Shifts in Volatility Driven by Large Stock Market Shocks

Yiannis Dendramis†  George Kapetanios*  Elias Tzavalis†

January, 2012

Abstract

This paper presents a new stochastic volatility model which allows for shifts in volatility of stock returns, referred to as breaks, driven endogeneously by large market pieces of news, referred to as large skocks. The values of these shocks are identified from the data as being bigger than two threshold parameters of the model: one for the negative shocks and another for the positive shocks. The model can be employed to investigate different economic sources of volatility shifts, without relying on exogenous information from the sample. In addition to this, it has a number of interesting properties. It can allow for level type of shifts in volatility which are stochastic both in timing and magnitude. Second, it can explain asymmetries of the impact of return innovations, or their dynamic effects, on stochastic volatility observed by implied (or realised) measures of volatility, beyond those predicted by leverage effects assumed by standard stochastic volatility models. The above properties of the model are shown based on estimates of it for the volatility of the US stock market. 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@aeub.gr.
1 Introduction

There is recently considerable evidence indicating the existence of discrete-time and persistent shifts in the conditional variance process volatility of asset (stock) returns. These shifts, referred to as structural breaks, are related to large negative, or positive, stock market return shocks, which reflect substantial pieces of market news (see Diebold and Pauly (1987), Lamoureux and Lastrapes (1990), Tzavalis and Wickens (1995), Hamilton and Lin (1996), Diebold and Inoue (2001), Ang and Bekaert (2002), Andreou and Ghysels (2002), Mikosch and Starica (2004), Morana and Beltratti (2004), Smith (2009), inter alia). There are many economic reasons for which larger, in magnitude, return shocks can cause a permanent shift in the level of market volatility. These can reflect changes in agents’ beliefs about the optimal asset allocation due to monetary regime changes or realignments, financial market crises, institutional changes or other market events. If they are not accounted for, they tend to overstate evidence of persistence in the volatility process (see, e.g., Psaradakis and Tzavalis (1999)).

The empirical literature mentioned above tends to treat the large stock return shocks as exogenous. To capture their effects on volatility relies on the intervention (dummy) variable analysis of Box and Tiao (1975), relying on exogenous information from the sample to determine the time points that the breaks occur. Of course, more sophisticated multibreak testing procedures can be applied to find out from the data the timing that these breaks occur, like those employed for breaks in the mean of 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 of volatility shifts and to study the dynamic effects of their associated shocks on
volatility functions. Separating the impact of these shocks on volatility from those of the ordinary volatility shocks may also have important implications for long-term portfolio management and hedging, as it should bring more focus on controlling large risks based on long-term shifts in volatility leaving aside its short-term ones.

To overcome the above pitfalls of the intervention analysis, in this paper we suggest a parametric model of breaks in volatility function which are driven by large in magnitude stock return shocks. Depending on their sign, these shocks are identified by being larger (or smaller) than a positive (or negative) value of threshold parameter which can be estimated from the sample based on a search procedure. The different threshold parameters that our model considers can incorporate asymmetric threshold effects in volatility function, which can capture extra asymmetries of this function to those implied by leverage effects (see, e.g., Yu (2005) and Smith (2009)). Our model 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 distinguishes our model from other parametric volatility models allowing for endogenous break processes which assume shifts in the volatility function parameters of fixed magnitude, see, e.g., Hamilton’s (1989) Markov regime-switching model, Glosten’s et al (1993) threshold GARCH model, and the stochastic volatility (SV) threshold models of So et al (2003), and Asai and McAleer’s (2004) and Smith (2009). The last category of threshold models assume known values of threshold parameters. These have been introduced in the literature to capture possible asymmetries in volatility, which may explain leverage effects.

The econometric framework employed to build up our 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 search procedure which involves the following steps. First, conditioned on the threshold parameters we use a mixture of ten normal distributions to approximate the joint distribution of log squared stock return innovations and the volatility function innovations. We then sample the structural parameters of the model and its state variables from their posterior distributions, and calculate the marginal densities. The values of the threshold parameters which give the maximum value of the marginal likelihoods over the threshold parameters space are considered to be as the optimum ones. The paper proves that the above estimation procedure of the model leads to consistent parameter estimates, including those of threshold parameters. To evaluate the performance of the above procedure, the paper conducts a Monte Carlo exercise, which shows that it works very efficiently.

Our model is employed to explain level shifts in the volatility of the US stock market aggregate return. The results of this study indicate that, indeed, these shifts can be attributed to large market return shocks. The most dominant of these shifts, which are associated with persistent changes in the slope of volatility function, are triggered by sequences of large return shocks which are related to market news about financial crises. Our model identifies as large negative return shocks those 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. This can explain shapes of news impact functions based on implied, or realized, values of volatility which are asymmetric towards large negative return innovations (see, e.g., Ederington and Guan (2010)). Our empirical results indicate that the standard stochastic volatility model with leverage effects, which does not allow for level shifts, cannot sufficiently capture the above market news asymmetries.

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 impact news functions and generalized impulse response functions, with the aim of 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 $t$, $r_t$:

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

(1)

with stochastic volatility process

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

(2)

and a break process, given as

$$b_{t+1} = b_t + I(A_t)\gamma_t, \quad (3)$$

5
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 (news), respectively, distributed as

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

(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 by the process (3), are driven by large return innovations $\varepsilon_t$, referred to as large shocks. They are identified by indicator function $I(A_t)$ taking the value 1 if the event $A_t = \{\varepsilon_t > r^R \text{ or } \varepsilon_t > r^L\}$ occurs, and zero otherwise, where $(r^L, r^R)$ is a pair of threshold parameters which can be estimated based on sample information.\footnote{In the literature, these news are sometimes recognised as outliers in the level of series $y_t$ (see, e.g., Huang (2007)).} The events captured by set $A_t$ can be thought of as reflecting large pieces of positive (or negative) stock market news when $\varepsilon_t > r^R$ (when $\varepsilon_t < r^L$), affecting stock market returns at time $t$. Since the breaks captured by the above SV model, defined by equations (1)-(3), are endogenously driven by stock return innovations $\varepsilon_t$, this volatility model will be henceforth denoted with the acronym SVEB, which stands for 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 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 $I(A_t)$, 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 allows for a more flexible approach of modelling cyclical shifts in volatility relaxing the assump-
tion that these shifts are of constant magnitude over time, as is considered by existing volatility models allowing for breaks like the regime-switching volatility model of Hamilton (1989) and the threshold volatility models of Glosten et al (1993), So et al (2003), Asai and McAleer (2004) and Smith (2009). These threshold models have been introduced in the literature to capture possible asymmetries in volatility, which may explain leverage effects.\(^2\)

