Kelvin Balcombe

University of Reading

and

Iain Fraser, University of Kent and La Trobe University

Manuscript forthcoming in Economic Modelling

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license

http://creativecommons.org/licenses/by-nc-nd/4.0

Do Bubbles have an Explosive Signature in Markov Switching Models?

Abstract

We investigate nine data series previously identified as containing bubbles using Bayesian Markov switching models. Nearly all series appear to display strong regime switching that could possibly be induced by 'bubble' processes, but in each case the type of model that best describes each price differs substantively. We pay particular attention to whether these series contain transient explosive roots, a feature which has been suggested to exist in several bubble formulations. Bayesian model averaging is employed which allows us to average across a range of submodels, so that our empirical findings are not based on only one well performing model. We show that explosive regimes may exist in many submodels, but only when the flexibility of the model is limited in other important respects. In particular, when Markov switching models allow for switching levels of error variance, explosive root regimes occur in only a minority of the series.

Key Words: Explosive Root Regimes, Transient Explosive Roots, Bubbles, Bayesian Model Averaging.

JEL Classification: C52, E3

1 Introduction

There has been a long standing interest in the idea that asset prices may exhibit bubbles (e.g., Garber 1990; Malkiel, 2012). This interest has been particularly apparent in the analysis of prices in markets that are subject to speculation with numerous papers supporting the existence of bubbles. For example, aggregate

prices (Hall et al., 1999, henceforth referred to as HPS), oil prices (e.g., Shi and Arora, 2012; Zhang and Yao, 2016), stock market prices (e.g., Narayan et al., 2013; Chen et al., 2016; Shi and Song, 2016) and house prices (e.g., Phillips and Yu, 2011) have all been found to contain bubbles. There is also extensive behavioural experimental evidence supporting the existence of bubbles (e.g., Shiller, 2003). Yet, while there is widespread support[1] for the idea that some economic series exhibit bubbles, there are several ways of defining bubbles. Bubbles are an evocative yet imprecise metaphor when applied to prices, perhaps inviting people to think that prices must 'float' somewhat unpredictably upward then 'pop' in the sense that they suddenly collapse. To others the term bubble might signify that a market or price is somehow cut off (albeit temporarily) from the fundamental forces shaping the wider economy. Different types of bubbles are commonly differentiated in terms of being fundamental (intrinsic) or speculative (extrinsic), rational or irrational (Gurkaynak, 2008). This complicates the empirical identification of bubbles, since the term can describe a range of phenomena. Commonly, however, the word 'explosive' has been employed to describe bubbles and a strand of the literature has gone further by proposing that the Transient Explosive Root (TER) property of a series is a 'bubble signature' (e.g., HPS, 1999; Phillips and Yu, 2011; Phillips et al., 2015; Shi and Song, 2016).[2]

In this paper, we first inquire as to whether alternative bubble formulations necessarily imply that we should see explosive regimes in Markov switching models. We observe that while the wider literature often refers to the 'explosive' nature of bubbles, it does not generally infer that TERs are a necessary condition for the existence of bubbles. For example, in the rational intrinsic bubble model of Froot and Obstfeld (1991), a large and increasing divergence of a price from its fundamental value is labelled 'explosive', even though there may be no sudden change upwards or downwards, and the series need not display TER behaviour. Likewise, the bubble of Blanchard and Watson (1982) is regime switching, but not of an explosive root regime (ERR) variety and Evans (1991) provides a formulation, which has often been considered as explosive, which need not strictly be explosive in the TER/ERR sense.

The significance of identifying a TER within the bubble literature is mostly due to Phillips and Yu (2011). Subsequently, Phillips (and various co-authors) developed the econometric theory enabling the testing for 'right valued alternatives' to the unit root hypothesis (see Phillips et al., 2011; Phillips et al., 2015). They interpret the rejection of the unit root in favour of the right valued alternative as the rejection of the 'no bubble' hypothesis. In related research, a Bayesian Markov switching approach has been used by Shi (2013) and Shi and Song (2016), searching for ERRs as a bubble signature following HPS (1999) who showed that

[1] The awarding of the 2013 Nobel prize to economists holding a different perspective about the existence of bubbles has reinvigorated interest in this topic. It was awarded to Robert Shiller, Eugene Fama and Lars Peter Hansen, with a number of commentators highlighting the receipents differences in opinion regarding bubbles e.g., A very Rational Award, The Economist, Oct 19th 2013.

[2] We use the term TER to refer any case where the series displays an explosive root which is not permanent. We also use the term ERR as a specific form of TER, that it is specific to regime switching models.

Evans (1991) bubbles can be identified using ERRs in an autoregressive Markov switching model.[3]

The posited connection between explosivity and certain types of bubbles serves as an additional motivation for searching for TERs. But, we do not only focus on TERs. In the empirical part of our paper, we examine nine data series that have previously been identified as containing bubbles. These include the WuLiang Put Warrant data (Xiong and Yu, 2011) and Bitcoin prices. Both are strong candidates for containing bubbles. The former because, as demonstrated by Xiong and Yu (2011), there is a convincing case that the prices observed could not possibly be justified from their fundamentals, and the latter because the fundamental price should be close to zero, yet Bitcoins continue to hold substantial value. We examine whether these series contain common features that may constitute a 'bubble signature'.

Identifying a signature is difficult because the same data can be often be explained well by very different models. Existing work has established that for many series thought to contain bubbles, an autoregressive representation with a constant error variance does a poor job of modelling their behaviour. In contrast, permitting parameters to switch back and forward between regimes improves model performance. But, allowing all parameters to switch is less than ideal if only a few need to. Our empirical work shows that if one part of the model is held constant (e.g., not being regime specific or a parameter is set to zero) what remains flexible will do the 'explanatory work'. In this sense there is 'exchangeability' between models. There is nothing particularly new about this. For example, an ARMA model has both an AR and MA representation, and if one part is suppressed, the other will become more important. Unlike the ARMA example, we do not believe there is an exact exchangeability[4], between our submodels. However, models may be nearly exchangeable in the sense that several models may perform almost equally well. The classical statistical approach is to select models based on some testing down procedure, or because they are the better of two (or more) models no matter how small the difference in performance criteria. By contrast the Bayesian model averaging (BMA) approach, we employ, recognises that many models have merits and averages across the results/estimates in a way that gives better performing models more weight. We use the marginal likelihood (a Bayesian measure of the adequacy of a model) to discriminate among many models and to construct the model weights, employing the approach introduced by Frühwirth-Schnatter (2001, 2004). The marginal likelihoods allow us to see whether price behaviour can sometimes be explained by quite different models. But, unlike standard model selection procedures, BMA does not require us to select a specific model. Nor does it require us to adopt arbitrary sequential reduction sequences that determine parsimonious model selection. Thus, we see the BMA approach as being critical to addressing whether TERs are a feature of the dominant submodels of which there may be

[3] As shown in Homm and Breitung (2012), tests based on right valued alternatives to the unit root also identify Evans (and other) bubbles in Monte Carlo studies.

[4] We cannot prove this conjecture, but we are not aware of any work that would suggest that this is the case

many.

Our research also differs to the existing Bayesian research by Shi and Song (2016). We allow for averaging over lag lengths and a number of other restrictions, the selection of which can be pivotal, whereas Shi and Song (2016) allow for endogenous selection of multiple regimes with fixed lag lengths. Importantly, we offer a more definitive approach than previous Bayesian papers for testing for ERRs, by calculating the marginal likelihood with and without the unit root imposed, since the imposition of a unit root throughout the sample period excludes the possibility of an ERR. In addition, our model specification also extends the existing literature by simultaneously allowing for shifts in error variance as well as t-distributed errors where degrees of freedom for the t–distribution are estimated endogenously. The importance of allowing for shifts in the error variance within a Markov switching specification has been noted by Shi (2013), who found that there may be bias in favour of finding bubbles unless provision is made for error variance shifts. Our results support Shi (2013) such that when using a Markov switching approach, clear ERRs will more commonly manifest when a constant variance is imposed. Thus, for many of the series considered in this paper the imposition of a constant variance is often pivotal.

The paper proceeds by first examining whether all bubbles should contain TERs from a theoretical perspective in Section 2. In Section 3, we introduce the regime switching model that we employ in the empirical section. In this section we also detail the restrictions that identify the special cases of interest. Given the use of Bayesian methods in this paper, we also describe and explain our choice of priors. Next we discuss estimation with details being relegated to appendices. The following sections present the empirical results with discussion and the last section concludes.

2 Models Specifications and Statistical Tests for Bubbles

We begin by considering various bubble model specifications, what this implies for data and what are the implications for econometric testing. We restrict our attention here to some formalised structures that exist in the literature, rather than the wider literature that considers the nature of bubbles. Thus, we focus on several rational extrinsic and intrinsic bubble models that have played an important role in the literature.

2.1 Rational Intrinsic and Extrinsic Bubbles

These models posit the existence of a 'fundamental price', which is the sum of the discounted future dividends of an asset. Extrinsic models posit that the actual price of an asset is equal to this fundamental price plus a bubble component which is not a function of dividends (e.g., Blanchard and Watson, 1982; Evans, 1991;

Brooks and Katsaris, 2005). Intrinsic models either specify the bubble component and make it a function of dividends (e.g., Froot and Obstfeld, 1991) or posit that the bubble is itself within the fundamental price (e.g., Phillips and Yu, 2011).

Bubble models are rational providing they observe the sub-martingale property, such that for a normal and constant rate of return r, the bubble at time t (B_t) obeys

$$E_t B_{t+1} = (1 + r) B_t \tag{1}$$

where E_t is the expectations operator at time t. As already noted, Phillips and Yu (2011) posit a time varying rate of return that can unambiguously generate bubbles of an explosive nature. However, not all bubbles necessarily have this property.

2.1.1 Non-Explosive Rational Bubbles

The model of Blanchard and Watson (1982) posits that the bubble component at time t (B_t) is of a regime switching type with one component having a significantly faster upward trajectory than the other. They observe that the sub-martingale property can easily be satisfied in a regime switching context. The model in Froot and Obstfeld (1991) also satisfies this condition, however, they show that the bubble can be made to depend on dividends in the instance where the dividends are a random walk. Likewise the model of Brooks and Katsaris (2005) has the sub-martingale property. Importantly, however, these models are not necessarily explosive in the TER sense. For example, while Blanchard and Watson (1982) is essentially a regime switching model between high and low growth rates, the model in Brooks and Katsaris (2005) has potentially complex non-linear behaviour, with high growth rates depending inter alia on high trading volumes, such that it is not explosive in the narrow TER sense.

Economists generally distinguish pure volatility from bubbles, and rightly so. However, as observed by Froot and Obstfeld (1991), rational bubbles that are prone to periods of growth and collapse can be generated from changes in the variances of a process. For example, let $B_{t+1} = (1 + r) B_t u_{t+1}$ where u_{t+1} is log normal with a mean of one, but with regime switching variance. In periods of very high volatility B_{t+1} will in probability fall and in periods of very low volatility it will rise. When using logged data this will manifest as a regime switching variance and intercept component, since $\ln\left(\frac{B_{t+1}}{B_t}\right) = \ln(1 + r) + \ln u_{t+1}$ and $E(\ln u_{t+1}) = \frac{-\sigma_t^2}{2}$ where σ_t^2 is the variance of $\ln u_{t+1}$. If σ_t^2 has periods where $\sigma_t^2 > 2\ln(1 + r)$, there will be periods that tend to 'collapse' even though expected returns are unchanged.

2.2 Evans (1991)

A model that has garnered particular interest in the literature is Evans (1991). The specification of the Evans' bubble (B_t) is:

$$\begin{aligned} B_{t+1} &= (1+r) B_t u_{t+1} \text{ if } B_t \leq \alpha \\ B_{t+1} &= \left(\delta + \pi^{-1}(1+r)\theta_{t+1}\left(B_t - (1+r)^{-1}\delta\right)\right) u_{t+1} \text{ if } B_t \geq \alpha \end{aligned} \tag{2}$$

where t is time, u_{t+1} is strictly positive, *iid* and with mean one and $0 < \delta < \alpha(1+r)$. The variable θ_t takes the value one with probability π otherwise it is 0. In the event that $B_t > \alpha$ and $\theta_t = 0$ the process 'collapses' to a (potentially) small value δu_{t+1}. However, if $\theta_{t+1} = 1$ the bubble persists and has a reasonable probability of a very high jump. We refer to the upper regime as the unexcited phase, with the lower being the excited phase.

The Evans (1991) bubble model also has the sub-martingale property $E(B_{t+1}) = B_t(1+r)$. The fact that this property holds can be understood intuitively by the fact that when $B_t > \alpha$ the bubble might collapse, but it also has a countervailing probability of a very large gain. For this reason B_t can be labelled as potentially explosive when in the excited phase.

