followed by those of East-Asian and Russian of years 1997-1998. These crises seem to cause persistent and upward sloping shifts in volatility. Finally, comparison of the estimated break process $b_{t|n}$ with return series $r_t$

[9]Note that, as was expected, these plots of $h_{t|n}$ and $b_{t|n}$ are consistent with those generated by our Monte Carlo analysis (see, e.g. Figure 1).

reveals that the SVEB model can indeed capture shifts of volatility of different size, observed over time.

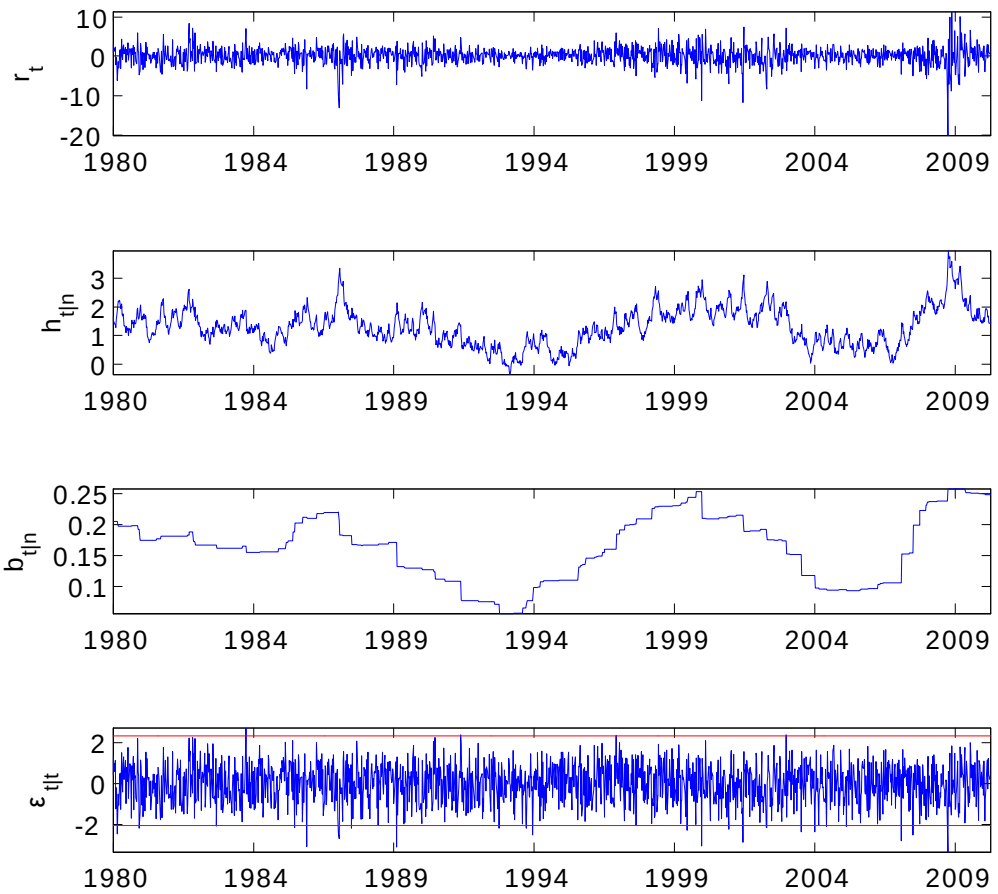


Figure 2: Smoothed estimates of processes $h_t$ (top plot) and $b_t$ (middle plot), and filtered estimates of $\varepsilon_t$ (denoted $\varepsilon_t|t$ (bottom plot)), for the S&P500

## 5.1 Evaluating the forecasting performance of the SVEB model

To further assess the ability of the SVEB model to fit satisfactorily the data compared to the SV or other parametric volatility models, in this section we conduct an in-sample forecasting exercise for the one-period ahead volatility $h_{t+1}$. Our forecasting exercise is focused on examining if $t$-time conditional estimates of volatility at time

$t$, defined as $h_{t+1|t} \equiv E\left(h_{t+1}|I_t\right)$, can successfully forecast the logarithm of squared demeaned return $\log y_{t+1}^2$, which is an observed variable as shown by observation equation (7).

Table 4 presents the results of the above exercise. In particular, the table reports values of the mean square error (MSE) of $h_{t+1|t}$ to forecast $\log y_{t+1}^2$ implied by equation (7), as well as estimates for this equation (see the table) based on the following linear regression:

$$\log y_{t+1}^2 = a + bh_{t+1|t} + u_t, \tag{13}$$

where $u_t = \left(\log\left(\varepsilon_t^2\right) - (-1.27)\right)$ and $E(\log\left(\varepsilon_t^2\right))$ =-1.27. The intercept and slope estimates of this forecasting regression can be used to test if they correspond to their theoretical values implied by (7), given as $a = -1.27$ and $b = 1.0$ (see, e.g., Pagan and Schwert (1990)). If this is true, then the SVEB model can be thought of as providing level and slope forecasts of the logarithm of squared demean return $\log y_{t+1}^2$ which are in the right direction. Since regressor $h_{t+1|t}$ is generated by the estimates of the volatility models, the estimates of coefficients a and b are obtained based on the generalized method of moments (GMM), using as instruments lagged values of $\log y_{t+1}^2$. Apart from the SVEB and SV models, Table 4 also reports values of the MSE, and the intercept and slope coefficients of the regression of $\log y_{t+1}^2$ on $h_{t+1|t}$ for the following volatility models: the EGARCH(1,1) and GJR-GARCH(1,1) estimated using the same set of data. These models share some common features with the SVEB model, as they consider the impact of stock return innovations on volatility function.

| | MSE | a | b | $\sigma_u^2$ | F$_{(p\text{-value})}$ |
|---|---|---|---|---|---|
| SVEB | 4.93 | -1.20 | 0.99 | 4.94 | 0.48 |
| | | (0.17) | (0.12) | | |
| SV | 4.96 | -1.10 | 0.96 | 4.96 | 0.11 |
| | | (0.16) | (0.11) | | |
| EGARCH(1,1) | 4.96 | -1.34 | 1.03 | 4.96 | 0.44 |
| | | (0.19) | (0.12) | | |
| GJR-GARCH(1,1) | 4.96 | -1.38 | 1.03 | 4.97 | 0.44 |
| | | (0.19) | (0.13) | | |
| Instruments: $\log y_t^2$, $\log y_{t-1}^2$, $\log y_{t-2}^2$ and $\log y_{t-3}^2$ | | | | | |

**Table 4:** Volatility forecasts

Notes: The table reports MSE values of $h_{t+1|t}$ to forecast $\log y_{t+1}^2$implied by observation equation (7), for different volatility models. It also provides GMM estimates of a linear regression model of (7) and carries out a joint test (F-type) of joint hypothesis a$= -1.27$ and b$= 1.0$. The $p$-value of this test is reported in the table.