In addition to the above property, 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 \(r^R\), the SVEB model can capture asymmetries of stock market news on volatility function beyond those implied by leverage effects of stochastic volatility models. This will result in producing patterns of stock market news impact functions (NIFs) or impulse response functions (IRFs) with higher degree of asymmetries between negative and positive values of large stock return innovations, compared to those implied by the SV model, which does not allow for breaks. This may explain evidence in the literature, based on measures of realized or implied volatility, indicating that volatility responses more

\(^2\)For instance, the model suggested by Asai and McAleer (2004) assumes that, in terms of our notation, processes \(h_t\) and \(b_t\) are given as \(h_{t+1} = b_t + \phi h_t + \eta_t\) and \(b_t = \gamma (I(\epsilon_t < 0) - E[I(\epsilon_t < 0)])\), where \(I(.)\) is an indicator function taking the value of 1 if \(\epsilon_t < 0\), and zero otherwise. This function of \(b_t\) implies that stochastic volatility \(h_{t+1}\) switches between the following two values of \(b_t\):

\[
b_t = \begin{cases} 
-\gamma \{ E[I(\epsilon_t < 0)] \} & \text{if } \epsilon_t \geq 0 \\
\gamma \{ 1 - E[I(\epsilon_t < 0)] \} & \text{if } \epsilon_t < 0, 
\end{cases}
\]

which are of fixed magnitude. The \(\rho\) and \(\gamma\) parameters of the above model, capturing leverage and threshold effects, respectively, can not be identified, separately. That is, when \(E(\varepsilon_t b_t) = 0\) (implying \(\rho = 0\)), this model can yield the same magnitude of leverage effects to those implied by the SV model allowing for leverage effects (see equation (4) of Asai and McAleer (2004)). This problem is not present in our model since \(Cov(\varepsilon_t, I(A_t) \gamma_t) = E[\gamma_t]Cov(\varepsilon_t, I(A_t)) = 0\).
sharply to large negative stock return shocks compared to large positive shocks of the same magnitude (see, e.g., Bekaert and Wu (2000), Kane et al (2000), Yu (2004) and Ederington and Guan (2010)).

The SVEB model can nest different stochastic volatility models, which may be applied 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 of 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 (1989)). 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 \( |\phi| < 1 \).

\(^3\)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.
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 retrieve sample estimates of the parameters of the volatility function of the SVEB model, collected in vector $\theta = (\phi, \sigma_n, \sigma_\gamma, \rho)'$, and its state variables, collected in vector $a_t = (h_t, b_t, \gamma_t)'$. To this end, we will rely on the Bayesian MCMC estimation method suggested by Omori et al (2006). This is extended to provide estimates of threshold parameters $r^L$ and $r^R$.

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.

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

where $\varepsilon_t^*$ is a transformed IID innovation process, which follows a $\log \chi^2_1$ 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:

$$y_t^* = h_t + \varepsilon_t^* \quad (7)$$

and

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

respectively, where $w_{t+1} \sim NIID (0, \sigma^2_w)$.

Let us assume that the vector of threshold parameters $r = (r^L, r^R)'$ is known, at the moment. The key feature of the MCMC method suggested by Omori et al (2007) 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 \( s_t \in \{1, 2, \ldots, 10\} \), for \( t = 1, 2, \ldots, n \). 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 \( s_n = \{s_t\}_{t=1}^n \) 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 \( s_n \) and then iterating the following steps to obtain posterior samples:

1. Draw \( \mathbf{a}_n, \theta \mid \mathbf{y}_n^*, \mathbf{d}_n, s_n \) by
   
   (a) drawing \( \theta \mid \mathbf{y}_n^*, \mathbf{d}_n, s_n \)
   
   (b) drawing \( \mathbf{a}_n \mid \mathbf{y}_n^*, \mathbf{d}_n, s_n, \theta \)

2. Draw \( s_n \mid \mathbf{y}_n^*, \mathbf{d}_n, \mathbf{a}_n, \theta \),

until convergence is achieved. To derive the posterior distribution \( \theta \mid \mathbf{y}_n^*, \mathbf{d}_n, s_n \) based on the above algorithm, we will assume a prior distribution of \( \theta \), denoted as \( \pi (\theta) \), and will calculate the likelihood \( g (\mathbf{y}_n^* \mid \theta, \mathbf{d}_n, s_n) \) based on the approximation of the bivariate joint density \( f (\varepsilon_t^* \mid \theta, \mathbf{d}_n, s_n) \) by the following mixture of \( K = 10 \) normal distributions with mean \( m_j \), variance \( v_j^2 \), denoted as \( N (\varepsilon_t^* \mid m_j, v_j^2) \), for \( j = 1, 2, \ldots, K = 10 \):

\[
\begin{align*}
f (\varepsilon_t^* \mid \eta_t \mid d_t) &= f (\eta_t \mid \varepsilon_t^* \mid d_t) f (\varepsilon_t^*) \\
&\approx \sum_{j=1}^{10} p_j N \left( \eta_t; d_t \rho \sigma e^{m_j} (a_j + b_j (\varepsilon_t^* - m_j)), \sigma^2 \left( 1 - \rho^2 \right) \right) N (\varepsilon_t^* \mid m_j, v_j^2),
\end{align*}
\]

(9)

where \( \{p_j, a_j, b_j\}_{j=1}^{10} \) constitute the mixing parameters. These parameters are chosen to make the approximation of the true density \( f (\varepsilon_t^* \mid \eta_t \mid d_t) \) as tight as possible.\(^4\)

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

\[
\left\{ \begin{array}{l}
(\varepsilon_t^*) \\
\eta_t
\end{array} \right\} \mid d_t, s_t = j \xrightarrow{L} \left[ \begin{array}{c}
d_t \rho \sigma \left( a_j + b_j (\varepsilon_t^* - m_j) \right) \\
\frac{m_j + v_j \varepsilon_t}{2} \exp \left( \frac{m_j}{2} \right) + \sigma \sqrt{1 - \rho^2} z_t
\end{array} \right],
\]

(10)