If the bubble is in its excited phase ($B_t \geq \alpha$) and does not collapse then

$$\ln\left(\frac{B_{t+1}}{B_t}\right) = w_t + \ln(1+r) + \ln(u_{t+1}). \tag{3}$$

The difference in behaviour when excited (but not collapsing) relative to it being in the unexcited phase is that it has the additional component (when excited)

$$w_t = \ln\left(\frac{(\pi-1)\delta}{B_t(1+r)} + 1\right) - \ln(\pi) \tag{4}$$

We observe that

$$\ln\left(\left[\frac{(\pi-1)\delta}{\alpha(1+r)} + 1\right]\pi^{-1}\right) < w_t < \ln\left(\pi^{-1}\right) \tag{5}$$

by recalling that $\left[\frac{(\pi-1)\delta}{\alpha(1+r)} + 1\right]$ must be between zero and one due to the condition $0 < \delta < \alpha(1+r)$. Therefore, as $B_t \to \infty$ the more explosive the series potentially becomes, yet w_t has an asymptote $\ln\left(\pi^{-1}\right)$. Thus, as the bubble becomes very large but does not collapse it approaches the process

$$\ln\left(\frac{B_{t+1}}{B_t}\right) = -\ln(\pi) + \ln(1+r) + \ln(u_{t+1}). \tag{6}$$

The bounded nature of w_t means B_{t+1} is not strictly explosive in the TER sense[5]. However, in one sense a bubble is more explosive when π is small, though more prone to collapse.

There is a further point worth making here in relation to the explosivity of the Evans' bubble. If instead of specifying δ as a constant, but instead we assume $\delta_t = \psi B_t$, then the sub-martingale property is preserved without having a smooth transition to a regime with a higher rate of growth (but with probability of collapse). Thus, this process need not exhibit TER tendencies, only the tendency to partially collapse periodically.

2.3 Detecting Explosive Bubbles

HPS (1999) show that for some parameterisations, an Evans' bubble will exhibit ERRs that are detectable using Markov switching models. Moreover, Phillips et al. (2015) investigate the performance of the Sup Augmented Dickey Fuller (SADF) and Generalized Sup Augmented Dickey Fuller (GSADF) tests to detect an Evans' bubble.[6] Likewise Homm and Breitung (2012) also conduct Monte Carlo tests to evaluate the power of tests to detect an Evans bubble.

Homm and Breitung (2012, Table 6) investigate a number of methods that have the power to reject the no-bubble hypothesis including the SADF test. They demonstrate that the SADF test has maximum (and good) power in detecting Evans' bubbles when π is close to one. According to the arguments discussed, this may seem contradictory, since tests for the Evans' bubble using unit root tests with right tailed alternatives are at their most powerful when the bubble is in one sense less explosive. This happens because when π is high the bubble often persists in its growth phase for long periods, meaning that tests based on explosiveness acquire greater power. This serves to demonstrate, however, why we might not see explosive behaviour when using a regime switching approach even when it is detected by a procedure such as the SADF.

On the other hand, if one simply specifies a process such as $\Delta y_t = \alpha_t + e_t$ for $t = 1, ..., 150$, and $e_t \overset{iid}{\sim} N(0, 1)$ where $\alpha_t = 1$, $t \in [101, 120]$ and 0 otherwise, then the application of the GSADF test to this process rejects the 'no bubble' hypothesis around 85% of the time[7]. Yet, the above process is neither explosive in the TER sense, nor would we would conjecture that many economists would agree that a regime switch or structural break necessarily constitutes a bubble, even though as shown above, it is entirely possible that this is generated by what might be called a bubble.

One further observation is that if a bubble is generated at one data frequency (e.g., daily) but we observe

[5] Note, this does not imply that an Evans bubble will not appear to have explosive regimes when π deviates from unity, as we demonstrate in our results.

[6] Phillips and Yu (2011) is a particularly important contribution, since it provides one of the principle theoretical mechanisms that have linked the TER process to a bubble. They propose that fundamental prices are being calculated using a time varying discount rate that can be shown to induce explosive behaviour, which is in contrast with many other models which posit a stable discount rate.

[7] This was based on 2,000 trials setting the lag length to one, with minimum window of 40.

the data at lower frequencies (e.g., monthly, quarterly) then the probability that the bubble may have collapsed between observations will again dampen any possibility to see explosive tendencies. The greater the difference between the actual bubble frequency and observational frequency of the bubble, the more this will tend to look like 'noise'.

2.4 Summary

To summarise, TERs are a property within some, but certainly not all, bubble formulations. Thus, whether TERs are present in empirical series is worthy of investigation, but caution needs to be exercised in treating TERs as a bubble signature within a Markov switching type model. While tests for right side alternatives to the unit root hypothesis certainly have the power to detect bubbles, they may also detect processes that some economists would not readily call a bubble. However, an empirically important question remains: Are TERs evident in series which have been thought to contain bubbles? We turn to this question in the remainder of the paper.

3 Model Specification

3.1 Regime Switching Models

Since the seminal work of Hamilton (1989), Markov switching models have been enormously popular. Hamilton (1989) limited the switching to take place only between the intercepts in models explaining GNP growth. However, since then models have been generalised to allow switching over all coefficients in models as well as the variances. Estimation can be performed classically or in a Bayesian fashion. The Bayesian approach has been greatly facilitated by the methods outlined in Chib (1996) which showed how to efficiently draw the state variables governing the regime in each time period. The class of models, we deal with in this paper are the autoregressive Markov switching models which are characterised by parameters that depend on a regime (τ_t) existing at a given time t.

Specifically, we represent the data $\{y_t\}$ as

$$\begin{aligned}
\Delta y_t &= \alpha_{\tau_t} + \rho_{\tau_t} y_{t-1} + \sum_{k=1}^{K} \theta_{k,\tau_t} \Delta y_{t-k} + \left(\lambda_t^{-\frac{1}{2}}\right) \sigma_{\tau_t} e_t \qquad (7) \\
&\quad \text{where} \\
t &= 1, ..., T \\
\tau_t &\in \{1, 2, ..., R\} \text{ s.t. } R = 1, 2, 3, ..., R_{\max} \\
K &= 0, 1, 2, ..., K_{\max} \\
&\quad e_t \overset{iid}{\sim} N(0,1)
\end{aligned}$$

In this model R is the number of regimes, where within this paper R is limited to the the values 1,2 or 3 ($R_{\max} = 3$ meaning that we consider models without regime changes along with models that have either 2 or 3 regimes). The (potentially) regime dependent parameters within [7] are α_{τ_t}, ρ_{τ_t} $\{\theta_{k,\tau_t}\}, \sigma_{\tau_t}$ and λ_t. We allow for K (i.e., lag length) to take the values 0 through to $K_{\max} = 12$ because we are dealing with monthly data. The volatility of the error is governed by σ_{τ_t} and λ_t, with σ_{τ_t} being the regime specific parameter governing the error variance. The term λ_t in [7] allows for an error specific variance, and as discussed in Geweke (1993). This term can be viewed as representing a model with a mixture of normals (across variances) but equivalently it can be interpreted as endowing the error with a Student t-distribution. The degrees of freedom (of the t-distribution) can be estimated along with the other parameters of the model, and when the degrees of freedom become large, the errors have an approximate normal distribution. This innovation was introduced into mixture models by Deschamps (2006). Finally, regime switching from state i to j is governed by the Markov process (where $P(|)$ denotes conditional probability)

$$P(\tau_t = j | \tau_{t-1} = i, \{\tau_{t-k}\}_{k=2}^{\infty}) = P(\tau_t = j | \tau_{t-1} = i) = \eta_{i,j} \text{ for all } t \qquad (8)$$

which indicates that the regime in any period only depends on the regime in the preceding period.

3.2 Explosive Root Regimes

The parameterisation for the autoregressive model that is adopted in [7] is the familiar form used in the Augmented Dickey-Fuller test, and adopted by HPS (1999), the first paper to investigate bubbles in price series. Importantly, if in one of the regimes τ, if $\rho_\tau > 0$, (or $\rho_\tau < -2$) this would be evidence of an ERR. However, even if $\rho_\tau = 0$ for all τ, this does not preclude ERRs in the differenced series. That is, with $\rho_\tau = 0$, $\tau = 1, ..., R$, imposed, the polynomial characterising the autoregressive process in differences may have roots

inside the unit circle. Also, because we can also reparameterise [7] for $K \geq 1$ as

$$\Delta^2 y_t = \alpha_{\tau_t} + \rho_{\tau_t} y_{t-1} + \left(\sum_{k=1}^{K} \theta_{k,\tau_t} - 1 \right) \Delta y_{t-1} + \sum_{k=1}^{K-1} \beta_{k,\tau_t} \Delta^2 y_{t-k} + \left(\lambda_t^{-\frac{1}{2}} \right) \sigma_{\tau_t} e_t \tag{9}$$

(where $\left\{\beta_{k,\tau_t}\right\}$ are derived from linear combinations of $\{\theta_{k,\tau_t}\}$) we are interested in whether the solution (z) to

$$1 - \sum_{k=1}^{K} \theta_{k,\tau} z^k = 0 \tag{10}$$

may have a modulus less than one for one or more regimes. If we find evidence of regimes with solutions inside the unit circle, then we also have non-stationary regimes.

For the sake of clarity, we shall distinguish between a first order ERR where $\rho_\tau > 0$ (or $\rho_\tau < -2$) and where the solutions to [10] lie inside the unit circle and a second order ERR where $\sum_{k=1}^{K} \theta_{k,\tau} > 1$ $\left(\text{or } \sum_{k=1}^{K} \theta_{k,\tau} < -1\right)$. The majority of work dealing with bubbles has focused on first order ERRs and whether series contain regimes where $\rho_\tau > 0$. However, 'second order' explosive behaviour may be even more important in terms of its impact on the overall behaviour of data series.

For the purpose of illustration, suppose that y_t is a logged price with only two regimes. For simplicity assume $\left\{\lambda_t^{-\frac{1}{2}} \sigma_{\tau_t} = 1\right\}$ and $\left\{\beta_{k,\tau_t} = 0\right\}$ and we impose for a specific regime τ_1, $\rho_{\tau_1} = 0$, and $\left(\sum_{k=1}^{K} \theta_{k,\tau_1} - 1\right) = \omega_1 \geq 0$ and $\alpha_{\tau_1} > 0$, then:

$$\Delta y_t = \alpha_{\tau_1} + (\omega_1 + 1) \Delta y_{t-1} + e_t \tag{11}$$

The growth rate Δy_t will contain unit root or explosive regimes. If we then have another 'stationary' regime of the form $-1 < \rho_{\tau_2} < 0$, $-2 < \left(\sum_{k=1}^{K} \theta_{k,\tau_2} - 1\right) = \omega_2 < 0$ and $\alpha_{\tau_1} = 0$, then

$$\Delta y_t = \rho_{\tau_2} y_{t-1} + (\omega_2 + 1) \Delta y_{t-1} + e_t \tag{12}$$

Such a series will transition between two regimes where by in one regime there is rapid growth, with a 'collapse' towards zero which is the mean of the stationary regime. Such behaviour is consistent with a 'boom and bust' scenario that is often associated with bubbles.

Monitoring $\sum_{k=1}^{K} \theta_{k,\tau}$ is also of relevance since if it is at or just below one, the series may have a regime that is I(2) or very close to I(2). If this is accompanied by high error variances, then such regimes may also be extremely volatile.[8] Finally, the existence of stationary regimes among non-stationary regimes may also

[8] It is worth noting that explosive roots need not be represented in a regime switching way. For example, the tests developed by Phillips et al. (2015) do not employ regime switching methods. Moreover, even if a regime switching model is employed, the types of bubbles considered in Brooks and Katsaris (2005) and in Shi and Arora (2012) do not employ Markov Switching. Instead these papers rely on the ability to construct a fundamental price from associated data.

have a dramatic 'implosive' impact on the behaviour of time series. For example, if a series has a random walk regime and a stationary regime, there are likely to be very large periodic corrections back towards the mean of the stationary regime. Consequently, consideration should be given to all regimes, not just those that happen to be 'explosive'. As noted in the Introduction, the 'bubble' metaphor arguably suggests not only can series rapidly inflate, but that the bubble can 'pop'. The representation above illustrates that a 'pop' can be empirically modelled within the regime switching framework.