The results of Table 4 clearly indicate that the SVEB model improves considerably the forecasting performance of the standard SV model, which does not allow for breaks. In terms of the values of MSE or the estimates of variance of the error term of forecasting regression (13), $\sigma_u^2$, reported in the table, this model is also found to outperform the three other parametric volatility models considered. Further support to the SVEB model compared to the other volatility models can be taken from the estimates of the forecasting regression coefficients a and b reported in the table. The probability values ($p$-values) of the joint hypothesis test: $a = -1.27$ and $b = 1.0$ reported in the table indicate that we can not reject this hypothesis. The probability of rejecting this hypothesis and, thus, making type I error is higher for the SVEB model, than the other volatility models.

## 5.2 Estimates of the news impact function implied by the SVEB model

In this section, we estimate the news impact function (NIF) implied by the SVEB model, based on the sample estimates of this model reported in Table 3. This function captures the reaction of the expected value of volatility one-period ahead, defined as $\Delta\sigma_{t+1|t} = \ln\left(\sigma_{t+1|t}/\sigma_{t|t-1}\right)$ where $\sigma_{t+1|t} = \exp\left(\frac{h_{t+1|t}}{2}\right)$, to stock return innovation forecasts, defined as $\varepsilon_{t|t-1} = \frac{y_t}{\sigma_{t|t-1}}$. Is is often used in the literature as a model specification criterion to assess if volatility models can explain the asymmetric responses of volatility changes $\Delta\sigma_{t+1|t}$ observed in practice, based, for instance, on implied by option market prices (e.g.,VIX) measures of volatility, with respect to positive and negative values of $\varepsilon_{t|t-1}$. As is shown in the literature (see, e.g., Engle and Ng(1993), Kane et al (2005) and Li et al (2005)), large negative values of $\varepsilon_{t|t-1}$ have a greater impact on $\Delta\sigma_{t+1|t}$ than their corresponding positive values. In particular, the effects of $\varepsilon_{t|t-1}$ on $\Delta\sigma_{t+1|t}$ is described by a monotonic and negative relationship which is more asymmetric at its end points, corresponding to large negative or positive return shocks.[10] This relationship is analogous to that between expected volatility changes $\Delta\sigma_{t+1|t}$ and values of stock return $r_t$ (see Ederington and Guan (2010)).

Table 5 and Figure 3 present estimates of NIFs implied by the SVEB and SV models, for different sub-intervals of return innovations $\varepsilon_{t|t-1}$. In Tables 5 and Figure 4, we present estimates of the above NIFs with respect to different sub-intervals of return values $r_t$. This table and figure also include estimates of NIFs based on market measures of volatility given by VIX, defined $\Delta VIX_{t+1} = \ln\left(VIX_{t+1}/VIX_t\right)$.[11] These will be compared to the NIFs implied by the SVEB and SV models to see which

[10] Note that parametric volatility models like GARCH, GJR-EGARCH, EGARCH can not produce INFs with the monotonic mentioned above (see, e.g., Yu (2005), Ederington and Guan (2010)). In particular, GARCH, EGARCH and GJR-GARCH models imply a V-shaped INF, which is against the almost linear INFs reported in the empirical literature, based on realised or implied values of volatility.

[11] The term $VIX_{t+1}$ designates implied volatility for the month beginning on day $t+1$ and running through day $t+21$ calculated from option prices on day $t$ (see Ederington and Guan (2010)).

of these two volatility models can better explain observed shapes of NIFs. Note that, since VIX is a four-weeks measure of volatility, the expected volatility changes presented in Table 6 and Figure 4 are estimated as $\Delta\overline{\sigma}_{t+4|t} = \ln\left(\overline{\sigma}_{t+4|t}/\overline{\sigma}_{t+3|t-1}\right)$, where $\overline{\sigma}_{t+4|t} = \frac{1}{4}\sum_{i=1}^{4}\sigma_{t+i|t}$.

The results of Tables 4-5 and Figures 3-4 clearly indicate that the SVEB model can produce shapes of NIFs which correspond to those observed in reality, based on the VIX values. These are negative and monotonic, and they become asymmetric towards their end points. Compared to the SV, the SVEB model can produce NIFs with higher in magnitude values and asymmetry, especially at the sub-intervals of large negative or positive return innovations. Note that, due to the allowance for different threshold parameter values $r^L$ and $r^R$, the NIFs implied by the SVEB model are more asymmetric with respect to large negative return innovations, rather than the positive. These functions imply that a large negative shock $\varepsilon_{t|t-1}$ will lead to a higher in magnitude change of future volatility $\Delta\sigma_{t+1|t}$ than that of large positive shock. As Figure 4 clearly indicates, the shape of these NIFs is consistent with that implied by the changes of VIX, which is an observable risk neutral measure of market volatility.

Another interesting conclusion which can be drawn from the inspection of Figure 4 is that the impact function of returns $r_t$ on expected volatility changes implied by the SVEB (or SV) model is lower than that implied by the VIX index, especially for large negative values of returns, i.e. for $r_t < -3.5$. This can be obviously attributed to the fact that the implied by option prices estimates of the stock market volatility values are adjusted for risk premia effects. Note that the differences between the VIX based estimates of volatility changes and those predicted by the SVEB model become larger, the larger the negative values of returns $r_t$ are. This result can be taken to support the view that stock market risk premia effects or their associated price of risk

may increase in terms of magnitude with the size of negative return innovations.