\(^4\)Optimal values of \( a_j \) and \( b_j \), as well as of \( p_j \) and \( m_j, v_j \) for \( j = 1, \ldots, 10 \), are reported by Omori et al (2006).
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 \( \overset{L}{\Rightarrow} \) 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 \( g (y_n^* \mid \theta, d_n, s_n) \). In so doing, we will replace the innovations \( \varepsilon_t \) of set \( \mathcal{A}_t \) by 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 \( \mathbf{r} = (r^L, r^R)' \). The above linear Gaussian state space form of the SVEB model allows us to draw posterior samples from the density of \( a_n \mid y_n^*, d_n, s_n, \theta \) (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 \( \pi (\theta \mid y_n^*, d_n, s_n) \propto g (y_n^* \mid \theta, d_n, s_n) \pi (\theta) \) 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 \( g (y_n^* \mid \theta, d_n, s_n) \pi (\theta) \). Then, we will generate

\[
y_t^* = \mathbf{c}'_t \mathbf{a}_t + m_j + \mathbf{G}_t \mathbf{u}_t,
\]

\[
a_{t+1} = \mathbf{T}_t \mathbf{a}_t + \mathbf{W}_t + \mathbf{H}_t \mathbf{u}_t,
\]

where

\[
\mathbf{T}_t = \begin{bmatrix}
\phi & 1 & I (\mathcal{A}_t) \\
0 & 1 & I (\mathcal{A}_t) \\
0 & 0 & 0
\end{bmatrix},
\mathbf{H}_t = \begin{bmatrix}
d_t \rho \sigma_1 \exp \left( \frac{m_j}{2} \right) b_j v_j z_{1t} & \sigma_\gamma \sqrt{1 - \rho^2} & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix},
\mathbf{G}_t = [v_j, 0, 0, 0]' \\
\mathbf{W}_t = \begin{bmatrix}
d_t \rho \sigma_\eta \exp \left( \frac{m_j}{2} \right) a_j, 0, 0
\end{bmatrix}',
\mathbf{u}_t \sim NIID (0, \mathbf{I}) \quad \varepsilon_t = \exp \left( -\frac{\mathbf{c}'_t \mathbf{a}_t}{2} \right) y_t
\]
a candidate \( \theta^* \) from the truncated normal distribution 
\[
TN_R \left( \bar{\theta} ; \left( - \frac{\partial^2 \log \pi(y_n^*|\theta, d_n, s_n)\pi(\theta)}{\partial \theta \partial \theta} \right)_{\theta=\bar{\theta}} \right)^{-1},
\]
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:

\[
\pi(s_n|y_n^*, d_n, a_n, \theta) \propto Pr(s_t=j)N(y_t^*-h_t;m_j,\sigma_j^2)
\]

\[
N(h_{t+1} - \phi h_t - b_t - I(A_t) \gamma_t; d_t \rho \sigma e^{\sum_{j} \{ a_j + b_j (y_t^* - m_j) \} , \sigma_j^2 (1 - \rho^2)}) ,
\]

for all \( j \).

Since the MCMC method presented above involves a small approximation error due the sampling variability of \( s_n \), to control this, we will reweight the samples of \( \{h_t^k, b_t^k, \gamma_t^k, \theta_t^k\}_{k=1}^M \) obtained through the MCMC iterations, denoted as \( k = 1, 2, ..., M \), using the following weights:

\[
w_t^* = \prod_{t=1}^{n} \frac{f(\varepsilon_t^*, \eta_t^k|d_t, \theta_t^k)}{\tilde{f}(\varepsilon_t^*, \eta_t^k|d_t, \theta_t^k)}, \quad \text{for} \quad k = 1, 2, ..., M
\]

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

The estimation procedure presented above assumes that the vector of threshold parameters \( \mathbf{r} = (r^L, r^R)' \) is known, which is not true in reality. 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 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 \( a_t \) corresponding to this value of \( \mathbf{r} \) will

\(^6\)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.
constitute the maximum likelihood estimates of the model. These estimates will be consistent provided that vector \( r \) is also consistently estimated. The consistency of \( r \) based on grid search estimation method can be proved following analogous arguments to that of Kapetanios and Tzavalis (2010), considering a stochastic break process like (8) for the mean of economic series. Below, we present a useful practical remark for the estimation of threshold \( r \).

**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 \( r \) 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 \( m(y|\theta) \) (or its logarithmic value \( \ln m(y|\theta) \)) 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 \( g(y|\theta^*, r) \), and Chib’s and Jeliaskov (2001) method to calculate the posterior ordinate, \( \pi(\theta^*|y, r) \). The marginal likelihood \( m(y|\theta) \) can be calculated from \( g(y|\theta, r) \) using Bayes’ theorem, i.e.

\[
m(y|\theta) = g(y|\theta^*, r) \pi(\theta^*) / \pi(\theta^*|y, r),
\]

where \( \pi(\theta^*|y, r) \) and \( \pi(\theta^*) \) are the posterior and prior ordinates. Note that the above expression of marginal density \( m(y|\theta) \) 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 \( m(y|\theta) \), the auxiliary particle filter algorithm can be employed to provide forecasted and filtered values of the state vector

---

7This 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).
\( a_{t+1} \) based on information sets \( I_t = \{y_t, y_{t-1}, \ldots y_1\} \) and \( I_{t+1} \), respectively, denoted as \( a_{t+1|t} \) and \( a_{t+1|t+1} \). These can be employed to calculate goodness of fit or forecasting performance measures of the SVEB model, which are very 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. This is done for a sample where the number of breaks is relatively small. 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 \( a_t = (h_t, b_t, \gamma_t)' \) and the vector of parameters \( \theta \), while we treat the threshold vector \( r \) as known.

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: \( \phi = 0.85, \sigma_y = 0.28, \sigma_\gamma = 0.05, \rho = -0.59 \). 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 \( a_t \), we assume that its initial conditions are given as \( a_0 = (\phi/ (1 - b_0), b_0, 0)' \), where \( b_0 = 0.10 \). For the vector of threshold parameters, \( r \), we consider the following values: \( (r^L, r^R) = (-1.96, 2.05) \). 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: \( \phi + 1 \sim Beta (20, 1.5), \frac{1}{\sigma_y^2} \sim Gamma (4.5, 0.15), \frac{1}{\sigma_\gamma^2} \sim Gamma (1.5, 0.005), \rho \sim U (-1, 1) \). These distributions are often used in the litera-
ture 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 $Corr(\ldots)$, of the generated state variables $h_t$ and $b_t$ over these experiments, based on their smoothed estimates, denoted as $h_{t|n}$ and $b_{t|n}$, respectively. The table also reports average estimates of the mean and variance values of posterior distributions of the structural parameters of the SVEB model, over the 10 experiments.

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. Together with the above graphs, Figure 1 also presents a plot of a return process $r_t$ generated by our SVEB model.