3.3 Submodels by Restrictions

The equations in [7] nest a number of submodels that can be characterised by combinations of restrictions that we now outline. We will denote a model without any restrictions on the parameters as H0). This model nests a number important special cases as follows:

H1) Variance switching only (no switching of intercepts or autoregressive coefficients)

$$\alpha_{\tau_t} = \alpha,\ \rho_{\tau_t} = \rho \text{ for all } t \text{ and } \theta_{k,\tau_t} = \theta_k \text{ for all } t \text{ and } k$$

H2) No switching of autoregressive coefficients (switching of intercepts and variance possible)

$$\rho_{\tau_t} = \rho \text{ for all } t \text{ and } \theta_{k,\tau_t} = \theta_k \text{ for all } t \text{ and } k$$

H3) No variance switching (switching in the autoregressive coefficients and intercepts possible)

$$\sigma_{\tau_t} = \sigma \text{ for all } t$$

H4) Normal Errors

$$\lambda_t = 1 \text{ for all } t$$

H5) A first order unit root in the autoregressive lags:

$$\rho_{\tau_t} = 0 \text{ for all } t$$

The inclusion of the term λ_t allows for the error term to be a mixture of normals (across variances), whereby there is a formal equivalence between this representation and one where the errors ($\tilde{e}_t = \lambda_t^{-\frac{1}{2}} e_t$) are t-distributed with degrees of freedom that can be estimated.

This set of restrictions implies that there are a large number of potential submodels which form the 'model space' over which we average. For example, for a given lag length, if the restrictions defining H1 are not imposed then there are $2^4 = 16$ possible versions. With H1 imposed there are only four possible models that do not imply a non-regime shifting system (since if the restrictions relating to both H1 and H3 are imposed there are no regime shifts). So for $R = \{1, 2, 3\}$ there are approximately $(2 \times 20 + 4)(K_{\max}+1) = 44(K_{\max}+1)$ possible models. We shall denote these models as $m = 1, ..., M$ where $M = 44(K_{\max} + 1)$. However, we note that although it is theoretically possible to be more general in terms of the number of lags and/or regimes, the list of potential restrictions is clearly not exhaustive. Importantly, the associated increase in computational burden for whatever type of generalisation is considered appropriate is significant and beyond the scope of the current study.[9]

Our approach to estimation and inference for models H0 through to H5 is to estimate all the models for each of the series and rank them using model marginal likelihoods. The parameters that we present will be averaged using the marginal likelihoods once they are converted into 'model weights'. This approach imbues considerable freedom to the estimation process, since the weighted parameter estimates are in effect composed of some models which allow all the parameters to vary simultaneously, and others that hold one or more parameters constant while allowing others to vary. Moreover, by having three regimes, it is not necessary that a change in one parameter is 'paired' with another in time. For example, there could be an explosive regime and a stationary regime with an error variance that is approximately the same over these two regimes, but with a further regime that has a variance that deviates from the other two. Thus, a switch in one parameter is not forced to take place at the same time as the others. Increasing the number of regimes to four would naturally give another level of flexibility[10]. However, this would involve a considerable expansion of the model space, and we would remark that while many of our series are more consistent with three regimes rather than two, limiting the number of regimes to two does not change our general findings. Therefore, we surmise that expanding the analysis by having up to four regimes would not substantively impact on our results (in most cases).

3.4 Prior Distributions

Prior distributions are required with respect to models and parameters. With regard to the former we attach a uniform prior distribution to any given model that is determined by R, K and a given combination of

[9] For example, we could allow for each regime to have different lag lengths. This would indeed represent an extension of the model space, but at quite a computational cost.

[10] A particular point which has been raised (informally) is that, in particular, a seperate regime switching process could operate for the error variance and the autogregressive component. Such a system would be a special case of a four regime model.

restrictions outlined in the previous sections (all models are considered equally likely *a priori*). With respect to the model parameters we outline the general structure of the priors, with the precise specification of the hyperparameters that characterise these distributions (where applicable) consistent with Frühwirth-Schnatter (2004).

For fixed R, K and restrictions (i.e., a given model m), the priors for the model parameters are best discussed by grouping the parameters into blocks. Again, we shall suppress model subscripts in discussing the prior distributions. The following notation is used: $\Theta_\tau = (\rho_\tau, \theta_{1,\tau}, ..., \theta_{K,\tau})$, $\eta_i = \left(\eta_{i,1}, ..., \eta_{i,R}\right)$ where $\eta_{i,j} = P\left(\tau_t = j | \tau_{t-1} = i\right)$, $\phi_\tau = \sigma_\tau^{-2}$; and $\Lambda = (\lambda_1, ..., \lambda_T)$. For all $\tau = 1, ..., R$, the parameters (should they enter a given model and are not constrained to a given value) have the following priors:[11]

$$\begin{aligned}
\Theta_\tau &\sim Normal\left(0, A\right) \text{ for } \tau = 1, ..., R \\
\phi_\tau &\sim Gamma\left(s_0, v_0\right) \text{ for } \tau = 1, ..., R \\
\eta_i &\sim Dirichlet\left(a_{i,1}, ..., a_{i,R}\right) \text{ for } i = 1, ..., R \\
\lambda_t &\sim Gamma\left(\varepsilon, \varepsilon\right) \text{ for } t = 1, 2, ..., T \\
\text{where } \varepsilon &\sim Gamma(\frac{2}{d_0}, 2)
\end{aligned} \tag{13}$$

Note, the prior mean for λ_t is equal to one, and recalling that ε represents the degrees of freedom for the t-distribution for the errors, this means that d_0 represents the prior mean for this parameter.

Next, we adopt the following notation for ease of derivation. Denote the parameter sets $\Theta = \{\Theta_\tau\}_{\tau=1}^R$, $\eta = \{\eta_i\}_{i=1}^R$, $\Lambda = \{\lambda_t\}_{t=1}^T$ and $\phi = \{\phi_\tau\}_{\tau=1}^R$, and since we have $\{\tau_t\}$, which can take the values 1 to R, let us denote the vector S_τ=$(t_{1,\tau}, ..., t_{T_\tau,\tau}) = \{t : \tau_t = \tau, t_{1,\tau} < t_{2,\tau}, ..., < t_{T_\tau,\tau}\}$ (which is the set of indices denoting the time points in a given regime) and T_τ denotes the number of periods within regime τ. We let the collection of parameters be denoted as $\Gamma = (\Theta, \eta, \phi, \Lambda, \varepsilon)$ with $\Gamma_{-\Theta}$ denoting the parameters Γ excluding Θ (and $\Gamma_{-\eta}$ the parameters excluding η and so on) and the collection of the indices denoting regime partitions $S = \{S_\tau\}$. The data up to time t is denoted as D_t.

Defining:

$$x_t = (1, y_{t-1}, \Delta y_{t-1}, ..., \Delta y_{t-K})' \tag{14}$$

[11] The Gamma distribution is parameterised as $Gamma\left(\theta | s, v\right) = \frac{1}{\Gamma\left(\frac{v}{2}\right)\left(\frac{2}{s}\right)^{\frac{v}{2}}} \theta^{\frac{v}{2}-1} e^{-\frac{\theta s}{2}}$

then for each τ we can collect the vectors corresponding to a particular regime as follows:

$$\begin{aligned} X_\tau &= \left(x_{t_{1,\tau}}, ..., x_{t_{T_\tau,\tau}}\right)' \\ Y_\tau &= \left(y_{t_{1,\tau}}, ..., y_{t_{T_\tau,\tau}}\right)' \\ E_\tau &= \left(e_{t_{1,\tau}}, ..., e_{t_{T_\tau,\tau}}\right)' \\ &\quad \text{and} \\ \Lambda_\tau &= diag\left(\lambda_{t_{1,\tau}}, ..., \lambda_{t_{T_\tau,\tau}}\right)' \end{aligned} \tag{15}$$

Additionally, we construct (for a $k \times 1$ vector of ones j_k)

$$\Phi_\tau = diag\left(j_{T_\tau}\phi_\tau\right) \tag{16}$$

Then for a given regime we can construct the linear model,

$$Y_\tau = X_\tau\Theta_\tau + \Lambda_\tau^{-\frac{1}{2}}\Phi_\tau^{-\frac{1}{2}}E \tag{17}$$

and denoting

$$\begin{aligned} \dot{Y}_\tau &= \Lambda_\tau^{\frac{1}{2}}\Phi_\tau^{\frac{1}{2}}Y_\tau \\ &\quad \text{and} \\ \dot{X}_\tau &= \Lambda_\tau^{\frac{1}{2}}\Phi_\tau^{\frac{1}{2}}X_\tau \end{aligned} \tag{18}$$

and given $\Lambda_\tau^{\frac{1}{2}}\Phi_\tau^{\frac{1}{2}}$, the model can be re-expressed as a conventional linear model

$$\dot{Y}_\tau = \dot{X}_\tau\Theta_\tau + E_\tau \tag{19}$$

where

$$\tau = 1, ..., R \tag{20}$$

3.5 Posterior Distributions and Estimation

Since [19] has the form of a standard linear model, then given the independent normal and gamma priors, the conditional posterior distributions for Θ_τ and ϕ_τ are of a standard form that has been derived in numerous places (e.g., Koop, 2003). For more details see Appendix A1.

In estimating the model, the posterior distributions for the above are used to sequentially draw from the conditional posterior using Gibbs sampling, with ε being sampled using a Metropolis-Hastings step. The procedure runs the sampler from an arbitrary starting point for a fixed number before recording samples at each p^{th} iteration, whereby p is set sufficiently large to achieve draws that are not highly dependent. We collect N draws on which to base our posterior estimates. Convergence of the sampler is monitored both visually and by conducting statistical tests that evaluate whether the values drawn at different stages within the sampling process are consistent with being drawn from the same (stable) distribution. Failure of a particular model to converge according to our tests triggered thinning of the sampler and a doubling of the number of draws from the sampler. While such methods are not infallible, the fact that model averaging is being used means that the impact of the results from any one model is considerably reduced, thus limiting the impact of a poorly converged model.[12]

Finally, as is well documented, this sampler can involve 'label switching'. Label switching occurs when a particular regime which has been labelled as τ switches its label to τ^*. Frühwirth-Schnatter (2001) suggested using a permutation at the end of each step in which labels are randomly reassigned. Using this approach there is no particular meaning given to the ordering of the labels[13]. The parameters associated with a particular regime can be recovered by assuming that a particular restriction identifies the model through the use of 'loss functions'. In this paper, we do not pay attention to identifying the regimes in this way. Rather, we concentrate on the time dependent estimates as outlined below.

3.6 Time Dependent Estimates

For a given model m, the estimate of a given parameter (e.g., intercept) at a given point in time can be calculated as the expected value of that parameter by weighting the regime specific parameters by the probability that they are in a given regime. Take a particular specification with the nth draw of parameters $\{\Theta_\tau^n\}$ from the posterior distribution. We recover the estimates of a given set of parameters at a given time point (say $\hat{\Theta}_t$) by taking a given draw of Θ_τ and $P\left(\tau_t = \tau | D_T, \Gamma\right)$,

$$\hat{\Theta}_t^n = \sum_{\tau=1}^{R} \Theta_\tau^n P\left(\tau_t = \tau | D_T, \Gamma\right)^n \tag{21}$$

[12] For readers unfamiliar with Monte Carlo Markov Chains (MCMC) and associated methods, coverage can be found in a number of textbooks, including Koop (2003).

[13] We recognize that strictly speaking that the permuation step is not neccessary within the sampler as argued in Geweke (2007).

This estimate is a 'draw' of the parameter that has been averaged over the regimes. Averaging over these draws gives the (model specific) estimate

$$\hat{\Theta}_t = N^{-1} \sum_{n=1}^{N} \hat{\Theta}_t^n \tag{22}$$

In the empirical section we present the time dependent estimates for each specification for the four key time dependent parameters estimated. Noting that these estimates depend on the specification m, we also give them model subscripts (m)

- The intercept $\hat{\alpha}_{t,m}$
- The deviation from the unit root $\left(\hat{\rho}_{t,m} - 1\right)$
- The 'Second Order Root' $\hat{\delta}_{t.m} = \sum_{k=1}^{K} \hat{\theta}_{t,m}$
- The Error Standard Deviation $\hat{v}_{t,m} = \left(\hat{\lambda}_{t,m}^{-\frac{1}{2}}\right) \hat{\sigma}_{t,m}$

4 Model Selection and Model Averaging

4.1 The Marginal Likelihood

The Marginal Likelihood (MARGLIK) is a Bayesian measure of the adequacy of a model (m) to represent the data where Γ are the model parameters with support Ω_m. It is represented by:

$$f_m\left(D_T\right) = \int_{\Omega_m} f\left(D_T|\Gamma, m\right) f\left(\Gamma\right) d\Gamma \tag{23}$$

As outlined in Frühwirth-Schnatter (2001,2004) some popular methods of estimating the MARGLIK such as the Chib (1995) method and Gelfand and Dey (1994) can perform very poorly. The method of Frühwirth-Schnatter (2001, 2004) is based on the 'bridge sampling' method. It uses a construction of the proposal density (required by the bridge sampling method) that is built on the basis of the parameters sampled from the posterior. This method is outlined in Appendix A2.