| Return shocks | SVEB with $(r^L, r^R)=(-2.05, 2.33)$ | | | SV | | |
|---|---|---|---|---|---|---|
| | obs. | $\Delta\sigma_{t+1\|t}$ | mean of $\varepsilon_{t\|t-1}$ | obs | $\Delta\sigma_{t+1\|t}$ | mean of $\varepsilon_{t\|t-1}$ |
| $\varepsilon_{t\|t-1} \leq -2.05$ | 27 | 0.2257 | -6.9449 | 25 | 0.1815 | -7.1389 |
| $-2.05 < \varepsilon_{t\|t-1} \leq -1.5$ | 74 | 0.1493 | -3.7282 | 78 | 0.1103 | -3.7376 |
| $-1.5 < \varepsilon_{t\|t-1} \leq -1$ | 143 | 0.1024 | -2.4008 | 160 | 0.0686 | -2.3572 |
| $-1 < \varepsilon_{t\|t-1} \leq -0.5$ | 233 | 0.0650 | -1.3845 | 215 | 0.0493 | -1.2998 |
| $-0.5 < \varepsilon_{t\|t-1} \leq 0$ | 262 | 0.0162 | -0.3338 | 261 | 0.0104 | -0.3454 |
| $0 < \varepsilon_{t\|t-1} \leq 0.5$ | 339 | -0.0291 | 0.6881 | 319 | -0.0221 | 0.6624 |
| $0.5 < \varepsilon_{t\|t-1} \leq 1$ | 246 | -0.0611 | 1.6265 | 266 | -0.0429 | 1.5425 |
| $1 < \varepsilon_{t\|t-1} \leq 1.5$ | 181 | -0.0931 | 2.7306 | 177 | -0.0722 | 2.8019 |
| $1.5 < \varepsilon_{t\|t-1} \leq 2.33$ | 64 | -0.1316 | 4.4018 | 68 | -0.0823 | 4.2908 |
| $2.33 < \varepsilon_{t\|t-1}$ | 7 | -0.1709 | 6.6817 | 7 | -0.1227 | 6.6817 |

**Table 5:** NIFs of the SVEB and SV models

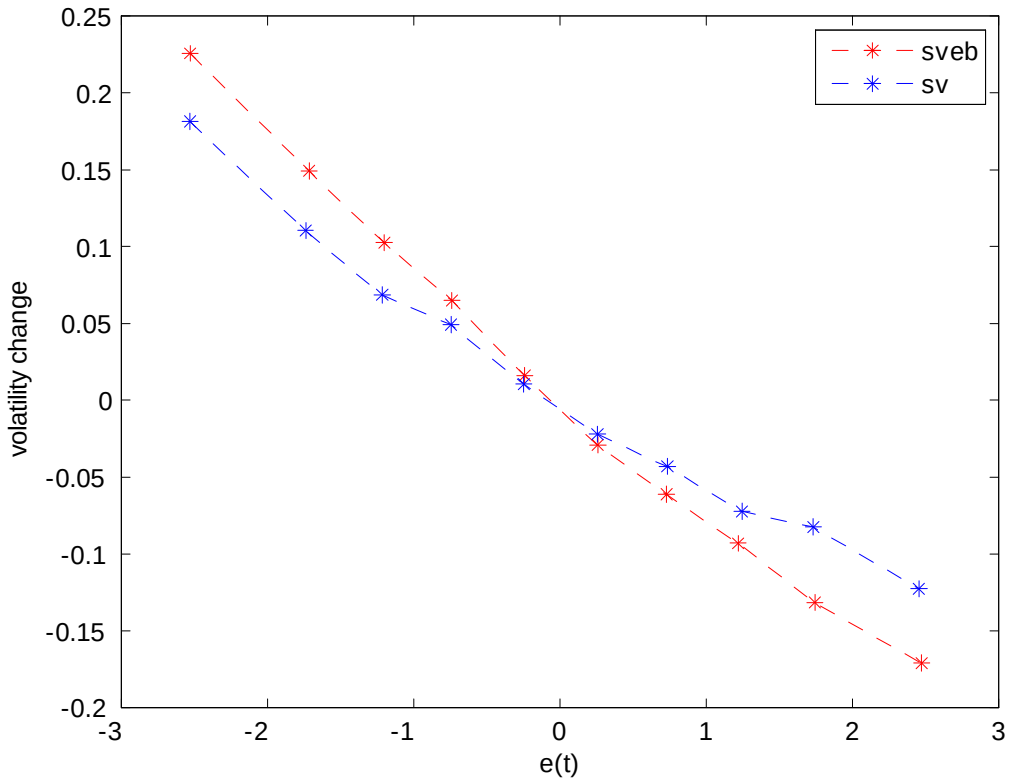


Figure 3: NIFs of the SVEB and SV models.

| Values of $r_t$ | SVEB with $(r^L, r^R)=(-2.05, 2.33)$ | | | SV | VIX | | |
|---|---|---|---|---|---|---|---|
| | obs. | mean of $r_t$ | $\Delta\sigma_{t+4}$ | $\Delta\sigma_{t+4}$ | obs. | mean of $r_t$ | $\Delta\sigma_{t+4}$ |
| $r_t < -3.5$ | 76 | -5.48 | 0.133 | 0.111 | 55 | -5.52 | 0.183 |
| $-3.5 < r_t \leq -2.50$ | 73 | -2.93 | 0.094 | 0.076 | 46 | -2.87 | 0.131 |
| $-2.50 < r_t \leq -2.0$ | 66 | -2.25 | 0.079 | 0.063 | 45 | -2.24 | 0.099 |
| $-2.0 < r_t \leq -1.5$ | 101 | -1.75 | 0.064 | 0.052 | 65 | -1.76 | 0.072 |
| $-1.50 < r_t \leq -1$ | 109 | -1.25 | 0.059 | 0.048 | 76 | -1.26 | 0.051 |
| $-1 < r_t \leq -0.5$ | 123 | -0.75 | 0.035 | 0.028 | 80 | -0.76 | 0.029 |
| $-0.5 < r_t \leq 0$ | 141 | -0.26 | 0.021 | 0.019 | 106 | -0.24 | 0.010 |
| $0 < r_t \leq 0.5$ | 155 | 0.24 | -0.004 | -0.003 | 108 | 0.23 | -0.012 |
| $0.5 < r_t \leq 1$ | 180 | 0.73 | -0.027 | -0.024 | 124 | 0.73 | -0.030 |
| $1 < r_t \leq 1.5$ | 176 | 1.23 | -0.038 | -0.030 | 112 | 1.23 | -0.055 |
| $1.5 < r_t \leq 2.0$ | 110 | 1.73 | -0.059 | -0.050 | 73 | 1.71 | -0.065 |
| $2.0 < r_t \leq 2.5$ | 86 | 2.23 | -0.069 | -0.060 | 49 | 2.22 | -0.063 |
| $2.5 < r_t \leq 3.5$ | 92 | 2.94 | -0.091 | -0.075 | 59 | 2.97 | -0.105 |
| $r_t > 3.5$ | 84 | 4.95 | -0.11 | 0.086 | 58 | 5.01 | -0.128 |