<table>
<thead>
<tr>
<th></th>
<th>True value</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.85</td>
<td>0.8664</td>
<td>0.0067</td>
</tr>
<tr>
<td>$\sigma_\eta$</td>
<td>0.28</td>
<td>0.2832</td>
<td>0.0128</td>
</tr>
<tr>
<td>$\sigma_\gamma$</td>
<td>0.05</td>
<td>0.0485</td>
<td>0.0010</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.59</td>
<td>-0.5803</td>
<td>$1.3\times10^{-2}$</td>
</tr>
<tr>
<td>$b_0$</td>
<td>0.10</td>
<td>0.0224</td>
<td>$2.15\times10^{-4}$</td>
</tr>
<tr>
<td>$Corr(h_t, h_{t</td>
<td>n})$</td>
<td>0.9453</td>
<td>$1.5\times10^{-4}$</td>
</tr>
<tr>
<td>$Corr(b_t, b_{t</td>
<td>n})$</td>
<td>0.9638</td>
<td>$1.7\times10^{-4}$</td>
</tr>
</tbody>
</table>

**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 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 $b_t$.

![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 upper plot), implied by the SVEB model.](image)

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

In this section, we employ the SVEB model to investigate if shifts in volatility of the US stock market aggregate return are driven by large return shocks. 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.

Tables 2A presents the estimates of the SVEB model for all possible combinations of the sets of threshold parameters \( r^L \) and \( r^R \) considered: \( r^L = \{-1.65, -1.96, -2.05, -2.33\} \) and \( r^R = \{1.65, 1.96, 2.05, 2.33\}\). These are obtained based on the Bayesian MCMC method, presented in Section 3. The above values of \( r^L \) and \( r^R \) correspond to negative and positive shock market events \( \left( \frac{\hat{\omega}}{\sigma} < r^L \right) \) and \( \left( \frac{\hat{\omega}}{\sigma} > r^R \right) \), with probabilities to occur 5%, 2.5%, 2% and 1%, respectively. The bottom line of 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 are the same to those assumed by our Monte Carlo study. Analogous priors are used for the SV model. The logarithms of the marginal likelihood values reported in the table, i.e. \( \ln m(y|r) \), indicate that, among all possible combinations of \( r^L \) and \( r^R \) considered, the SVEB model with threshold parameter values \( r^L = -2.05 \) and \( r^R = 2.33 \) constitutes the best specification of the data, as it gives the highest value of \( \log m(y|r) \). This model is also found to describe better the data than the SV model. These results can be also confirmed by the values of log-likelihood ordinate \( \ln g(y|\theta, r) \), reported in the tables.

\footnote{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.}
Table 2A: Estimates of SVEB model for different values of $r=(r^L, r^R)^T$

| $r^L$ | $r^R$ | $\phi$ | $\sigma_g$ | $\sigma_s$ | $\rho$ | $b_0$ | $\ln g(y|\theta, r)$ | $\ln m(y|r)$ |
|-------|-------|-------|-----------|-----------|-------|-------|---------------------|----------------|
| -1.65 | 1.65  | 0.7996| 0.3000    | 0.0501    | -0.62 | 0.3262| -3357.7 $\times 10^3$ | -3369.5 $\times 10^3$ |
| -1.65 | 1.96  | 0.8156| 0.2970    | 0.0510    | -0.616| 0.3151| -3358.2 $\times 10^3$ | -3368.7 $\times 10^3$ |
| -1.65 | 2.05  | 0.8207| 0.2966    | 0.0514    | -0.6144| 0.3080| -3360.5 $\times 10^3$ | -3370.3 $\times 10^3$ |
| -1.65 | 2.33  | 0.8286| 0.299     | 0.0511    | -0.619 | 0.2948| -3357.6 $\times 10^3$ | -3366.0 $\times 10^3$ |
| -1.96 | 1.65  | 0.8063| 0.3055    | 0.0563    | -0.6247| 0.3042| -3356.1 $\times 10^3$ | -3364.2 $\times 10^3$ |
| -1.96 | 1.96  | 0.8312| 0.2950    | 0.0575    | -0.6025| 0.3279| -3361.6 $\times 10^3$ | -3367.9 $\times 10^3$ |
| -1.96 | 2.05  | 0.8363| 0.2902    | 0.0581    | -0.6164| 0.2745| -3359.4 $\times 10^3$ | -3367.5 $\times 10^3$ |
| -1.96 | 2.33  | 0.8683| 0.2798    | 0.0504    | -0.5855| 0.2196| -3359.4 $\times 10^3$ | -3365.5 $\times 10^3$ |
| -2.05 | 1.65  | 0.8193| 0.2966    | 0.0548    | -0.6259| 0.2829| -3357.6 $\times 10^3$ | -3367.7 $\times 10^3$ |
| -2.05 | 1.96  | 0.8428| 0.2835    | 0.0572    | -0.6174| 0.2677| -3362.1 $\times 10^3$ | -3370.2 $\times 10^3$ |
| -2.05 | 2.05  | 0.8437| 0.2875    | 0.0578    | -0.6069| 0.2600| -3362.5 $\times 10^3$ | -3370.4 $\times 10^3$ |
| -2.05 | 2.33  | 0.8772| 0.2724    | 0.0506    | -0.5837| 0.2047| -3357.1 $\times 10^3$ | -3363.1 $\times 10^3$ |
| -2.33 | 1.65  | 0.8336| 0.2876    | 0.0546    | -0.6125| 0.2672| -3366.63 $\times 10^3$| -3375.7 $\times 10^3$ |
| -2.33 | 1.96  | 0.8568| 0.2745    | 0.0600    | -0.6121| 0.2394| -3365.0 $\times 10^3$ | -3372.4 $\times 10^3$ |
| -2.33 | 2.05  | 0.8623| 0.2725    | 0.0611    | -0.6119| 0.2290| -3360.0 $\times 10^3$ | -3366.8 $\times 10^3$ |
| -2.33 | 2.33  | 0.8989| 0.2569    | 0.0518    | -0.5703| 0.1600| -3358.3 $\times 10^3$ | -3364.0 $\times 10^3$ |

More results for the SVEB model with $(r^L = -2.05, r^R = 2.33)$ and the SV model are reported in Table 2B. In addition to posterior means, this table reports posterior standard deviations of the structural parameters and the inefficiency factor (IF) for these two models. The inefficiency factor (or autocorrelation time) 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 values of the standard deviations reported in the table, which are very small for all structural parameters of the SVEB model. A final conclusion that can be drawn from the results of Table 2B concerns the estimates of stochastic volatility autoregressive coefficient $\phi$. This is found to be smaller for the SVEB, compared to the SV. This can be obviously attributed to the fact that the SVEB model allows for structural
breaks.

| Estimates of the SVEB model ( $r^L = -2.05$ and $r^R = 2.33$) |
|------------------|------------------|------------------|------------------|------------------|
| $\phi$          | $\sigma_\eta$   | $\sigma_\gamma$ | $\rho$          | $b_0$          |
| Mean            | 0.8772           | 0.273            | 0.051           | -0.584         | 0.205           |
| Std. Dev.       | 0.001            | 0.002            | 0.003           | 0.007          | 0.006           |

| Estimates of the SV model ($\sigma^2 = 0$) |
|------------------|------------------|------------------|
| Mean            | 0.95             | 0.223            |
| Std. Dev.       | 0.002            | 0.001            |
| IF              | 1.795            | 4.104            |

Table 2B: Estimates of the SVEB and SV models

The estimates of threshold parameters of the SVEB model which are found to better describe the data, i.e. $r^L = -2.05$ and $r^R = 2.33$, indicate that there are significant asymmetries between the large negative and positive return shocks. These asymmetries reveal 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 innovations. 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 will be analyzed more thoroughly in a next section, which presents estimates of the impact news function of the SVEB model.