4.2 Model Averaged Estimates of the Time Dependent Parameters

For each of the four key parameters of interest, we also present a model averaged version using the MARGLIKs as weights normalised to sum to one over all specifications (implicitly assuming that models are equally likely, *a priori*) to give

$$w_m = \frac{f_m(D_T)}{\sum_{m=1}^{M} f_m(D_T)}$$

The model averaged estimates of the key time dependent parameters are:

- The intercept $\hat{\alpha}_t = \sum_{m=1}^{M} w_m \hat{\alpha}_{t,m}$
- The deviation from the unit root $\left(\hat{\rho}_{t,m} - 1\right) = \sum_{m=1}^{M} w_m \left(\hat{\rho}_{t,m} - 1\right)$
- The 'Second Order Root' $\hat{\delta}_t = \sum_{m=1}^{M} w_m \hat{\delta}_{t,m}$
- The Error Standard Deviation $\hat{\upsilon}_t = \sum_{m=1}^{M} w_m \left(\hat{\lambda}_{t,m}^{-\frac{1}{2}}\right) \hat{\sigma}_{t,m}$

5 Data and Model Priors

5.1 The Data

Within this paper we explore the following data series[14],

1. Argentinian Money Supply (AMS) Monthly data between 1983:1 to 1989:11 (Source: HPS, 1999).
2. Argentinian Consumer Prices (AP) Monthly data between 1983:1 to 1989:11 (Source: HPS, 1999).
3. Case-Shiller US National Quarterly Home Price (USHP) Index between 1953:1 to 2012:2 (Source: Shiller, 2005, updated).
4. Japanese House Prices (JHP) Quarterly, between 1970:1 to 2012:1 (Source: Japanese Real Estate Institute).
5. US Standard & Poors Stock Price to Dividends Ratio (S&P), Monthly logged composite price to dividends ratio between 1871:1 to 2010:12 (Source: Shiller, 2005, updated).
6. NASDAQ Monthly Price to Dividends Ratio (NASD) between Jan 1973 to Feb 2013 (Source: Shi and Song, 2016).
7. IMF monthly index of oil prices (OIL) (average of UK, Brent, Dubai and West Texas intermediate) between 1980:1 to 2012:2 (Source: IMF).
8. Chinese WuLiang Warrant Prices (WWP) Daily Data between 2006:4:3 to 2008:1:09 (Source: Xiong and Yu 2011).
9. BitCoin Prices (BITC) Daily Data between 2006:6:26 to 2014:2:25 (Source: https ://www.quandl.com/c/markets/data).

[14] We also tested the exchange rate series in HPS (1999) but do not include it due to space constraints. However, this series was not found to contain explosive roots.

All analysis of variables is done in terms of logged nominal values.[15]

The AMS and AP were found to have bubbles by HPS(1999), USHP were found to have bubbles by Phillips and Yu (2011) and Homm and Breitung (2012), JHPs have been labelled bubbles in the media and are commonly considered part of a general asset price bubble, the S&P has been found to contain bubbles in Phillips et al. (2015) and Brooks and Katsaris (2005), NASD has been found to have bubbles by Homm and Breitung (2012) and Shi and Song (2016). The WWP was found to have bubbles by Xiong and Yu (2011), and the BITC has been argued to be a pure bubble by among others Quiggin (2013).

Although we do not present the results here, the unit root tests with left hand alternatives as outlined in Harvey et al. (2009) suggest that the series are consistent with being I(2), with the exception of the S&P series and oil series which are I(1). The GSADF tests for all series reject the unit root in favour of right hand alternatives.[16]

5.2 Model Priors and Prior Hyperparameter Values

As noted the model space was given uniform priors (such that any submodel was treated *a priori* as being equally likely to any other). The priors for the intercepts were set to have a zero mean with a standard deviation of one. For all coefficients, we set the prior means to zero with a standard deviation of unity. The prior covariances for the coefficients were set to zero.

The hyper parameters for the Gamma priors were $s_0 = 1$ and $v_0 = 2$, and for the Dirichlet priors we used $a_{i,i} = 2$ and $a_{i,j} = \frac{1}{R-1}$ for $i \neq j$ which are similar to that of Frühwirth-Schnatter (2004). The prior for the degrees of freedom for the t-distribution was set to $d_0 = 25$. This means that our prior weakly supported an approximate normal distribution for the errors. This prior is consistent with that suggested in the literature (e.g., Koop, 2003). These priors are also consistent with our beliefs that the coefficients $\{\theta_{k,\tau}\}$ are less than unity in absolute value, and are also consistent with there being a unit root in the series in 'levels'.

While our priors provided the basis for the model averaging results, we also estimated some key models under alternative priors to see if the results were different. Since our particular focus is on the explosive root properties of the data, a key prior is for the parameter ρ_τ. Our prior for this parameter is normal with a mean of zero (thus giving some prior mass close to the unit root)[17]. However, with a Bayesian approach, the larger the variance attached to this parameter the greater the penalty for the inclusion of y_{t-1}. As the variance

[15] The software used for calculation was Gauss, vs 13. The data can be found in the supplementary materials (available from the authors on request).

[16] These results can be found in the Supplementary materials (available on request from the authors). The tests in Harvey et al. (2009) automate lag length selection and allow for structural breaks and changes in volatility.

[17] As already noted, since all series are broadly consistent with being at least I(1) using unit root tests with left handed alternatives, this prior is quite reasonable in this respect.

shrinks, the more likely that models containing y_{t-1} will be given higher posterior weights. But, on the one hand, we do not want the prior to dominate the data and drive the estimate of ρ_τ towards zero by placing too much mass close to zero. And on the other, we wish to detect small deviations from the unit root (including mildly explosive models) within the model averaging process. Therefore, we estimated all models using two different prior standard deviations for ρ_τ which were set at 0.1 and 0.01. We present results for both priors and discuss the implications for the change in the variance ρ_τ in the results section. However, as we will show for only one variable did variations in this hyperparameter make a substantive difference to the findings.

5.3 Power

The very large model space combined with the need for considerable computation for a given model limits the scope for Monte Carlo exploration of the power of the procedures to detect bubbles. However, in order to verify that the procedures had reasonable power we conducted 5,000 Monte Carlo trials with a 'toy' process $\Delta y_t = \rho_t y_{t-1} + e_t$ with $t = 1, ..., 500$, where ρ_t was Markov switching between the values 1 and $1 + \epsilon$, and where ϵ took the values $0.00, 0.01, 0.025$ and 0.05. The transition probability from one state to another was set at 10%. We then model averaged over models (lag length set to 0) with and without the exclusion of y_{t-1}. Our search was limited across normal errors, though we allowed for regime switching in the intercept as well as 1 or 2 regimes.

At the higher values of $\epsilon = (0.05, 0.025)$ explosive roots were detected in over 99% of trials, regardless of whether the higher or lower prior on the variance of ρ_τ was used. For the case of $\epsilon = 0$, y_{t-1} was included around 14% of the time. However, at $\epsilon = 0.01$, ERRs were detected in around 42% of cases with the higher prior variance, though this rose to around 72% when the smaller variance prior was used. Of course, the actual power of tests to detect Markov switching ERRs will depend on a number of other factors including sample size and the transition probabilities. However, from this we would assess that our procedures will have reasonable power to detect moderately sized ERRs, although when these become very small $\epsilon \leq 0.01$ such explosivity may be difficult to detect.

6 Results

To summarise our findings, we present the Logged Bayes Ratios (LBRs) for comparing model attributes in Table 1.[18] In producing the LBRs for each class within Table 1, there is averaging over all models within a specific class, defined by a particular model attribute. We use the convention that the LBRs represent

[18] The Logged Marginal Likelihoods generating these can be found in the Supplementary materials (available on request from the authors).

the posterior odds in favour of the expression in the numerator relative to the denominator in the fractional expression which is given in the first column. So for example, the first comparison is $\frac{\rho_r=0}{\rho_r\neq 0}$, and, therefore, positive LBRs in the two rows below this mean that the unit root model is supported (in the sense of the 'first order' unit root as discussed earlier). For each comparison of attributes, we give two LBRs for two values of σ_{prior} where this represents the prior standard deviation for the ρ_r parameters. We will focus on Table 1 first, and then discuss the results in further depth using the Figures 3 through 18, generated under the prior which had the highest support in terms of the logged marginal likelihoods. In all cases except the AMS this was for the smaller value of σ_{prior}.

[Position of Table 1]

At the top of Table 1, we report the optimal number of lags (K) supported by the logged marginal likelihoods. As can be seen these did not differ greatly across the two sets of priors. Only in two cases is there a difference, and even then the differences in the logged marginal likelihoods were not large. Note, in terms of deriving these results that for the HPS (1999) series (AMS and AP), we have only included a maximum of four (differenced) lags (that is $K = 4$ in [7]) so as to be consistent with the original research, even though this is monthly data. USHP is quarterly data, so we have only employed a maximum of four (differenced) lags in the regression specification. For OIL, we have a maximum of 12 lags, since it is monthly and we had no prior reason to suspect less lags. For the monthly stocks (S&P) three differenced lags, which is consistent with previous studies. For the daily data (WWP and NASD) we had a maximum of five lags.

Turning to the bottom of Table 1, all nine series support two or three regimes, in most cases three. Thus, all data series seem to have Markov switching in some or all of the parameters. The nature of the regimes is explored below.

6.1 y_{t-1} Exclusion $\{\rho_\tau = 0\}$

We first consider the inclusion or otherwise of the undifferenced lag variable (restriction H5) in Table 1. The restriction that y_{t-1} is excluded imposes the condition that $\{\rho_\tau = 0\}$ for all regimes, thus, prohibiting a 'first order' ERR. As indicated in the prior section, a key hyperparameter is the prior variance for ρ_τ. As this decreases, the prior support for the unit root increases. However, for approximate unit root models the penalty for the inclusion of y_{t-1} decreases, lending prior support to potential ERRs. If the explosive behaviour is very slight, then this prior is more likely to support the inclusion of y_{t-1} and therefore the possibility of finding an explosive regime (as demonstrated in Section 5.3).

Table 1 presents the LBRs likelihoods for the models with and without y_{t-1}, under the two different prior variances. All series with the exception of the AMS, AP, and USHP support the exclusion of y_{t-1} regardless of the prior variance on $\{\rho_\tau\}$. There is only one series for which the support for the exclusion of y_{t-1} changes depending on the prior (USHP). For this series, under the more diffuse prior on $\{\rho_\tau\}$, y_{t-1} is excluded. However, when the prior variance is small, USHP marginally supports the inclusion of y_{t-1}. As we discuss below, the higher or lower prior variance on $\{\rho_\tau\}$ tends to have a small impact on the other model aspects that are supported for each of the variables, except USHP, since the higher support given to the inclusion of y_{t-1} has implications for the selection of other model attributes.

6.2 Error Distribution and Regime Switching Error Variance

In terms of the normal versus t-distribution (restriction H4) for the errors, the majority of the series marginally support the t-distribution relative to the normal and this does not change substantively according to the alternative priors. The BITC series is the exception for which there is strong support for the normal distribution. Considering H3, which concerns regime switching error variances, eight out of nine series (AMS being the exception) have considerably more support for regime switching variances thus not supporting H3. Therefore, regime switching in the error variance is a fairly pervasive finding. We will consider this further when we examine Figures 3 through 18. Considering H1 (whether regime switching variance is the only regime switching required), we need to consider switching in the other components which we do so below.

6.3 Regime Switching in Autoregressive Coefficients and Intercept

Here we make three pairwise comparisons. For example, where all coefficients are allowed to be regime specific versus fixed coefficients is compared in the LBRs under $\frac{\text{Regime All Coeffs}}{\text{Fixed Coeffs}}$. When making this comparison, we are averaging over models with and without potentially time varying values for ρ_t so 'All Coeffs' should be interpreted as all coefficients excluding ρ.

With regard to H1, the results demonstrate there is variability across the series, with four of the series supporting all coefficients being regime specific relative to being fixed and four the opposite. Only, the USHP results depend on the prior, with marginally more support for the full regime switching model when y_{t-1} is excluded.