**Table 6:** NIFs of the SVEB and SV models, and VIX

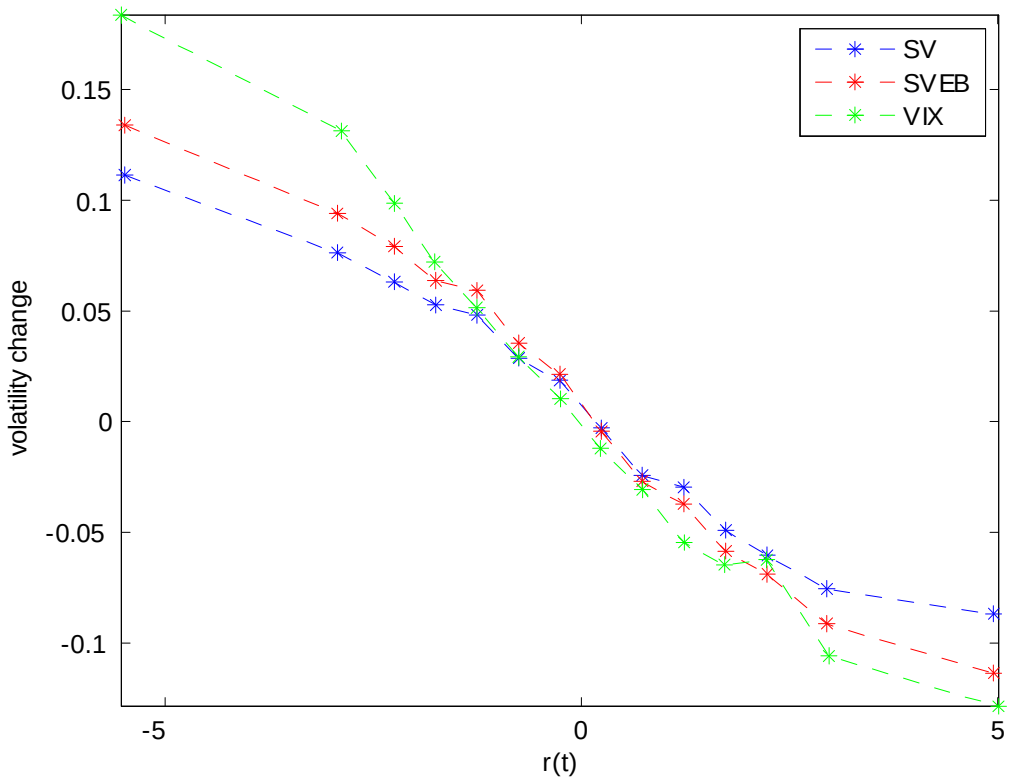


Figure 4: NIFs of the SVEB and SV models, and VIX with respect to stock return $r_t$.

## 5.3 Generalized impulse response functions of large stock return shocks

To study the pattern of the dynamic effects of large return shocks $\varepsilon_t$ on the expected future values of volatility $h_{t+\tau}$ net of the effects of possible future or past return innovations on volatility, in this section we will employ the SVEB model to calculate impulse response functions of future sequences of volatility $h_{t+\tau|t}$, for horizons $\tau = 1, 2, 3, ....$, with respect to innovations $\varepsilon_t$. Our analysis will enables us to examine the relative importance of large and ordinary return shocks on volatility, either for negative or positive shocks. Since the SVEB model is nonlinear and has a multivariate structure, we will calculate the Generalized IRF ($GIRF$) (see, e.g., Koop et al (1996), Pesaran and Shin (1998)), instead of the traditional $IRF$. The $GIRF$ considers impulse responses that are history and shock independent, while it treats the problem of future realizations of traditional $IRFs$. It is defined as the difference of the following conditional expectations $E\left[h_{t+\tau}|\,\varepsilon_t, \mathbf{a}_t\right]$ and $E\left[h_{t+t}|\,\mathbf{a}_t\right]$, i.e.

$$GIRF\left(\tau, \varepsilon_t, \mathbf{a}_t\right) = E\left[h_{t+\tau}|\,\varepsilon_t, \mathbf{a}_t\right] - E\left[h_{t+t}|\,\mathbf{a}_t\right], \text{ for } \tau = 1, 2, ..., \tag{14}$$

and it can provide impulse responses of $h_{t+\tau|t}$ to $\varepsilon_t$ given only the past of state vector $\mathbf{a}_t$. If we consider $\varepsilon_t$ and $\mathbf{a}_t$ as particular realizations of random variables, $GIRF\left(\tau, \varepsilon_t, \mathbf{a}_t\right)$ can be thought itself as a realization of a random variable whose distribution can be estimated.

The random nature of $GIRF\left(\tau, \varepsilon_t, \mathbf{a}_t\right)$ provides a more flexible approach to analyzing the effects of $\varepsilon_t$ on $h_{t+\tau}$. Someone can condition on a particular shock $\varepsilon_t$ and treat $\mathbf{a}_t$ as a random vector, or she/he can condition on a particular history $\mathbf{a}_t$, treating $\varepsilon_t$ as a random variable. Another possibility is to condition on particular subsets of the history of $\varepsilon_t$, i.e. large or ordinary positive, or negative, values of innovations $\varepsilon_t$. To this end, we will first define the following sets: $B^{(1)} = (-\infty, -2.05)$ and

$B^{(2)} = (2.33, \infty)$ consisting of the large negative and positive return shocks, respectively, identified by the estimates of the threshold parameters of the SVEB model, obtained in Section 5. Given $B^{(1)}$ and $B^{(2)}$, the following complementary sets of return shocks: $\tilde{B}^{(1)} = (-2.05, 0)$ and $\tilde{B}^{(2)} = (0, 2.33)$ define the ordinary negative and positive return shocks, respectively. Then, for the above all sets, we will calculate $GIRF(\tau, \varepsilon_t, \mathbf{a}_t)$, at time $t$. Below, we describe how this can be done for the case of large return shock sets, i.e. for $\varepsilon_t \in B^{(i)}$, $i = 1, 2$. An analogous procedure can be followed for the ordinary shocks $\varepsilon_t$, i.e. for $\varepsilon_t \in \tilde{B}^{(i)}$, $i = 1, 2$.