Another interesting conclusion that can be drawn from the results of Table 2A, or Table 2B, is that the estimates of correlation coefficient $\rho$, which captures the leverage effects of stock return innovations on volatility, change very little with the different values of the threshold parameters considered. The higher estimate of $\rho$ found through the SVEB model, compared to that found by the standard SV model,
can be attributed to the better fit (specification) of the SVEB model into the data. It can also be attributed to the ability of our 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, captured by indicator function $I(\mathcal{A}_t)$. The last function can amplify the leverage effects through the state variable $\gamma_t$, and thus can explain higher values of $\rho$. Finally, the results of Table 2A indicate that the estimates of $\sigma^2_\gamma$, capturing different over time magnitude effects of large return shocks on volatility, remain almost the same across all different values of the vector of threshold parameters $(r^L, r^R)$ considered. This result emphasizes the existence of shifts in volatility of stochastic magnitude over time.

To have a better picture how stock market volatility is affected by large return shocks, Figure 2 graphically presents smoothed estimates of the volatility and break processes, $h_{t|n}$ and $b_{t|n}$, respectively, together with ex-post filtered estimates of return innovations, defined as $\varepsilon_{t|t} = y_t \exp(-\frac{h_{t|t}}{2})$, where $h_{t|t}$ is obtained through the auxiliary particle filter method. Inspection of the plots of this figure indicates that the estimates of volatility $h_{t|n}$ exhibit persistent and non-linear shifts over time. The shifts of $h_{t|n}$ can be explained by those implied by the break process $b_t$. The latter have different magnitude over time. 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).

The estimates of $\varepsilon_{t|t}$, presented in Figure 2, indicate that the changing points of the cyclical shifts of processes $b_{t|n}$ and $h_{t|n}$, observed during our sample, are 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.
5.1 Forecasting volatility based on the SVEB model

To further assess the ability of the SVEB model to fit satisfactorily into the data compared to the SV model, in this section we conduct in-sample and out-of-sample forecasting exercises for the one-period ahead volatility $h_{t+1}$. Table 3 presents the results of these two exercises. In particular, the table reports values of the mean of forecasting error, and the MSE (mean square error) and MAE (mean absolute error) metrics, often used in forecasting exercises as performance measures. Since there are no actual measures of volatility process, in our analysis we will evaluate the
relative forecasting performance of the above models based on filtered estimates of
$h_{t+1}$, defined as $h_{t+1|t} = \mathbb{E}(h_{t+1}|I_t)$. These are obtained by the method of the auxiliary
particle filter. To calculate the out-of-sample volatility forecasts, we will rely on the
parameter estimates of the two models reported in Table 2B based on weekly data
from 07-01-1980 to 31-03-2010. These estimates will be then used to provide volatility
forecasts $h_{t+1|t}$ for the out-of-sample period from 01-04-2010 to 04-04-2011, consisting
of 58 weekly observations which are not used in the estimation of the models. Apart
from the SVEB and SV models, Table 3 also reports values of the mean, and the
MSE and MAE metrics of the forecasting errors for the following volatility models:
the MRS (Markov regime-switching), EGARCH(1,1) and GJR-GARCH(1,1), which
share some common features with the SVEB model.

<table>
<thead>
<tr>
<th>Models</th>
<th>Mean</th>
<th>MSE</th>
<th>MAE</th>
<th>Mean</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVEB($r^L=-2.05, r^R=2.33$)</td>
<td>0.0085</td>
<td>0.9180</td>
<td>0.7760</td>
<td>0.0128</td>
<td>0.8250</td>
<td>0.7319</td>
</tr>
<tr>
<td>SV</td>
<td>0.0089</td>
<td>0.9416</td>
<td>0.7927</td>
<td>0.0219</td>
<td>0.9425</td>
<td>0.7904</td>
</tr>
<tr>
<td>EGARCH(1,1)</td>
<td>-0.0024</td>
<td>1.0017</td>
<td>0.7812</td>
<td>0.0449</td>
<td>1.0073</td>
<td>0.7894</td>
</tr>
<tr>
<td>MRS</td>
<td>-0.0025</td>
<td>0.9648</td>
<td>0.7963</td>
<td>0.0445</td>
<td>0.8719</td>
<td>0.7414</td>
</tr>
<tr>
<td>GJR-GARCH(1,1)</td>
<td>-0.0014</td>
<td>1.0011</td>
<td>0.7789</td>
<td>0.0467</td>
<td>0.9641</td>
<td>0.7696</td>
</tr>
</tbody>
</table>

Table 3: In-and out-of-sample forecasts of volatility $h_{t+1}$

The results of Table 3 clearly indicate that the SVEB model improves considerably
the forecasting performance of the SV model. It also outperforms the three other
parametric volatility models examined. This result holds for both the in-sample
and out-of-sample exercises conducted. It is also robust across the two forecasting
evaluation metrics employed, i.e. the MSE and MAE.
5.2 Estimates of the impact news function implied by the SVEB model

Estimates of the impact news function (INF) implied by the estimates of the SVEB model are reported in Table 2B. This function is frequently used in practice to investigate if the reaction of the expected value of volatility one-period ahead, defined as $\Delta\sigma_{t+1|t} = \ln \left( \frac{\sigma_{t+1|t}}{\sigma_{t|t-1}} \right)$ where $\sigma_{t+1|t} = \exp \left( \frac{h_{t+1|t}}{2} \right)$, to return innovation forecasts, defined as $\varepsilon_{t|t-1} = \frac{h_{t}}{\sigma_{t|t-1}}$, is asymmetric between positive and negative values of $\varepsilon_{t|t-1}$ (see, e.g., Engle and Ng (1991), Bekaert and Wu (2000), Li et al (2005), inter alia). In addition to this, it is employed in the empirical literature as a model specification test and, in particular, to assess if parametric volatility models are consistent with patterns of volatility changes observed in reality. The latter are based on implied by market option prices (e.g., VIX), or realized, measures of volatility. As is shown in the above literature, negative return innovations $\varepsilon_{t|t-1}$ have a greater impact on $\Delta\sigma_{t+1|t}$ than positive return innovations. Their effects on $\Delta\sigma_{t+1|t}$ can be described by a monotonic and negative relationship, which is more asymmetric at its end points, corresponding to large negative or positive return shocks.\(^9\) This relationship is analogous to that between expected volatility changes $\Delta\sigma_{t+1|t}$ and stock returns $r_t$ (see Ederington and Guan (2010)).