Collectively, when we consider coefficient and/or intercept switching, under $\sigma_{prior} = 0.01$, seven series, and under $\sigma_{prior} = 0.1$, six series, out of nine series, show support for switching in either the intercept or all of the coefficients in the model (thus not supporting H1), though in a number of cases this is marginal. The notable exceptions are the JHP and WWP for which the regime switching is entirely in the error variance,

thus supporting H1. With regard to H2, there is mixed support, with an even split across series supporting the regime intercept only rather than regime specific autoregressive coefficients.

6.4 Time Dependent Estimates of Key Parameters

In an effort to further clarify the model results presented in Table 1, we now provide figures of the time dependent estimates of the four key parameters. We note that models supporting the exclusion of y_{t-1} will tend to have a very small values of $\hat{\rho}_t$, since the models where $\{\rho_\tau = 0\}$ will get higher weights in the computation of $\hat{\rho}_t$.

Before we discuss each series individually, it is worth noting that the time dependent estimates have a high degree of covariance through time. This is not unexpected given that when switches in regimes take place, a number of parameters may change rather than just one. While this complicates interpretation[19], it remains true that movements in the price series can at some points in time be better explained by one model (e.g., high error volatility but with unit roots) and at other times by another (e.g., low volatility and explosive roots). While it may be difficult to rationalise, we believe that the desire for a simple interpretation should not override the fact that this reflects the complex nature of price behaviour. It is not the case that one model aspect simply compensates for another. Looking across the results, we can find cases where parameters move in the same direction, and others where the same two parameters move in the opposite direction.

Our time dependent estimates for the real series are presented in Figures 3 through 18. For seven out of the nine series (other than AMS and AP), we present results with and without a regime switching variance. Each of the figures contains six plots, the top two plots give the log of the series and differenced series (the series have been scaled so that the maximum and minimum values for the undifferenced series are 1 and 0). The bottom four plots within each figure are of the estimated parameters $(\hat{\alpha}_{t,m},\ \left(\hat{\rho}_{t,m} - 1\right), \hat{\delta}_{t,m}, \hat{\upsilon}_{t,m})$ as defined in Sections 3.2 and 4.2 for the top 10 forming models (the ones with the highest log marginal likelihoods)[20] along with the dark line which represents the model averaged results obtained using $(\hat{\alpha}_t, \left(\hat{\rho}_t - 1\right), \hat{\delta}_t, \hat{\upsilon}_t)$, where averaging has been performed over all models (not just the top 10). The estimates for the top 10 performing models appear as grey lines on the graphs, though it is not always apparent that there are 10 separate estimates because in many instances they are very similar for all the top performing models, and also very similar to the model averaged estimates. A summary of the meaning of each of the six panels is presented in Figure 2.

In addition, and for illustrative purposes, we start by examining a pure Evans' type bubble (this is for hypothetical data generated under the settings of Evans, 1991). These results are presented in Figure 1. We

[19] A point emphasised by a referee.

[20] A plot of the results of all the models results in such large scales being on the vertical access so that the changes in the model averaged parameters become very small. Therefore we use just the top 10.

present Figure 1 in order to gain some perspective on the results which follow. As can be seen in Figure 1, that for this particular bubble, the growth and collapse phases have been captured by all of the coefficients (including in this instance some moderate TER behaviour) and changes in variance. It is important to remember that we are seeing the outcome of a mis-specified model in Figure 1. If there are complex non-linear processes happening then supposed regime switching will occur, but only in a way that approximates this non-linear system. The Evans' bubble that is presented in Figure 1 is one such non-linear process. Additionally, we would not expect to see such extreme behaviour where a process also contained a fundamental component along with the fact that the parameter $\pi = 0.85$ is capable of generating some degree of TER type behaviour, whereas if π approaches one this explosive component will tend to disappear.

6.4.1 Argentinian Money Supply - AMS

Turning to the real data series, taking the figures in order, we first examine the result for AMS presented in Figure 3. Although HPS (1999) analysed three series, we do not consider the exchange rate series as our aim is to show that our methods also pick up the explosiveness found in HPS (1999). Starting with Figure 3, we can observe that the 'deviation from the unit root' in the middle left hand plot for the AMS has grey shaded areas that show that the majority of the high performing models yield estimates of very slightly explosive behaviour ($\hat{\alpha}_{t,m} - 1 > 0$) over the whole range of the series. However, the dominant models and the estimates have two peak ERRs after 1985 and 1989. These coincide with the peaks in Figure 3 in HPS(1999), with the estimates being slightly explosive. This appears to be quite consistent with HPS (1999). However, we note that these explosive phases have been accompanied by a fall in the second order root according to the model estimate along with a fall in the error variance in this phase. It is also worth observing that in many of the lesser performing models (together giving the shaded areas) these two episodes are instead modelled by explosive second order roots and/or increases in the error variance.

6.4.2 Argentinian Consumer Prices - AP

Next consider the AP series presented in Figure 4, for which we obtain some interesting results. The behaviour of this series is very different from that of the AMS series. The AP series appears to exhibit some explosive behaviour after 1989. Of particular interest, however, is the behaviour of the 'second order root'. The AP series appears to have near I(2) behaviour with the series actually having explosive behaviour in first and second order senses in the post 1989 phase. Thus, while the results are consistent with HPS (1999), there is also other interesting behaviour in the propagation mechanisms including explosive behaviour in the second order sense.

6.4.3 All Other Series

For all the remaining data series, we provide a comparison of plots for Figures 5 through 18. The comparison is between pairs of plots that are aligned side-by-side. Specifically, the plots are the same except that the right side plots demonstrate collectively that if there is suppression of the regime switching in error variance, there is a marked increase in the variation in the other parameters of the models. As can be seen from all of these plots, the conclusions about ERRs, in particular, would be very different under constant variances, with both series appearing to display explosive episodes whereas with a regime specific error variance this was not the case. These results echo those of Shi (2013) who demonstrated a tendency for explosive roots to be found when no changes in the variance was permitted.

USHP

The results for USHP are shown in Figures 5 (left) and 6 (right). Recalling the results from Table 1, USHP supported the inclusion of y_{t-1} for the tighter prior. However, we can also see that the deviations from the unit root $\hat{\rho}_t$ are extremely small with the dips below zero being as prominent as the episodes above. What is perhaps just as notable is the behaviour of the sum of differenced coefficients ($\hat{\delta}_t$) which is mostly less than one but with episodes where δ_t is equal to and even greater than unity. There appears to be a clear anomalous phase in the data with the unprecedented rise after 2000 and collapse towards the end of the decade. However, the regime switching model characterises this with a slightly higher intercept along with high error variance in conjunction with movements in $\hat{\delta}_t$ and $\hat{\rho}_t$. Figure 6 gives the results where the error is assumed to be normal with a non-regime specific variance.

JHP

The JHP results are shown in Figures 7 and 8. In some respects they parallel the results for USHP. The periods of rapid expansion and contractions are primarily modelled through changes in error variance, with there being significant propagation of shocks as illustrated through the sum of the difference coefficients (in the bottom left hand panel of each figure). Switching off the error variance has an odd effect as it does not induce an ERR in the sense of the deviation from the unit root, yet there is some quite extreme behaviour induced in the other coefficients, being periodically explosive in the 'second order'.

S&P

The S&P results are presented in Figures 9 and 10. This series exhibits little regime switching in terms of the lag coefficients. Again, for some of the lesser supported models the shaded areas indicate explosive

behaviour at both first and second order levels. However, the model averaged results indicate a stable first order unit root and stable lag coefficients. There is some evidence of regime switching in the intercept term, but the variation is in the error variance, which has a sustained higher error variance during the 1930s and early 1940s. Phillips et al. (2015) and Brooks and Katsaris (2005) found bubbles in these series, but in contrast our approach sees the series movements as predominantly changes in volatility and with some shifts in the intercept.

NASD

The NASD results are presented in Figures 11 and 12. This is the series that typically is used to represent the 'dot.com bubble' given the sharp rise in prices up until 2000 and the subsequent rapid collapse. This series has been found by Shi and Song (2016) to have explosive roots during some of the boom period. Were we to use the criteria of Shi and Song (2016), we would also conclude that there was a bubble in the sense that we have a very slightly explosive root ($\hat{\rho}_t > 0$) over substantial periods of time. However, this is of such a small order (10^{-5}) that it is empirically negligible. Moreover, as we noted from Table 1, y_{t-1} can be excluded from the model (supporting the imposition of a unit root at the first order level). This finding confirms our view that the criteria used by Shi and Song (2016) is particularly generous in its tendency towards finding TERs. In contrast, the regime switching model treats the boom and bust period prior to and after the 'dot.com bubble' as a highly volatile period in terms of the error variance combined with a switch from an upward trend (positive intercept) to a downward trend (a negative intercept). Again, we observe the highly volatile and periodic explosive behaviour of $\hat{\delta}_t$ if the the error variance is held constant.

OIL

Turning to the OIL results in Figures 13 and 14, the plots indicate that there is some switching in the 'second order' roots but with the series not displaying any explosive behaviour, at least according to the model averaged results. We note again, however, that for some of the lesser supported models there is some explosive behaviour after 1985 and 1990, in particular. However, the more highly supported models depict the volatility in the series ostensibly through an estimated increase in the error variance. There is also a very clear collapse in 2008, which is modelled within our framework as (within the full regime switching model) as a negative intercept, combined with a stationary (though near unit root) regime which was similar to three other periods. In some respects this result is similar to Shi and Arora (2012), Phillips and Yu (2011) and Zang and Yao (2016) which see 2008 as a 'collapse' (in the sense of it being the end of a bubble) though in our framework this period is seen as being similar to one just after 1985 and 1990, which is not a strong feature in these other papers.

WWP

The results for WWP are presented in Figures 15 and 16. Here we have estimated the model while ignoring the last 50 days prior to maturity. During this last phase the warrant prices went into a free fall. Using the entire sample alters the estimates substantively and we believe requires an additional regime. Given that we already have three regimes prior to the collapse, it is interesting to see that prices during this phase behave like a sub-martingale process, except that there are notable changes in the error variance and the intercept term. What we do not see is any evidence of a TER. Thus, we would conclude that while there was a clear basis for believing that the price would collapse based on fundamentals, there was little or no evidence for this based on the univariate behaviour of the series. On the other hand it was equally obvious that this series is prone to radically switching regimes with huge swings in volatility in the error. Imposing a constant variance as in Figure 16 again induces very odd 'second order' explosive behaviour in the $\hat{\delta}_t$ coefficients.

BITC

The BITC results presented in Figures 17 and 18 show that there is no real TER behaviour of the first order, with most of the volatility in the series being modelled by switches in the error variance. As with the warrant price data, imposing constant variance has large consequences for the behaviour of the other coefficients $\hat{\delta}_t$ which in absolute terms exceeds unity, though notably imposing a constant variance did not have a large impact on $(\hat{\rho}_t - 1)$. The wild swings in error variances can be shown to imply wide swings in expected returns for this series, which can be calculated to be large and positive at all points in time. There are also periods where the price would in all probability, be expected to fall. On the other hand, we can see nothing in this process at this point that would imply inevitable collapse.

7 Discussion and Conclusions

We explored autoregressive regime switching models for nine economic data series that have previously been argued to contain bubbles, with a view to establishing whether they had a common bubble signature. Our model averaging approach mitigated the tendency for findings to be based on single model specifications that differ substantively from close competitors, thus increasing the robustness of the findings. All series were consistent with regime switching of some sort. Notably, two of the three series explored by HPS (1999) displayed explosive roots even when allowing for regime switching variances. However, we did not find that stock price to dividend ratios (NASD or S&P), house prices (US or Japanese), oil, WuLiang Warrant, or

Bitcoin prices were best represented by models with TERs. We would maintain that there is a strong case these series containing bubbles of a sort, but not the sort that has gained recent attention within the literature. For example, the case made by Xiong and Yu (2011) for Chinese Warrant prices containing bubbles is highly persuasive. Likewise, the case for Bitcoin prices being bubbles is persuasive given that these prices have the propensity to collapse to zero.

While tests based on right hand alternatives to a unit root may detect bubbles (even those that are not strictly explosive in the TER sense), researchers should be mindful that such tests may identify certain types of phenomena as bubbles, even though these phenomena may not be unanimously accepted as being representative of a bubble. This at least includes simple regime switching processes with periods of higher and lower (perhaps negative) growth. Our findings across the range of series suggests that Markov switching models do not find TERs in series that are widely thought to contain bubbles, and researchers should not expect confirmatory results in this respect when comparing Markov regime switching models with the unit root tests under right tailed alternatives. It is possible of course that the TERs exist but that they are not sufficiently explosive to be detected using a Markov switching approach.