We will estimate the distribution of $GIRF(\tau, \varepsilon_t, \mathbf{a}_t)$ by means of a Monte Carlo integration. First, we will pick up 500 series of $\mathbf{a}_t$ from its sample estimates. To compute the first expectation of (14), i.e. $E[h_{t+\tau} | \varepsilon_t, \mathbf{a}_t]$, we proceed as follows: When $\tau = 1$, for each of the 500 series of $\mathbf{a}_t$, we will draw 30 realizations of the large return shock $\varepsilon_t$ from the truncated normal distribution $TN_{B^{(i)}}(0, 1)$, which is truncated at $B^{(i)}$, for $i = 1, 2$. That is, we will generate 15000 ($= 500 \times 30$) realizations of $GIRF$. Then, for each of the 30 realizations of $\varepsilon_t$, we will simulate 1000 volatility shocks $\eta_t$ conditional on each choice of $\varepsilon_t$, i.e. $\eta_t \sim N\left(\rho\sigma_\eta\varepsilon_t, \sigma_\eta^2(1-\rho^2)\right)$, and will average out to compute expectation $E[h_{t+\tau} | \varepsilon_t, \mathbf{a}_t]$. For $\tau > 1$, to estimate $E[h_{t+\tau} | \varepsilon_t, \mathbf{a}_t]$ we will simulate 1000 vectors of innovations $\varepsilon_{t+\tau-1}$ and $\eta_{t+\tau-1}$, according to its distributional assumption given in (4) and we will calculate the values of future volatility $h_{t+\tau}$ recursively, based on its data generating process. The average of the 1000 series of $h_{t+\tau}$ generated will give the estimate of $E[h_{t+\tau} | \varepsilon_t, \mathbf{a}_t]$, for $\tau > 1$. To compute the second expectation of (14), $E[h_{t+\tau} | \mathbf{a}_t]$, for all $\tau$, we will draw 1000 samples of the vector of innovations $\varepsilon_{t+\tau-1}$ and $\eta_{t+\tau-1}$, and then we will calculate $h_{t+\tau}$ recursively, according to its data degenerating process. The average of the series of $h_{t+\tau}$ generated will give an estimate of $E[h_{t+\tau} | \mathbf{a}_t]$. Given the above estimates of $E[h_{t+\tau} | \varepsilon_t, \mathbf{a}_t]$ and $E[h_{t+\tau} | \mathbf{a}_t]$, we can then obtain estimates of $GIRF(\tau, \varepsilon_t, \mathbf{a}_t)$, for all $\tau$, through

equation (14). To obtain the density of these estimates of $GIRF(\tau, \varepsilon_t, \mathbf{a}_t)$, we will use a normal kernel.

Figures 5 and 6 present the marginal densities of $GIRF(\tau, \varepsilon_t, \mathbf{a}_t)$ implied by the SVEB model, for horizons $\tau = 1, 2, ..., 20$, based on the procedure described above. This is done for both negative and positive large return shocks $\varepsilon_t$, i.e. $\varepsilon_t \in B^{(1)}$ and $\varepsilon_t \in B^{(2)}$, as well as for their complementary sets $\tilde{B}^{(i)}$, for $i = 1, 2$, consisting of ordinary return shocks. To better understand the economic meaning of distribution features of the above densities related to the traditional $IRFs$, in Table 7 we present the following descriptive statistics of them: the mean, variance, and skewness and kurtosis coefficients, for different values of horizon $\tau$. Inspection of the results of Figures 5-6 and Table 7 lead to a number of very useful conclusions about the dynamic effects of large or ordinary stock return shocks on the future paths of volatility.

First, both the plots of the $GIRF$ densities and the values of their descriptive statistics reported in the table clearly indicate that the effects of large return shocks on future paths of volatility tend to be much bigger than those of the ordinary shocks and to last for a higher number of future horizons $\tau$, e.g. $\tau = \{10\}$, ahead. This is much more likely to happen and it is true either for negative or positive return shocks. In terms of the distribution statistics reported in Table 7, it can be justified by the higher in magnitude values of the mean and skewness coefficients of the $GIRF$ densities for the large shocks, compared to those of the ordinary shocks for up to $\tau = 10$ horizons ahead. Note that the variance values of these densities between large and ordinary shocks are of the same order of magnitude, for all $\tau$.[12]

Second, either large or ordinary return shocks have the correct sign effect on volatility, predicted by the leverage hypothesis. That is, the negative shocks (reflect-

[12] As it can be seen from the descriptive statistics of Table 7, the distributional features of $GIRF$ for both the large and ordinary tend to those of the normal distribution, which has skeweness and kurtosis coefficients 0 and 3, respectively. But, this happens faster for the ordinary shocks. The mean of this distribution tends also faster to zero for the ordinary shocks. The latter can be obviously attributed to the specification of break process assumed by the SVEB model.

ing bad market news) have a positive effect on volatility, while the positive shocks (reflecting positive market news) have a negative effect. Note that these effects are much more clear and stronger than those implied by the estimates of the sample correlation coefficients $\text{Corr}(h_{t+\tau|n},\varepsilon_{t|t})$, for the different size of shocks, reported in Table 3, which undermine the effects of positive return shocks on volatility. The magnitude effects of these shocks on volatility differ slightly between the negative and positive shocks, with the positive shocks to have bigger mean effects than the negative. This can obviously attributed to the asymmetry of the estimates of the threshold parameters $r^L$ and $r^H$ found by our data.

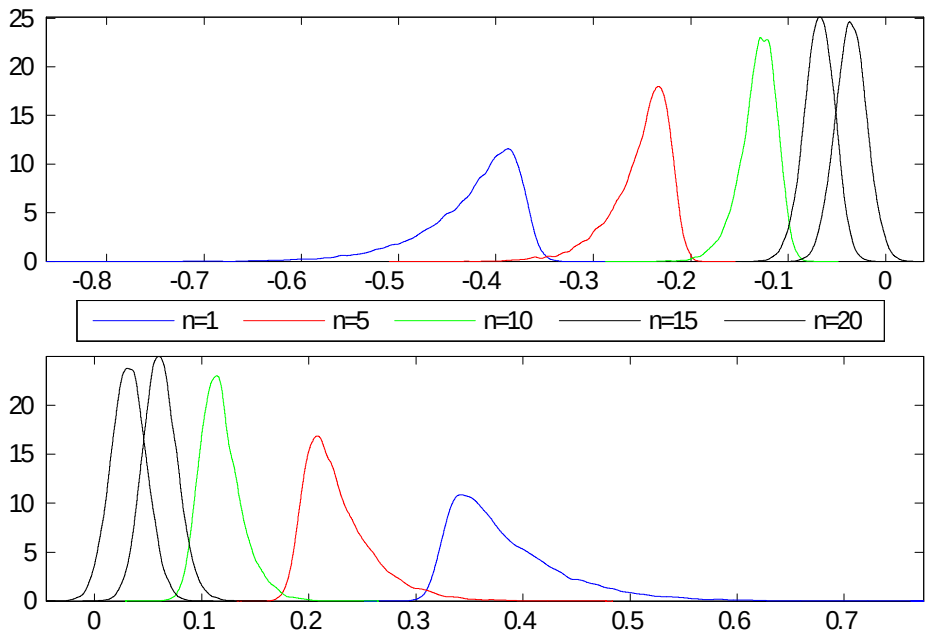


Figure 5: Estimates of the densities of GIRF of the SVEB model across $\tau$, for large positive and negative return shocks