Table 4A and Figure 3A present estimates of INFs implied by the SVEB and SV models, for different sub-intervals of return innovations $\varepsilon_{t|t-1}$. In Tables 4B and Figure 3B, we present estimates of the above INFs with respect to stock returns $r_t$. This table and figure also include estimates of INFs based on market measures of volatility given by VIX, defined $\Delta VIX_{t+1} = \ln \left( \frac{VIX_{t+1}}{VIX_t} \right)$\(^10\). These can be

\(^9\)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.

\(^10\)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)).
compared with the INFs implied by the SVEB and SV models. Note that, since VIX is a four-weeks measure of volatility, the expected volatility changes presented in Table 4B and Figure 3B are estimated as $\Delta \sigma_{t+4|t} = \ln \left( \sigma_{t+4|t}/\sigma_{t-3|t} \right)$, where $\sigma_{t+4|t} = \frac{1}{4} \sum_{i=1}^{4} \sigma_{t+i|t}$. The results of Tables 4A-4B and Figures 3A-3B clearly indicate that the SVEB model can produce shapes of INFs which correspond to those observed in reality, based on implied, or realized, estimates of stock market volatility. These are negative and monotonic, and asymmetric at their end points, as is found in the studies mentioned before. Compared to the SV model, the SVEB can produce INFs with higher asymmetry, especially at the sub-intervals of large negative or positive return innovations. Figure 3B clearly indicates that these asymmetries are more close to those 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 3B 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. This obviously can be attributed to the fact that the implied (risk-neutral) 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 are. This result supports the view that the premia effects may increase in terms of magnitude with the size of negative return innovations, which can be expected theoretically.
Return news (shocks) SVEB with \((r^L, r^R)=(-2.05, 2.33)\) SV

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

<table>
<thead>
<tr>
<th>e(t)</th>
<th>volatility change</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>-0.20</td>
</tr>
<tr>
<td>-2</td>
<td>-0.15</td>
</tr>
<tr>
<td>-1</td>
<td>-0.10</td>
</tr>
<tr>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
</tr>
</tbody>
</table>

**Table 4A:** NIFs of the SVEB and SV models.

**Figure 3A:** NIFs of the SVEB and SV models.
<table>
<thead>
<tr>
<th>Values of $r_t$</th>
<th>SVEB with $(r^L, r^R)=(-2.05, 2.33)$</th>
<th>SV</th>
<th>VIX (02/01/90-29/03/10, 1056 obs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>obs. mean of $r_t$</td>
<td>$\Delta \sigma_{t+4}$</td>
<td>$\Delta \sigma_{t+4}$</td>
</tr>
<tr>
<td>$r_t &lt; -3.5$</td>
<td>76</td>
<td>-5.48</td>
<td>0.133</td>
</tr>
<tr>
<td>$-3.5 &lt; r_t \leq -2.50$</td>
<td>73</td>
<td>-2.93</td>
<td>0.094</td>
</tr>
<tr>
<td>$-2.50 &lt; r_t \leq -2.0$</td>
<td>66</td>
<td>-2.25</td>
<td>0.079</td>
</tr>
<tr>
<td>$-2.0 &lt; r_t \leq -1.5$</td>
<td>101</td>
<td>-1.75</td>
<td>0.064</td>
</tr>
<tr>
<td>$-1.50 &lt; r_t \leq -1$</td>
<td>109</td>
<td>-1.25</td>
<td>0.059</td>
</tr>
<tr>
<td>$-1 &lt; r_t \leq -0.5$</td>
<td>123</td>
<td>-0.75</td>
<td>0.035</td>
</tr>
<tr>
<td>$-0.5 &lt; r_t \leq 0$</td>
<td>141</td>
<td>-0.26</td>
<td>0.021</td>
</tr>
<tr>
<td>$0 &lt; r_t \leq 0.5$</td>
<td>155</td>
<td>0.24</td>
<td>-0.004</td>
</tr>
<tr>
<td>$0.5 &lt; r_t \leq 1$</td>
<td>180</td>
<td>0.73</td>
<td>-0.027</td>
</tr>
<tr>
<td>$1 &lt; r_t \leq 1.5$</td>
<td>176</td>
<td>1.23</td>
<td>-0.038</td>
</tr>
<tr>
<td>$1.5 &lt; r_t \leq 2.0$</td>
<td>110</td>
<td>1.73</td>
<td>-0.059</td>
</tr>
<tr>
<td>$2.0 &lt; r_t \leq 2.5$</td>
<td>86</td>
<td>2.23</td>
<td>-0.069</td>
</tr>
<tr>
<td>$2.5 &lt; r_t \leq 3.5$</td>
<td>92</td>
<td>2.94</td>
<td>-0.091</td>
</tr>
<tr>
<td>$r_t &gt; 3.5$</td>
<td>84</td>
<td>4.95</td>
<td>-0.11</td>
</tr>
</tbody>
</table>

Table 4B: NIFs of the SVEB and SV models, and VIX

Figure 3B: NIFs of the SVEB and SV models, and VIX with respect to stock return $r_t$.  