Recognising bubbles before they collapse remains a tantalizing possibility. It would open the way for potential market intervention so that a bubble could be deflated before its size threatened to destabilize a market and/or economies should the bubble suddenly collapse. While the idea of bubble detection by identifying a TER phase seems attractive, our results here suggest this is not warranted within a Markov switching framework. However, we do not imagine this is the last word on the matter. One particular path of investigation might be to adopt an asymmetric prior around the 'unit root', which gives greater prior mass to moderately explosive regimes while giving very little to unrealistic explosivity. From an estimation point of view the approach of Shi and Song (2016) which allows any number of regimes seems highly attractive, particularly if this approach can be extended to simultaneously endogenise the choice of lag lengths and other model aspects, not just the number of regimes.

The broadest message in this paper is that any policy framework based on approximating price behaviour by stable stochastic difference equations, with the occasional structural break, amounts to a heuristic that is likely to fail drastically at some point in the future. In all the series that we tested there appear to be quite extreme changes in either volatility, growth rates and/or propagation mechanisms. Of course, the work in this paper has deliberately analysed series that have previously been identified as having bubbles and we should not therefore be surprised to have identified what one might term 'fickle' behaviours. However, an analysis of a much wider set of prices than that presented here has lead us to suggest that many of price series are 'fickle' in their own peculiar way. It is not our proposal that regime switching models are remedial in the sense of offering pointwise

prediction of the direction of prices going forward. Nor, as we have already indicated, do they offer a precise diagnosis of bubbles. Rather, a strategic approach to help policy makers identify that a market or economy is on 'thin ice' would be to monitor data series of interest on an ongoing basis with the objective of looking out for 'odd behaviour' (e.g., TERs or very high error volatility). Monitoring could examine various model specifications simultaneously using regime switching models with BMA. If certain regimes can be characterised as potentially 'unstable', then early recognition that these regimes exist and that current behaviour indicates a high probability of being in such a regime, would strengthen any existing case for intervention, where possible, or 'cheap talk' that may influence uninformed risk averse investors. Moreover, using a BMA methodology, a sudden or rapid switch in the support for one model class to another due to the inclusion of recent data may itself be a clue that something substantive is taking place in the market. The case for real time monitoring of data series behaviour by and for policy makers is persuasive. There is no reason to assume that monitoring of economic series that might be subject to potential bubbles could not be implemented as a form of nowcasting (e.g., Giannone et al., 2008; Castle et al., 2013) if appropriate data series exist.

References

A Very Rational Award. The Economist 19th Oct 2013

Blanchard, O.J. and Watson, M.W. (1982). "Bubbles, Rational Expectations, and Financial Markets." In P.Wachtel (ed.), **Crisis in the Economic and Financial Structure**, 295–315. Lexington: Lexington Books.

Brooks, C. and Katsaris, A. (2005). A Three-Regime Model of Speculative Behaviour: Modelling The Evolution of The S&P 500 Composite Index. **The Economic Journal**, 115 (July): 763–793.

Castle, J.L., Hendry, D.F. and Kitov, O.I. (2013). Forecasting and Nowcasting Macroeconomic Variables: A Methodological Overview, Discussion Paper Number 674, Department of Economics, University of Oxford, Manor Road Building, Manor Road, Oxford OX1 3UQ.

Chen, S-W., Hsu, C-S. and Xie, Z. (2016). Are there periodically collapsing bubbles in the stock markets? New international evidence. **Economic Modelling**, 52: 442-451.

Chib, S. (1995). Marginal Likelihood from the Gibbs Sampler, **Journal of the American Statistical Association**, 96: 270-281.

Chib, S. (1996). Calculating posterior distributions and modal estimates in Markov mixture models. **Journal of Econometrics,** 75: 79-98.

Deschamps, P.J. (2006). A Flexible Prior Distribution for Markov Switching Autoregression with Student-t Errors. **Journal of Econometrics,** 133(1):153-190.

Evans, G.W. (1991). Pitfalls in testing for explosive bubbles in asset prices. **American Economic Review**, 81:922- 930.

Froot, K.,and Obstfeld, M. (1991). Intrinsic Bubbles: The Case of Stock Prices. **American Economic Review** 81: 1189–1214.

Frühwirth-Schnatter, S. (2001). Markov chain Monte Carlo estimation of classical and dynamic switching and mixture Models. **Journal of the American Statistical Association,** 96: 194-209.

Frühwirth-Schnatter, S. (2004). Estimating marginal likelihoods for mixture and Markov switching models using bridge sampling techniques. **The Econometrics Journal,** 7: 143-167.

Garber, P.M. (1990). Famous First Bubbles. **Journal of Economic Perspectives**, 4(2): 35-54.

Gelfand, A. and Dey, D. (1994). Bayesian Model Choice: Asymptotics and Exact Calculations. **Journal of the Royal Statistical Society Series B**, 56: 501-504.

Geweke, J. (1993). Bayesian Treatment of the Independent Student t Model. **Journal of Applied Econometrics**, Supplement. S19-S40.

Geweke, J. (2007). Interpretation and inference in mixture models: Simple MCMC works, **Computational Statistics & Data Analysis**, 51(7): 3529-3550.

Giannone, D., Reichlin, L. and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data, **Journal of Monetary Economics**, 55:665-676.

Gurkaynak, R.S. (2008). Econometric Tests of Asset Price Bubbles: Taking Stock, **Journal of Economic Surveys**, 22(1): 166-186.

Hall, S. G., Psaradakis, Z. and Sola, M. (1999). Detecting Periodically Collapsing Bubbles: A Markov-Switching Unit Root Test, **Journal of Applied Econometrics**, 14: 143-154.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. **Econometrica** 57: 357–84.

Harvey, D. H., Leybourne, S. and Taylor, R. (2009). Testing for Unit root in the Presence of Possible Break in Trend. **Econometric Theory**, 25: 1545–1585.

Homm, U. and Breitung, J. (2012). Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods. **Journal of Financial Econometrics**, 10(1): 198–231.

Koop. G. (2003). **Bayesian Econometrics**. Wiley, Sussex.

Malkiel, B.G. (2012). Bubbles in Asset Prices, in **Oxford Handbook of Capitalism**, Edt by Dennis Mueller, Oxford University Press.

Narayan, P.K., Mishra, S. Sharma, S. and Liu, R. (2013). Determinants of stock price bubbles. **Economic Modelling**, 35: 661-667.

Phillips, P.C.B., Shi, S.P. and Yu, J. (2015). Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500. **International Economic Review**, 56(4): 1043-1078.

Phillips, P.C.B., Wu, Y. and Yu, J. (2011). Explosive Behavior in the 1990s Nasdaq: When Did Exuberance Escalate Asset Values? **International Economic Review,** 52(1): 201–226.

Phillips, P.C.B. and Yu, J. (2011). Dating the timeline of financial bubbless during the subprime crisis. **Quantitative Economics**, 2: 455-491.

Quiggin J. (2013). The Bitcoin Bubble and a Bad Hypothesis.
http://nationalinterest.org/commentary/the-bitcoin-bubble-bad-hypothesis-8353.

Shi, S. and Arora, V. (2012). An application of models of speculative behaviour to oil prices. **Economics Letters**, 115: 469–472.

Shi, S. (2013). Specification sensitivities in the Markov-switching unit root test for bubbles. **Empirical Economics**, 45(2): 697-713.

Shi, S. and Song, Y. (2016). Identifying speculative bubbles with an infinite hidden Markov model. **Journal of Financial Econometrics,** 14(1): 159-184.

Shiller, R.J. (2000, 2005). **Irrational Exuberance**. Princeton University Press

Shiller, R.J. (2003). From Efficient Markets Theory to Behavioral Finance. **Journal of Economic Perspectives**, 17(1): 83-104.

Xiong, W. and Yu, J. (2011). The Chinese Warrants Bubble. **American Economic Review** 101: 2723–2753.

Zhang, Y-J. and Yao, T. (2016). Interpreting the movement of oil prices: Drive by fundamentals or bubbles? **Economic Modelling**, 55: 226-240.

Appendix A1. Posterior Distributions

These conditional posteriors are for each of the regimes ($\tau = 1, ..R$)

$$\begin{aligned} \Theta_\tau|\Gamma_{-\Theta_\tau}, S_\tau &\sim Normal\left(\hat{\Theta}_\tau, \hat{A}_\tau\right) \\ \hat{\Theta}_\tau &= \hat{A}_\tau \dot{X}'_\tau \dot{Y}_\tau \\ \hat{A}_\tau &= \left(A_0^{-1} + \dot{X}'_\tau \dot{X}_\tau\right)^{-1} \end{aligned} \tag{24}$$

$$\phi_\tau|\Gamma_{-\phi_\tau}, S_\tau \sim Gamma\left(E'_\tau E_\tau + s_0, v_0 + T_\tau\right) \tag{25}$$

Should the membership of a particular regime be empty, then the conditional distributions collapse to the prior distributions. Draws from the conditional posterior of the state variables $\{\tau_t\}$ can be done using the approach in Chib (1996). The Chib (1996) method uses the fact that

$$P\left(\tau_t = \tau|D_t, \Gamma\right) \propto P\left(\tau_t = \tau|D_{t-1}, \Gamma\right) \times f\left(y_t|D_{t-1}, \Gamma, \tau\right) \tag{26}$$

where $f\left(y_t|D_{t-1}, \Gamma, \tau\right)$ denotes the likelihood of y_t given D_{t-1} and Γ under regime $\tau_t = \tau$. The likelihood function $f\left(y_t|D_{t-1}, \Gamma, \tau\right)$ is normal under the assumption of $\lambda_t = 1$ for all t. However, by using the fact that $\tilde{e}_t$ are t-distributed (Geweke, 1993), the likelihood can be expressed as a t-distribution rather than as a normal. This means we also know $\tilde{f}\left(y_t|D_{t-1}, \Gamma_{-\Lambda}, \tau\right)$ and it is equally valid to state

$$P\left(\tau_t = \tau|D_t, \Gamma_{-\Lambda}\right) \propto P\left(\tau_t = \tau|D_{t-1}, \Gamma_{-\Lambda}\right) \times f\left(y_t|D_{t-1}, \Gamma_{-\Lambda}, \tau\right) \tag{27}$$

In the case where $\varepsilon \to \infty$ then $f\left(y_t|D_{t-1}, \Gamma, \tau\right)$ and $f\left(y_t|D_{t-1}, \Gamma_{-\Lambda}, \tau\right)$ become equal. The 'update' equation [27] is used in conjunction with

$$P\left(\tau_t = \tau|D_{t-1}, \Gamma_{-\Lambda}\right) = \sum_{r=1}^{R} P\left(\tau_t = \tau|\tau_{t-1} = r\right) \times P\left(\tau_{t-1} = r|D_{t-1}, \Gamma_{-\Lambda}\right) \tag{28}$$

The Chib (1996) method sets $P\left(\tau_1 = \tau|D_0, \eta\right)$ equal to the stationary distribution associated with η such that the posterior distributions for $\{\tau_t\}$ are generated by cycling between [27] and [28].

The posterior of η_i is also Dirichlet (where $s_{i,j}$ is the number of transitions from the state i to state j)

$$\eta_i|S, \Gamma_{-\eta_i} \sim Dirichlet\left(a_{i,1} + s_{i,1}, ..., a_{i,R} + s_{i,R}\right) \tag{29}$$

And finally, the distribution of Λ is derived as:

$$\begin{aligned} \lambda_t|D_T,\varepsilon &\sim Gamma\left(e_t^2+\varepsilon,\varepsilon+1\right) \\ for\ t &= 1,...,T \end{aligned} \tag{30}$$

and (where $f_{Gamma}(.)$ denotes the Gamma function)

$$\begin{aligned} \varepsilon|\Lambda &\propto \left(\frac{\varepsilon}{2}\right)^{\frac{T\varepsilon}{2}} f_{Gamma}\left(\frac{\varepsilon}{2}\right)^{-T}\exp(-w\left(\Lambda\right)\varepsilon) \\ & where \\ w\left(\Lambda\right) &= \frac{1}{d_0}+\frac{1}{2}\sum_{t=1}^{T}\left(\lambda_t-\ln\left(\lambda_t\right)\right) \end{aligned} \tag{31}$$

This last distribution is not of a known form.