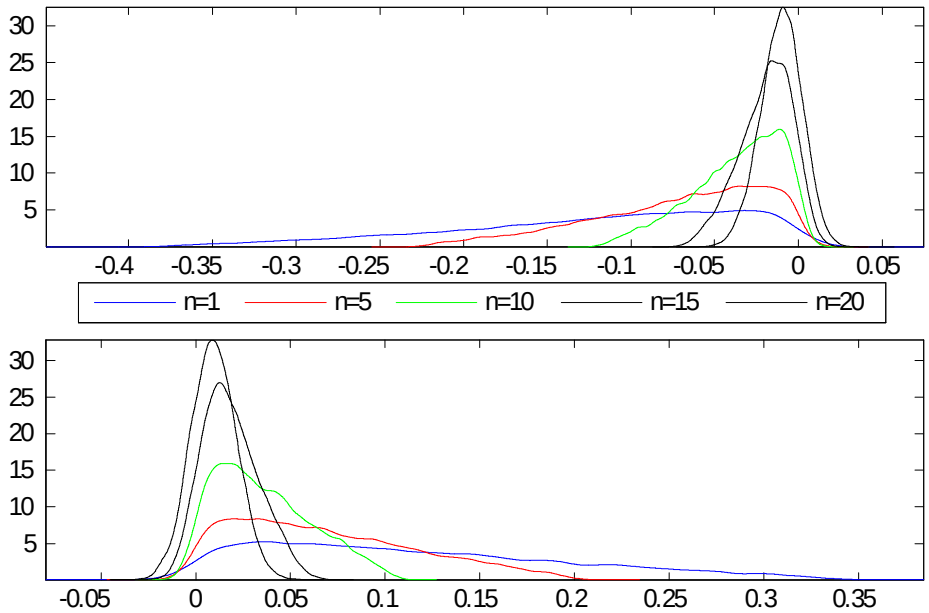


Figure 6: Estimates of the densities of GIRF of the SVEB model across $\tau$, for ordinary positive and negative return shocks

Summing up, the above results of the $GIRF$ density estimates can be taken to support the view that financial leverage effects and their persistence can be mainly attributed to large returns shocks. These results have important implications on forecasting volatility of stock returns.

| | $\varepsilon_t \in (-\infty, -2.05)$ | | | | $\varepsilon_t \in -2.05, 0.0)$ | | | |
|---|---|---|---|---|---|---|---|---|
| Horizon | Mean | Var | Sk | Ku | Mean | Var | Sk | Ku |
| $\tau = 1$ | 0.3842 | 0.0029 | 1.4794 | 6.1243 | 0.1153 | 0.0067 | 0.5775 | 2.4801 |
| $\tau = 2$ | 0.3372 | 0.0022 | 1.4423 | 5.7246 | 0.1029 | 0.0052 | 0.5433 | 2.4219 |
| $\tau = 3$ | 0.2954 | 0.0017 | 1.3968 | 5.5584 | 0.0887 | 0.0039 | 0.5683 | 2.4710 |
| $\tau = 4$ | 0.2603 | 0.0014 | 1.3946 | 5.5507 | 0.0786 | 0.0030 | 0.5542 | 2.4602 |
| $\tau = 5$ | 0.2278 | 0.0011 | 1.3691 | 5.5613 | 0.0685 | 0.0023 | 0.5536 | 2.4641 |
| $\tau = 10$ | 0.1185 | $3.9\times10^{-4}$ | 0.8923 | 4.4514 | 0.0358 | $6.6\times10^{-4}$ | 0.5331 | 2.4863 |
| $\tau = 15$ | 0.0613 | $2.5\times10^{-4}$ | 0.2173 | 3.2368 | 0.0184 | $2.2\times10^{-4}$ | 0.3693 | 2.7530 |
| $\tau = 20$ | 0.0319 | $2.6\times10^{-4}$ | 0.0586 | 3.0158 | 0.0096 | $1.4\times10^{-4}$ | 0.0797 | 2.9805 |
| | $\varepsilon_t \in (2.33, \infty)$ | | | | $\varepsilon_t \in (0.0, 2.33)$ | | | |
| Horizon | Mean | Var | Sk | Ku | Mean | Var | Sk | Ku |
| $\tau = 1$ | -0.4244 | 0.0026 | -1.4358 | 5.7257 | -0.1214 | 0.0077 | -0.6706 | 2.7022 |
| $\tau = 2$ | -0.3721 | 0.0020 | -1.4572 | 5.9072 | -0.1065 | 0.0059 | -0.6656 | 2.7086 |
| $\tau = 3$ | -0.3267 | 0.0015 | -1.4751 | 6.0474 | -0.0940 | 0.0046 | -0.6616 | 2.6730 |
| $\tau = 4$ | -0.2859 | 0.0012 | -1.4622 | 5.9919 | -0.0805 | 0.0034 | -0.6896 | 2.7452 |
| $\tau = 5$ | -0.2513 | $9.3\times10^{-4}$ | -1.3392 | 5.4702 | -0.0718 | 0.0027 | -0.6729 | 2.7057 |
| $\tau = 10$ | -0.1304 | $3.5\times10^{-4}$ | -0.8354 | 4.4029 | -0.0365 | $7.4\times10^{-4}$ | -0.6564 | 2.7549 |
| $\tau = 15$ | -0.0678 | $2.4\times10^{-4}$ | -0.2078 | 3.1435 | -0.0192 | $2.5\times10^{-4}$ | -0.4583 | 2.8454 |
| $\tau = 20$ | -0.0351 | $2.5\times10^{-4}$ | -0.0109 | 3.0047 | -0.0101 | $1.5\times10^{-4}$ | -0.1472 | 3.0269 |

**Table 7**: Descriptive statistics of GIRFs of the SVEB and SV models, for different $\tau$

Notes: Mean, Var, Sk and Ku stand for the mean, variance, skewness and kurtosis of the densities $GIRF\left(\tau, \varepsilon_t, \mathbf{a}_t\right)$, for $\tau = 1, 2, ..., 20$.

# 6 Conclusions

This paper suggests a new stochastic volatility model which extents the standard stochastic volatility model to allow for persistent level shifts in volatility, referred to as structural breaks in the empirical finance literature. These shifts are endogenously driven by large asset (stock) return shocks. The latter are defined as being bigger than the values of threshold parameters which can distinguish large negative shocks from positive shocks. Thus, the model is appropriate for studying the dynamic effects of large positive, or negative, pieces of stock market news on long-term paths of volatility.