26
5.3 Impulse response functions of large stock return shocks

To study the dynamic pattern of the effects of return innovations $\varepsilon_t$ on the future values of volatility, in this section we calculate the impulse response functions (IRFs) of future sequences of volatility $h_{t+\tau|t}$, for horizons $\tau = 1, 2, 3, \ldots$, with respect to $\varepsilon_t$. Our analysis is focused on the effects of large shocks on volatility, identified through the SVEB model. Since the 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. This function 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 $E[h_{t+\tau|t} \varepsilon_t, a_t]$ and $E[h_{t+\tau} a_t]$, i.e.

$$
GIRF(\tau, \varepsilon_t, a_t) = E[h_{t+\tau} | \varepsilon_t, a_t] - E[h_{t+\tau} | a_t], \text{ for } \tau = 1, 2, \ldots, \quad (13)
$$

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

The random nature of $GIRF(\tau, \varepsilon_t, a_t)$ provides a more flexible approach of analyzing the effects of $\varepsilon_t$ on $h_{t+\tau}$. Someone can condition on a particular shock $\varepsilon_t$ and treat $a_t$ as a random vector, or she/he can condition on a particular history $a_t$, treating $\varepsilon_t$ as a random variable. Another possibility is to condition on particular subsets of the history of $\varepsilon_t$, i.e. large positive, or negative, values of $\varepsilon_t$, which is the main focus of our analysis. To this end, we will first define the sets $B^{(1)} = (-\infty, -2.05)$ and $B^{(2)} = (2.33, \infty)$, consisting of the large negative and positive return shocks identified by the estimates of the threshold parameters of our SVEB model. Then, we will calculate $GIRF(\tau, \varepsilon_t, a_t)$, at time $t$, for $\varepsilon_t \in B^{(1)}$ and $\varepsilon_t \in B^{(2)}$, respectively.
We will estimate the distribution of $GIRF(\tau, \varepsilon_t, a_t)$ by means of a Monte Carlo integration. First, we will pick up 500 series of $a_t$ from its sample estimates. To compute the first expectation of (13), $E[h_{t+\tau} | \varepsilon_t, a_t]$, we proceed as follows: When $\tau = 1$, for each of the 500 series of $a_t$, we will draw 30 realizations of the large return shock $\varepsilon_t$ from the truncated normal distribution $TN_{B(\cdot)}(0, 1)$, which is truncated at $B(i), i = 1, 2$. 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. $\eta_t \sim N(\rho_\eta \varepsilon_t, \sigma_\eta^2 (1 - \rho^2))$, and will average out to compute expectation $E[h_{t+\tau} | \varepsilon_t, a_t]$. For $\tau > 1$, to estimate $E[h_{t+\tau} | \varepsilon_t, a_t]$ 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 $h_{t+\tau}$ recursively, based on its data generating process. The average of the 1000 series of $h_{t+\tau}$ generated will give the estimate of $E[h_{t+\tau} | \varepsilon_t, a_t]$, for $\tau > 1$. To compute the second expectation of (13), $E[h_{t+\tau} | a_t]$, 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 $E[h_{t+\tau} | a_t]$. Given the above estimates of $E[h_{t+\tau} | \varepsilon_t, a_t]$ and $E[h_{t+\tau} | a_t]$, we can then obtain estimates of $GIRF(\tau, \varepsilon_t, a_t)$, for all $\tau$, through equation (13). To obtain the density of these estimates of $GIRF(\tau, \varepsilon_t, a_t)$, we will use a normal kernel.

Figure 4A presents the marginal densities of $GIRF(\tau, \varepsilon_t, a_t)$ implied by the SVEB model, for $\tau = 1, 2, ..., 20$, based on the procedure described above. This is done for negative and positive large return shocks, i.e. $\varepsilon_t \in B^{(1)}$ and $\varepsilon_t \in B^{(2)}$. For useful comparisons, in Figure 4B we present densities of $GIRF$ for the SV model, when $\varepsilon_t \in B^{(1)}$ and $\varepsilon_t \in B^{(2)}$. These are estimated based on a Monte Carlo integration method analogous to that presented above, for the SVEB model. To better understand features of the above densities related to the traditional IRFs, in Table 5 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 3A-3B and Table 5 lead to a number of very useful conclusions about the dynamic effects of large return shocks on the future paths of volatility.

<table>
<thead>
<tr>
<th>Period</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau = 1 )</td>
<td>0.3842</td>
<td>0.0029</td>
<td>1.4794</td>
<td>6.1243</td>
<td>0.2644</td>
<td>0.0014</td>
<td>1.4304</td>
<td>5.8149</td>
</tr>
<tr>
<td>( \tau = 2 )</td>
<td>0.3372</td>
<td>0.0022</td>
<td>1.4423</td>
<td>5.7246</td>
<td>0.2516</td>
<td>0.0013</td>
<td>1.3618</td>
<td>5.6020</td>
</tr>
<tr>
<td>( \tau = 3 )</td>
<td>0.2954</td>
<td>0.0017</td>
<td>1.3968</td>
<td>5.5584</td>
<td>0.2399</td>
<td>0.0011</td>
<td>1.3807</td>
<td>5.6184</td>
</tr>
<tr>
<td>( \tau = 4 )</td>
<td>0.2603</td>
<td>0.0014</td>
<td>1.3946</td>
<td>5.5507</td>
<td>0.2288</td>
<td>0.0010</td>
<td>1.4772</td>
<td>6.2909</td>
</tr>
<tr>
<td>( \tau = 5 )</td>
<td>0.2278</td>
<td>0.0011</td>
<td>1.3691</td>
<td>5.5613</td>
<td>0.2179</td>
<td>9.4 \times 10^{-4}</td>
<td>1.4425</td>
<td>6.0307</td>
</tr>
<tr>
<td>( \tau = 10 )</td>
<td>0.1185</td>
<td>3.9 \times 10^{-4}</td>
<td>0.8923</td>
<td>4.4514</td>
<td>0.1715</td>
<td>5.9 \times 10^{-4}</td>
<td>1.5020</td>
<td>6.2402</td>
</tr>
<tr>
<td>( \tau = 15 )</td>
<td>0.0613</td>
<td>2.5 \times 10^{-4}</td>
<td>0.2173</td>
<td>3.2368</td>
<td>0.1348</td>
<td>3.7 \times 10^{-4}</td>
<td>1.4280</td>
<td>5.7860</td>
</tr>
<tr>
<td>( \tau = 20 )</td>
<td>0.0319</td>
<td>2.6 \times 10^{-4}</td>
<td>0.0586</td>
<td>3.0158</td>
<td>0.1058</td>
<td>2.2 \times 10^{-4}</td>
<td>1.4463</td>
<td>6.0751</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Period</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau = 1 )</td>
<td>-0.4244</td>
<td>0.0026</td>
<td>-1.4358</td>
<td>5.7257</td>
<td>-0.2920</td>
<td>0.0012</td>
<td>-1.3809</td>
<td>5.4554</td>
</tr>
<tr>
<td>( \tau = 2 )</td>
<td>-0.3721</td>
<td>0.0020</td>
<td>-1.4572</td>
<td>5.9072</td>
<td>-0.2788</td>
<td>0.0011</td>
<td>-1.3806</td>
<td>5.5904</td>
</tr>
<tr>
<td>( \tau = 3 )</td>
<td>-0.3267</td>
<td>0.0015</td>
<td>-1.4751</td>
<td>6.0474</td>
<td>-0.2653</td>
<td>0.0010</td>
<td>-1.3672</td>
<td>5.4841</td>
</tr>
<tr>
<td>( \tau = 4 )</td>
<td>-0.2859</td>
<td>0.0012</td>
<td>-1.4622</td>
<td>5.9919</td>
<td>-0.2527</td>
<td>9.2 \times 10^{-4}</td>
<td>-1.4478</td>
<td>5.9876</td>
</tr>
<tr>
<td>( \tau = 5 )</td>
<td>-0.2513</td>
<td>9.3 \times 10^{-4}</td>
<td>-1.3392</td>
<td>5.4702</td>
<td>-0.2407</td>
<td>8.2 \times 10^{-4}</td>
<td>-1.4252</td>
<td>5.6976</td>
</tr>
<tr>
<td>( \tau = 10 )</td>
<td>-0.1304</td>
<td>3.5 \times 10^{-4}</td>
<td>-0.8354</td>
<td>4.4029</td>
<td>-0.1894</td>
<td>5.2 \times 10^{-4}</td>
<td>-1.5065</td>
<td>6.4520</td>
</tr>
<tr>
<td>( \tau = 15 )</td>
<td>-0.0678</td>
<td>2.4 \times 10^{-4}</td>
<td>-0.2078</td>
<td>3.1435</td>
<td>-0.1487</td>
<td>3.2 \times 10^{-4}</td>
<td>-1.4267</td>
<td>5.7858</td>
</tr>
<tr>
<td>( \tau = 20 )</td>
<td>-0.0351</td>
<td>2.5 \times 10^{-4}</td>
<td>-0.0109</td>
<td>3.0047</td>
<td>-0.1170</td>
<td>2.0 \times 10^{-4}</td>
<td>-1.3982</td>
<td>5.5512</td>
</tr>
</tbody>
</table>