Appendix A2. The Calculation of the Marginal Likelihood

The Frühwirth-Schnatter method generates a 'proposal density' $q^*\left(\eta,\Theta,\phi,\varepsilon\right)$ using an iid sample from the posterior $\{\Gamma^n,S^n\}_{n=1}^N$ to calculate:

$$q\left(\Gamma_{-\Lambda}\right)=\frac{\sum_{n=1}^{N} f\left(\eta|S^n\right) f\left(\Theta|\Gamma^n_{-\Theta},S^n,D_T\right) f\left(\phi|\Gamma^n_{-\phi},S^n,D_T\right) f\left(\varepsilon|\Gamma^n_{-\varepsilon},S^n,D_T\right)}{N} \tag{32}$$

The first three terms in the numerator have known integrating constants. However, the last term needs to have the integrating constant calculated for each value but which can be estimated quickly by simulation. The estimate of the MARGLIK is constructed by defining the likelihood function. First, noting that

$$\tilde{f}\left(y_t|D_{t-1},\Gamma_{-\Lambda}\right)=\sum_{\tau=1}^{R}\tilde{f}\left(y_t|D_{t-1},\Gamma_{-\Lambda},\tau\right)p\left(\tau_t=\tau|D_{t-1},\Gamma_{-\Lambda}\right) \tag{33}$$

the likelihood can be calculated by using

$$f\left(D_T|\Gamma_{-\Lambda}\right)=\prod_{t=1}^{T}\tilde{f}\left(y_t|D_{t-1},\Gamma_{-\Lambda}\right) \tag{34}$$

This quantity is calculated during the sampling process for any given set of parameters. The posterior of these parameters is, therefore,

$$f\left(\Gamma_{-\Lambda}|D_T\right)=\frac{f\left(D_T|\Gamma_{-\Lambda}\right)f\left(\Gamma_{-\Lambda}\right)}{f\left(D_T\right)}=\frac{f^*\left(\Gamma_{-\Lambda}\right)}{f\left(D_T\right)} \tag{35}$$

Then for samples let f_0 be a first estimate for $f(D_T)$ and $p_0(\Gamma_{-\Lambda}) = \frac{f^*(\Gamma_{-\Lambda})}{f_0}$ such that the estimate of the MARGLIK is found using the recursion

$$f_j = f_{j-1} \frac{L^{-1} \sum_{l=1}^{L} \frac{p_{t-1}\left(\Gamma^l_{-\Lambda}\right)}{Mq\left(\Gamma^l_{-\Lambda}\right)+Lp_{t-1}}}{M^{-1} \sum_{m=1}^{M} \frac{q\left(\Gamma^m_{-\Lambda}\right)}{Mq\left(\Gamma^m_{-\Lambda}\right)+Lp_{t-1}\left(\Gamma^m_{-\Lambda}\right)}} \tag{36}$$

whereby $f_j \to \hat{f}(D_T)$ (is the estimate of the MARGLIK).

The L-sample $\left\{\Gamma^L\right\}_{l=1}^{L}$ can be obtained by resampling the N-sample $\{\Gamma^n\}_{n=1}^{N}$ and the M-sample is a sub-sample of the posterior, or the whole posterior. Here we set $N = 100$ and $L = M$ with M being equal to the number of draws made by the sampler.[21]

The MARGLIK can be used to discriminate between models. However, it can also be used to discriminate between model properties by averaging over models with a particular property. That is, if for a given model we can write $f_m(D_T)$ then if $\mathcal{M}(H)$ represents all models with property H, then (where # denotes the number of elements in a set)

$$f_{M(H)}(D_T) = \#^{-1}_{\mathcal{M}(H)} \sum_{m \in \mathcal{M}(H)}^{\#_{\mathcal{M}(H)}} f_m(D_T) \tag{37}$$

where equal prior weight has been given to all models. This can be used to compare models with H and those without H $\left(\bar{H}\right)$. Under equal prior odds the posterior odds in favour of H are

$$Odds_H = \frac{f_{\mathcal{M}(H)}(D_T)}{f_{\mathcal{M}\left(\bar{H}\right)}(D_T)} \tag{38}$$

[21] We verified that our procedures give almost identitical results as F-S (2004, page 162) when applied to the GNP data of Hamilton (1989).

Table 1: Model Class Results (Logged Bayes Ratios)

	AMS	AP	USHP	JHP	S&P	NASD	OIL	WWP	BITC
Lag Coefs (K)*									
σ_{prior}= 0.01	0	3	4	4	1	0	2	0	2
σ_{prior}= 0.1	2	3	4	2	2	0	2	0	2
$\frac{\rho=0}{\rho\neq 0}$									
σ_{prior}= 0.01	-3.3	-3.7	-1.0	2.1	3.7	2.4	1.3	0.4	3.2
σ_{prior}= 0.1	-6.4	-0.8	3.5	5.7	5.5	5.1	5.8	2.9	7.5
$\frac{\text{t-dist}}{\text{Normal}}$									
σ_{prior}= 0.01	1.5	0.1	0.1	0.3	2.3	0.2	0.7	0.5	-11.3
σ_{prior}= 0.1	-0.1	0.1	-0.3	0.2	2.2	0.2	0.5	0.6	-9.70
$\frac{\text{Regime Var}}{\text{Constant Var}}$									
σ_{prior}= 0.01	-1.8	4.0	9.8	8.4	38.3	10.6	9.7	26.2	102.5
σ_{prior}= 0.1	-0.4	3.1	14.9	8.2	39.0	11.6	10.4	26.1	102.5
$\frac{\text{Regime All Coeffs}}{\text{Fixed Coeffs}}$									
σ_{prior}= 0.01	2.2	10.4	1.0	-3.7	-1.9	-1.1	2.6	-0.8	1.4
σ_{prior}= 0.1	6.7	7.8	-4.7	-3.6	-1.8	-1.5	2.4	-3.1	1.8
$\frac{\text{Regime Intercept}}{\text{Fixed Coeffs}}$									
σ_{prior}= 0.01	-2.4	0.2	-1.1	-1.1	1.6	0.3	0.7	-0.4	2.0
σ_{prior}= 0.1	-0.8	0.2	-0.9	-1.2	1.4	0.5	1.0	0.4	2.0
$\frac{\text{Regime Intercept}}{\text{Regime All Coefs}}$									
σ_{prior}= 0.01	-4.6	-10.3	-2.1	2.7	3.5	1.5	-1.9	0.4	0.6
σ_{prior}= 0.1	-7.5	-7.6	3.8	2.4	3.2	2.0	-1.4	3.5	0.2
$\frac{\text{2 Regimes}}{\text{1 Regime}}$									
σ_{prior}= 0.01	0.1	6.9	22.7	5.3	41.6	14.1	16.5	23.2	84.8
σ_{prior}= 0.1	5.4	4.0	22.8	5.6	41.9	14.2	16.6	23.1	85.1
$\frac{\text{3 Regimes}}{\text{1 Regime}}$									
σ_{prior}= 0.01	1.1	14.0	32.8	7.8	45.5	12.9	24.9	25.8	105.6
σ_{prior}= 0.1	5.1	11.6	31.7	7.8	45.8	13.0	24.8	25.8	105.8
$\frac{\text{3 Regimes}}{\text{2 Regime2}}$									
σ_{prior}= 0.01	0.9	7.1	10.1	2.5	3.9	-1.2	8.4	2.6	20.8
σ_{prior}= 0.1	-0.3	7.6	8.9	2.2	3.9	-1.2	8.2	2.7	20.6

*Indicates the most supported lags, lags not fixed. Each model aspect is evaluated by averaging over the others

Fig 1. A Pure Evans Bubble (Full Regime Switching)

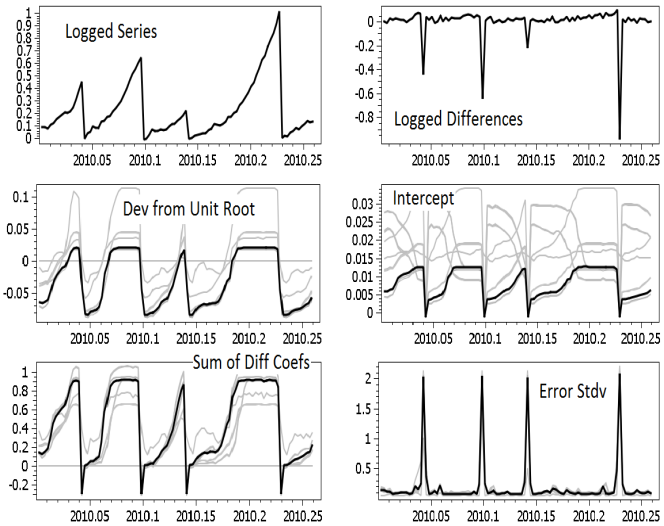

Fig 2. Definition of Panel Components

"Logged Series" $\ln y_t$	"Logged Differences" $\Delta \ln y_t$
"Dev from Unit Root" $(\hat{\rho}_t - 1)\&(\hat{\rho}_{t,m} - 1)$	"Intercept" $\hat{\alpha}_t \& \hat{\alpha}_{t,m}$
"Sum of Diff Coeffs" $\hat{\delta}_t \& \hat{\delta}_{t,m}$	Error Stdv $\hat{v}_t \& \hat{v}_{t,m}$

$\{m\}$ are for the top 10 performing models

Fig 3. AMS (Full Regime Switching)

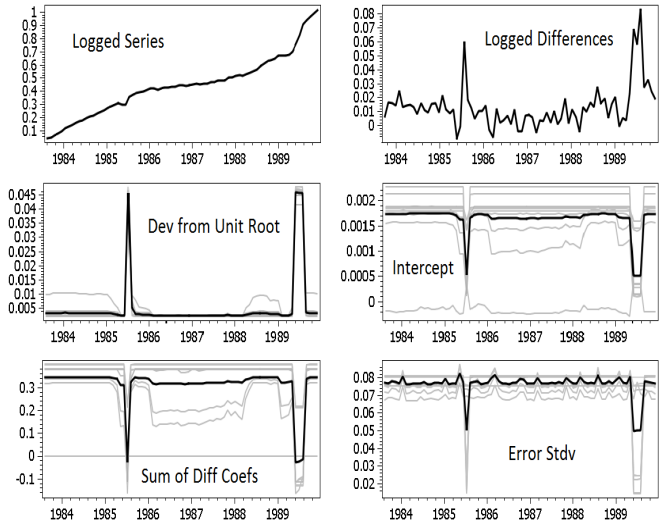

Fig 4. AP (Full Regime Switching)

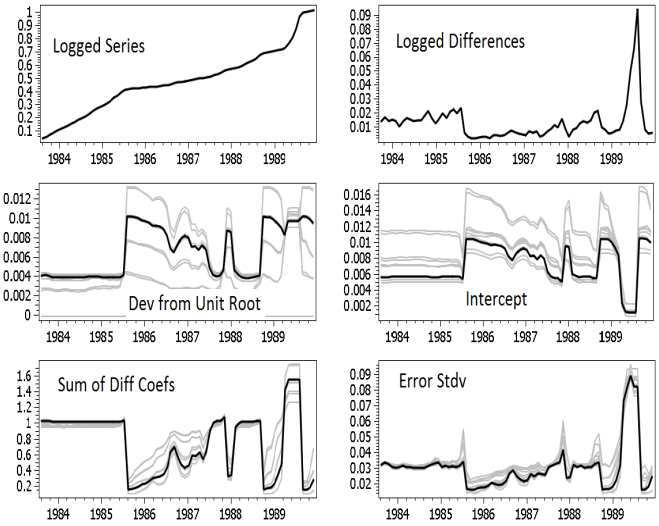

Fig 5. USHP (Full Regime Switching)

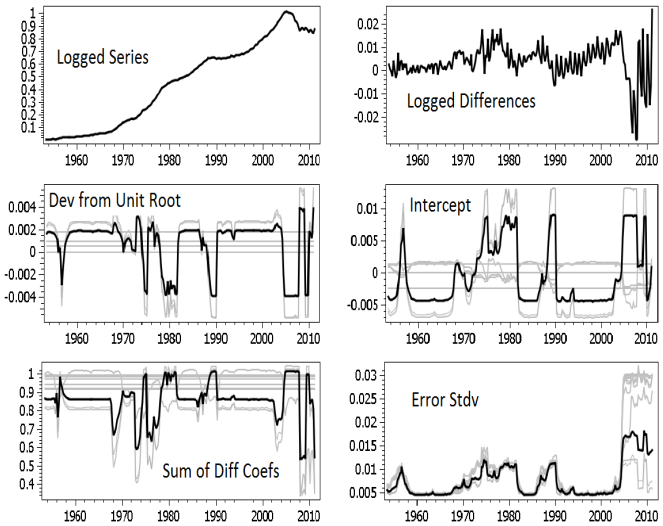

Fig 6. USHP (Constant Variance)