The suggested model allows for shifts in volatility which are stochastic both in

time and magnitude. The last property of the model can explain clusters of volatility of asset return series observed in reality, whose variability has different size over time. Apart from interpreting different sources of volatility shifts, the model can be also employed to reveal from the data the magnitude of stock market return innovations which can be considered as large shocks. Since the model is nonlinear, to estimate its parameters and state variables, namely volatility and break processes, the paper relies on a Bayesian MCMC method. A Monte Carlo exercise conducted by the paper shows that the above estimation method can efficiently retrieve from the data estimates of its parameters and state variables. To estimate the threshold parameters of the model, we rely on a grid search method. This chooses as sample estimates of them those which give the maximum value of the marginal likelihood of the model.

The paper employs the model to investigate if level shifts in the volatility of the US stock market aggregate return are endogenously driven by large negative or positive return shocks. Then, it examines the dynamic effects of these shocks on the future levels of this volatility. The empirical analysis of the paper leads to a number of interesting conclusions. First, it identifies as large negative shocks these which are less than the -2.05% on weekly basis and as large positive shocks those which are bigger than 2.33%, and it finds out that cyclical shifts of the US stock market volatility can be satisfactorily modelled through large return shocks. The slight asymmetry between the values of large negative and positive return shocks identified by the data through our model can explain shapes of stock market news impact curves of future levels of volatility changes which are more asymmetric with respect to negative large return shocks (or, simply, stock returns) than positive return shocks, as is observed in reality. Finally, based on the generalized impulse response functions calculated by the estimates of the SVEB model, the paper indicates that large positive stock market shocks are expected to substantially reduce future levels of volatility, as predicted

by the financial leverage effects hypothesis. The last relationship can not be easily diagnosed by sample data statistics due to the within-sample offsetting effects that positive and negative return shocks can have on volatility.

# References


[1] Andreou E. and E. Ghysels (2002), "Detecting multiple breaks in financial market volatility", *Journal of Applied Econometrics*, 17, 5, 579-600.

[2] Bai J., and Perron (2003), "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, 18, 1-22.

[3] Bekaert G. and G. Wu (2000), "Asymmetric volatility and risk in equity markets," *Review of Financial Studies*, 13(1), 1-42.

[4] Box G. E. P. and G. Tiao (1975), "Innovation analysis with applications to economic and environmental problems," *Journal of American Statistical Association*, 70, 70-79.

[5] Campbell J. and L. Hentschel (1992), "No news is good news: an asymmetric model of changing volatility in stock return", *Journal of Financial Economics*, 4, 353-384.

[6] Chan K. S. (1993), "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model," *The Annals of Statistics*, 21(1), 520-533.

[7] Chen C.W.S, Chiang T.C and M.K.P. So (2003), "Asymmetrical reaction to US stock-return news: evidence from major stock markets based on a double-threshold model", *Journal of Economics and Business,* 55, 487-502.

[8] Cristie A. A. (1992), "The stochastic behavior of common stock variances," *Journal of Financial Economics*, 10, 407-432.

[9] Chib S. (1995), "Marginal Likelihood from the Gibbs output," *Journal of the American Statistical Association*, 90, 1313-1321.

[10] Chib S. and I. Jeliazkov (2001), "Marginal likelihood from the Metropolis-Hastings output," *Journal of American Statistical Association*, 96, 270-281.

[11] Cogley T., and T.S. Sargent (2002), "Drifts and volatilities: Monetary policies and outcomes in the post WWII US" *New York University, Mimeo.*

[12] De Jong P. and N. Shephard (1995), "The simulation smoother for time series models", *Biometrica*, 82, 339-350.

[13] Denis P., Mayhew S. and C. Stivers (2006), "Stock returns, implied volatility innovations, and the asymmetric volatility phenomenon", *Journal of Financial and Quantitative Analysis* 41, 381-406.

[14] Diebold F.X. and A. Inoue (2001), "Long memory and structural change" *Journal of Econometrics*, 105, 131-159.

[15] Diebold F.X. and P. Pauly (1987), "Structural change and the combination of Forecasts", *Journal of Forecasting*, 6, 21-40.

[16] Durbin J. and S. J. Koopman (2001), Time series analysis by state space methods. Oxford University Press.

[17] Ederington L.H. and W. Guan (2010), "How asymmetric is U.S. stock market volatility?," *Journal of Financial Markets*, 13, 225-248.

[18] Engle R.F. and V.K. Ng (1991), "Measuring and testing the impact of news on volatility," *Journal of Finance,* 48, 1749-1778.

[19] Engle R.F. and A.D. Smith (1999), "Stochastic permanent breaks", *Review of Economics & Statistics*, 81, 553-574.

[20] Fleming J., Ostdiek B. and R. Whaley (1995), "Predicting stock market volatility: a new measure", *Journal of Futures Markets*, 15, 265-302.

[21] French K., Schwert G. W. and R. Stambaugh (1987), "Expected stock returns and volatility", *Journal of Financial Economics*, 19, 3-30.

[22] Glosten L.R., R. Jagannathan and D. Runkle (1993), "Relationship between the expected value and the volatility of the nominal excess return on stocks" *Journal of Finance*, 48, 1779-1802.

[23] Harvey A. C., Ruiz E. and N. Shephard (1994), "Multivariate stochastic variance models" *Review of Economic Studies*, 61, 247-264.

[24] Kane A., B.N. Lehmann and R.R. Trippi (2000), "Regularities in volatility and the price of risk following large stock market movements in US and Japan", *Journal of International Money and Finance*, 19, 1-32.

[25] Kapetanios G. (2000), "Small sample properties of the conditional least squares estimator in SETAR models" *Economic Letters*, 69(3), 267-276.

[26] Kapetanios, G. and E. Tzavalis (2010), "Modeling structural breaks in economic relationships using large shocks", *Journal of Economic Dynamics and Control,* 34, 417-436.

[27] Karolyi A. and R. Stulz (1996), "Why do markets move together? An investigation of US-Japanese stock return comovements", *Journal of Finance*, 51, 951-986.

[28] Kim S., Shephard N. and S. Chib, (1998), "Stochastic volatility: Likelihood inference and comparison with ARCH Models," *Review of Economic Studies*, 65, 361-393

[29] Koopman S. J. (2005), "On importance sampling for the state space model" *Timbergen Institute Discussion chapter No.* 2005-117.

[30] Koop G., Pesaran M.H. and S.M. Potter (1996), "Impulse response analysis in nonlinear multivariate models," *Journal of Econometrics,* 74, 119-147.

[31] Kramer W and B. Tameze (2007), "Structural change and estimated persistence in the GARCH-models", *Economics Letters*, 114, 72-75.