Table 5: Descriptive statistics of GIRQFs of the SVEB and SV models, for different \( \tau \)

First, the values of the descriptive statistics reported in the table indicate that the densities of \( GIRQF \) for both the SVEB and SV model converge to zero as \( \tau \) increases, given that their variance goes to zero with \( \tau \). However, the values of the mean, skewness and kurtosis statistics reported in the table indicate that the convergence rate of the SVEB model is higher than that of the SV model. The values of these statistics become smaller for the SVEB model compared to the SV model, as \( \tau \) increases, and, in particular, for \( \tau > 5 \). For \( \tau = 20 \), the mean of \( GIRQF \) is very close to zero for the SVEB model, while the skewness and kurtosis coefficients of \( GIRQF \)
are very close to the values of these coefficients implied by the normal distribution, which are 0 and 3, respectively. As can be also confirmed from Figures 4A and 4B, the densities of $GIRF$ for the SV model remain skewed and leptokurtic, and their mean is not very close to zero even for very large values of $\tau$, i.e. $\tau = \{10, 20\}$. The above results hold for both large negative and positive return shocks. They can be taken to support the view that the effects of large return shocks on future paths of volatility may be absorbed faster by the SVEB model, rather than the SV model. This is much more likely to happen. It can be attributed to the fact that, by allowing for abrupt shifts in volatility triggered by large return shocks, the SVEB model can absorb faster the impact of large return innovations on volatility, compared to the traditional SV model which does not allow for shifts in volatility. Moreover, the mean values of $GIRF$, reported in Table 5, indicate that the SVEB model is more likely to absorb the biggest effects of these innovations on volatility.
Figure 4A: Estimates of the densities of GIRF for the SV model, across different horizons $\tau$. 
Figure 4B: Estimates of the densities of GIRF for the SVEB model, across different horizons $\tau$.

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 breaks in the empirical finance literature. These shifts are endogenously driven by large asset (stock) return shocks, defined as being bigger than the values of threshold parameters which can distinguish large negative shocks from positive shocks. Some of these shocks can be associated with extraordinary events in asset (or stock) markets like financial crises.

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 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 volatility of the US stock market return are endogenously driven by large return shocks, reflecting substantial pieces of stock market news. Given this, then it employs the model to study the dynamic effects of these shocks on volatility. The empirical analysis of the paper leads to a number of interesting conclusions. First, it identifies as large negative shocks those which are less than the -2.05% on weekly basis, while as large positive shocks those which are bigger than 2.33%. Second, it indicates that cyclical shifts in the volatility of the US market can be triggered by sequences of large return shocks. These shifts are persistent and their magnitude change significantly over the time. The turning points of these shifts are found to be clearly driven by large return shocks associated with extraordinary stock market news such as financial crises. Third, it shows that the estimates of the model can explain asymmetric patterns of responses of stock market volatility to large return shocks is asymmetric observed in reality, especially for large negative or positive values of these shocks. Finally, it indicates that large stock market shocks are absorbed by the market faster than predicted by the standard stochastic volatility model.
References


Appendix

In this appendix, we provide proofs of Theorems 1, presented in the main text. It also presents the auxiliary particle filtering.

Proof of Theorem 1

We now prove strict stationarity for \( h_t \), given by

\[
y_t = \exp \left( -\frac{h_t}{2} \right) \epsilon_t
\]
\[ h_{t+1} = b_{t+1} + \phi h_t + \eta_t \quad \text{and} \quad b_{t+1} = \delta_t b_t + I(\mathcal{A}_t) \gamma_t, \]

where

\[ \delta_t = \begin{cases} 
1 & \text{if } I(|b_t| < b) \\
0 & \text{otherwise.} 
\end{cases} \quad (14) \]

The first step is to derive a recursive representation for \( h_t \). This is given by

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

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

\[ E(h_t^\alpha) < \infty. \]

By The Marcinkiewicz-Zygmund inequality we have that

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

which is finite as long as \( b_t \) is strictly stationary and \( E(b_{t-j})^\alpha < \infty \) and \( E(\eta_{t-j-1})^\alpha < \infty \). Thus it suffices to prove that \( b_t \) is strictly stationary and \( E(b_{t-j})^\alpha < \infty \). \( E(b_{t-j})^\alpha < \infty \) follows easily from strict stationarity and \( E(\gamma_t)^\alpha < \infty \). 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

\[ E[\|b_t\| \mid b_{t-1} = \vartheta \|d\| + L, \ \forall d \notin C], \quad (15) \]

\[ E[\|b_t\| \mid b_{t-1} = \vartheta \|d\|] \leq B, \ \forall d \in C, \quad (16) \]

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 (16) follows
easily. We need to show (15). (15) follows if the following condition holds

\[ E(\delta_t) < 1. \]  

(17)

To prove (17) 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 \( b \).