[32] Lamourex C. C. and W. Lastrapes (1990), "Persistence in variance, structural change and the GARCH models" *Journal of Business Economic Statistics*, 8, 225-234.

[33] Li Q., Yang J., Hsiao C. and Y. Chang (2005), "The relationship between stock returns and volatility in international stock markets," *Journal of Empirical Finance*, 12, 650-665.

[34] Ling S. and M. McAleer (1996), "Stationarity and the existence of moments of a family of GARCH processes," *Journal of Econometrics*, 106, 109-117.

[35] Low C. (2004), "The fear and exuberance from implied volatility of S&P500 index", *Journal of Business*, 77, 352-546.

[36] Malik F. (2011), "Estimating the impact of good news on stochastic market volatility", *Applied Financial Economics*, 8, 1-10.

[37] Mele A. (2007), "Asymmetric stock market volatility and the cyclical behaviour of expected returns", *Journal of Financial Economics*, 86, 446-478.

[38] Mikosch T. and Starica (2004), "Nonstationarities in Financial Time Series, the long range dependency and IGARCH Effects", *Review of Economics and Statistics*, 86, 378-390.

[39] Morales-Zumaquero A. and S. Sosvilla-Rivero (2010), "Structural breaks in volatility: Evidence for the OECD and non-OECD real exchange rates", *Journal of International Money and Finance*, 29, 139-168.

[40] Morana C. and A. Beltratti (2004), "Structural change and long-range dependence in volatility of exchange rates: Either, neither or both?," *Journal of Empirical Finance* 11, 629-658.

[41] Omori Y., Chib S., Shephard, N. and J. Nakajima (2006), "Stochastic volatility with leverage: Fast and efficient likelihood inference" *Journal of Econometrics*, 140, 425-449.

[42] Ozdagli, A.K. (2012), "Financial leverage, corporate investment and stock returns", *Review of Financial Studies*, 25, 1033-1069.

[43] Pagan A. and G. W. Schwert (1990), "Alternative models for conditional volatility", *Journal of Econometrics*, 45, 267-290.

[44] Pesaran H. and Y. Shin (1998), "Generalized impulse response analysis in linear multivariate models," *Economics Letters*, 58(1), 17-29.

[45] Pitt M.K. and N. Shephard (1999), "Filtering via simulation: Auxiliary particle filters," *Journal of American Statistical Association*, 94, 590-599.

[46] Psaradakis, Z. and E. Tzavalis (1999), "On regression-based tests for persistence in volatility models", *Econometric Reviews*, 18, 441-448.

[47] Rapach, D. and J.K. Strauss (2008), "Structural breaks and GARCH models of exchange rate volatility", *Journal of Applied Econometrics*, 23, 65-90.

[48] Schwert G. W. (1990), "Stock market volatility and the crass of 87", *Review of Financial Studies*, 3, 77-103.

[49] Taylor S. J. (1996), *Modeling financial Time Series, Wiley.*

[50] Tweedie R.L. (1975), "Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space," *Stochastic Processes Appl*, 3, 385-403.

[51] Tzavalis E. and M.R. Wickens (1995), "The Persistence in volatility of the US term premium," *Economic Letters*, 49, 381-398.

[52] Yu J. (2005), "On Leverage in a Stochastic Volatility Model," *Journal of Econometrics* 127, 165-178.

# Appendix

In this appendix, we provide the proof of Theorem 1, presented in the main text.

## Proof of Theorem 1

We now prove strict stationarity for $h_t$, given by

$$y_t = \exp\left(-\frac{h_t}{2}\right)\varepsilon_t$$

$$h_{t+1} = b_{t+1} + \phi h_t + \eta_t \quad \text{and} \quad b_{t+1} = \delta_t b_t + I(\mathcal{A}_t)\gamma_t,$$

where

$$\delta_t = \begin{cases} 1 \text{ if } I(|b_t| < b) \\ 0 \quad \text{otherwise.} \end{cases} \tag{15}$$

The first step is to derive a recursive representation for $h_t$. This is given by

$$h_t = \sum_{j=0}^{\infty} \phi^j \left(b_{t-j} + \eta_{t-j-1}\right).$$

Following Theorem 2.1 of Ling and McAleer (1996), the result will follow if we show that for some $\alpha \in (0,1)$

$$E\left(h_t^{\alpha}\right) < \infty.$$

By The Marcinkiewicz-Zygmund inequality we have that

$$E(h_t^{\alpha}) = E\left(\left(\sum_{j=0}^{\infty} \phi^j \left(b_{t-j} + \eta_{t-j-1}\right)\right)^{\alpha}\right) \leq c\left(\sum_{j=0}^{\infty} \phi^{2j}\right)^{\alpha/2} E\left(b_{t-j} + \eta_{t-j-1}\right)^{\alpha},$$

which is finite as long as $b_t$ is strictly stationary and $E\left(b_{t-j}\right)^{\alpha} < \infty$ and $E\left(\eta_{t-j-1}\right)^{\alpha} < \infty$. Thus it suffices to prove that $b_t$ is strictly stationary and $E\left(b_{t-j}\right)^{\alpha} < \infty$. $E\left(b_{t-j}\right)^{\alpha} < \infty$ follows easily from strict stationarity and $E\left(\gamma_t\right)^{\alpha} < \infty$. Thus we only need to prove strict stationarity for $b_t$. To do that we prove geometric ergodicity of $b_t$, which implies strict stationarity asymptotically. To prove geometric ergodicity, we use the drift criterion of Tweedie (1975). This condition states that a process is ergodic under the regularity condition that disturbances have positive densities

everywhere if the process tends towards the center of its state space at each point in time. More specifically, $b_t$ is geometrically ergodic if there exists constants $0 < \vartheta < 1$, $B, L < \infty$, and a small set $C$ such that

$$E\left[\|b_t\| \mid b_{t-1} = d\right] \leq \vartheta \|d\| + L, \quad \forall d \notin C, \tag{16}$$

$$E\left[\|b_t\| \mid b_{t-1} = d\right] \leq B, \quad \forall d \in C, \tag{17}$$

where $\|\cdot\|$ is the Euclidean norm. The concept of the small set is the equivalent of a discrete Markov chain state in a continuous context. It is clear that (17) follows easily. We need to show (16). (16) follows if the following condition holds

$$E(\delta_t) < 1. \tag{18}$$

To prove (18) it suffices to show that

$$\Pr(|b_t| > b) > 0.$$

This follows easily by the independence of $\varepsilon_{t-1}$ and $\gamma_t$, the fact that $\Pr(A_t) > 0$ and the fact that $\Pr(|\gamma_t| > 2b) > 0$ for all finite $b$.