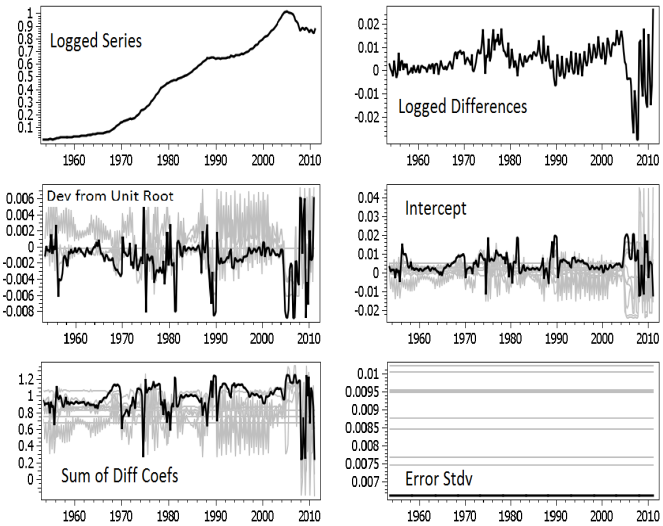

Fig 7. JHP (Full Regime Switching)

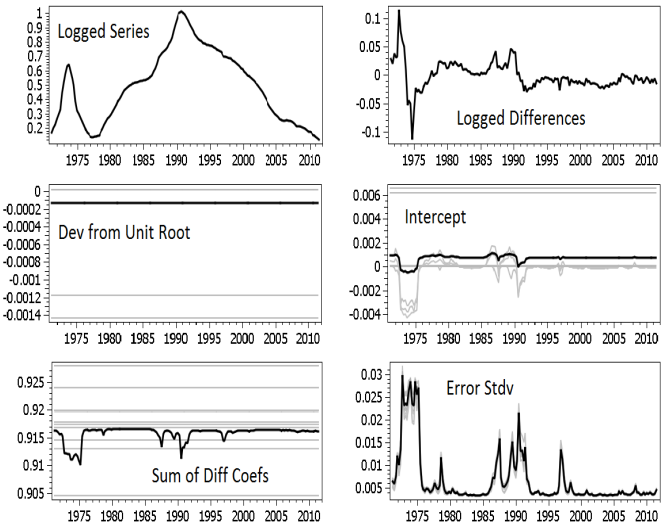

Fig 8. JHP (Constant Variance)

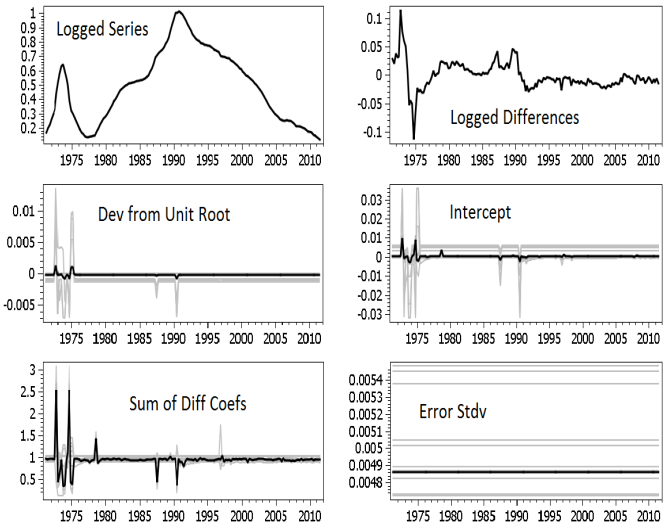

Fig 9. S&P (Full Regime Switching)

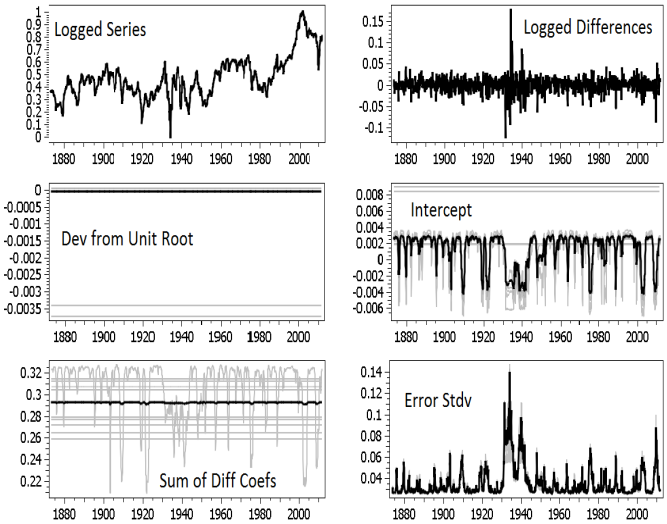

Fig 10. S&P (Constant Variance)

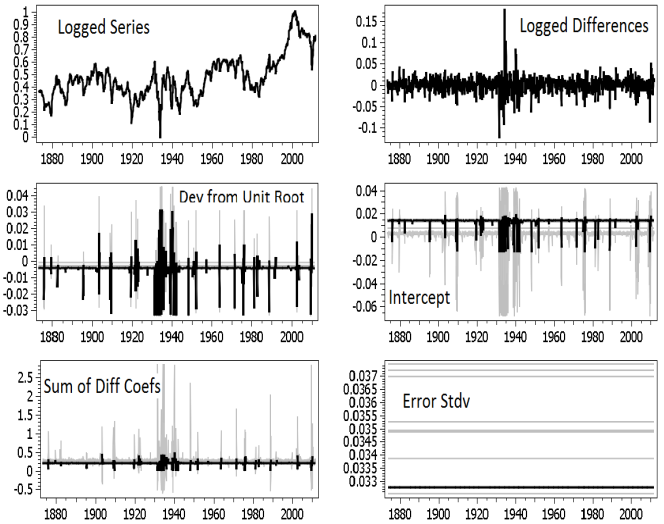

Fig 11. NASD (Full Regime Switching)

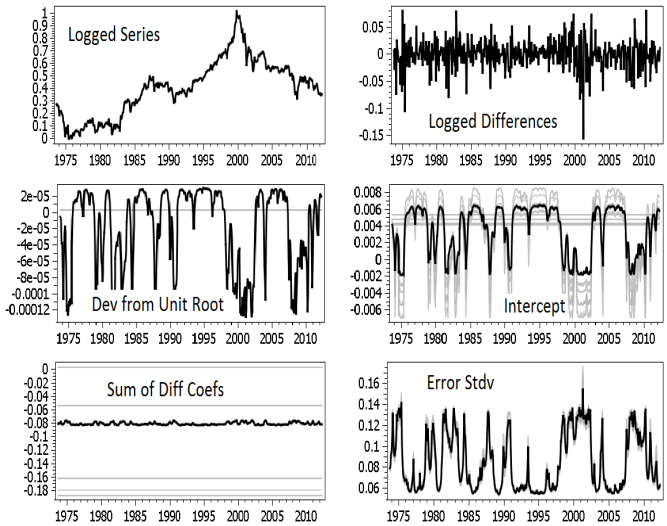

Fig 12. NASD (Constant Variance)

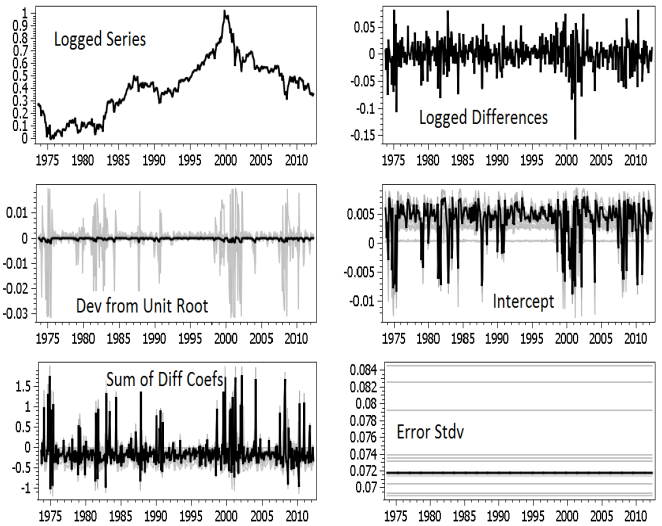

Fig 13. OIL (Full Regime Switching)

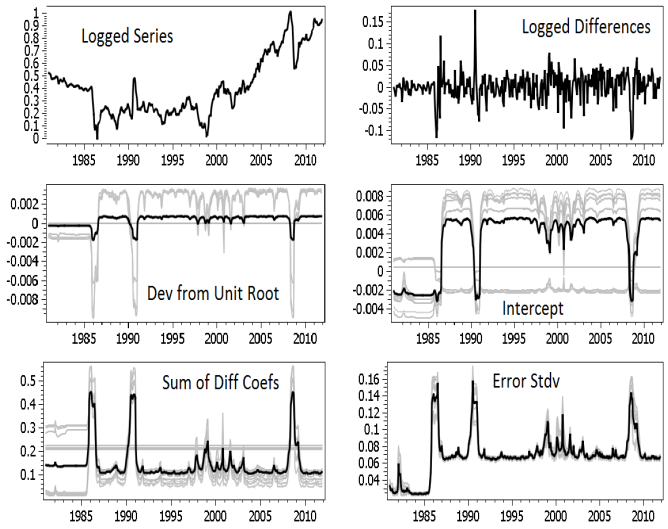

Fig 14. OIL (Constant Variance)

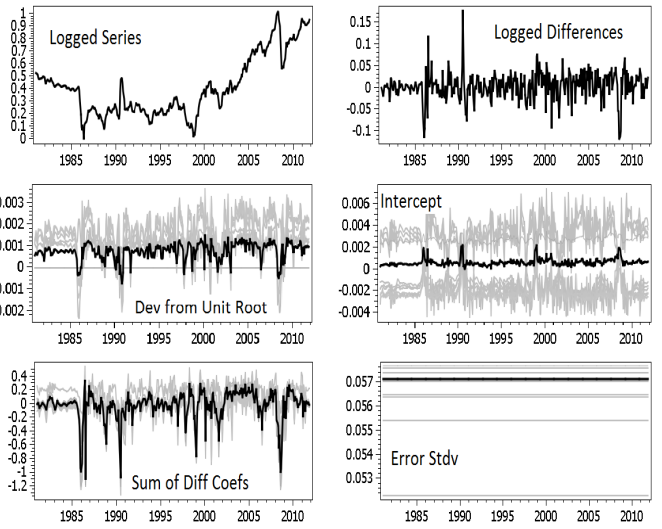

Fig 15. WWP (Full Regime Switching)

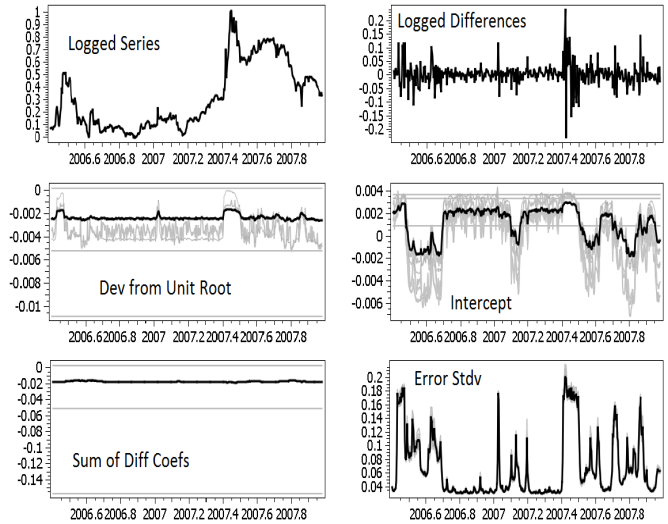

Fig 16. WWP (Constant Variance)

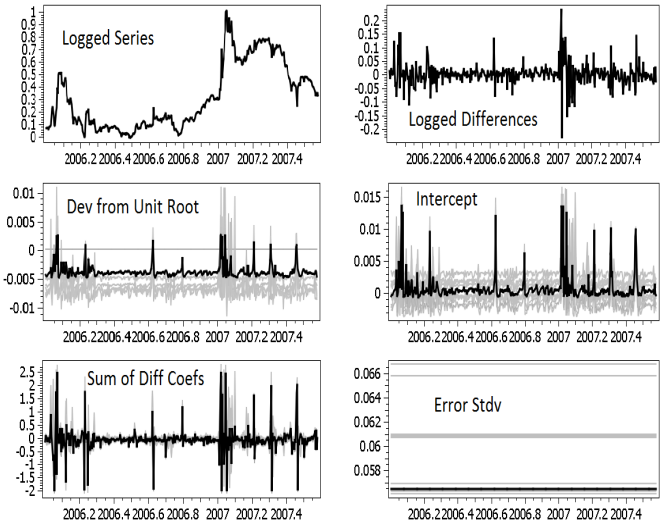

Fig 17. BITC (Full Regime Switching)

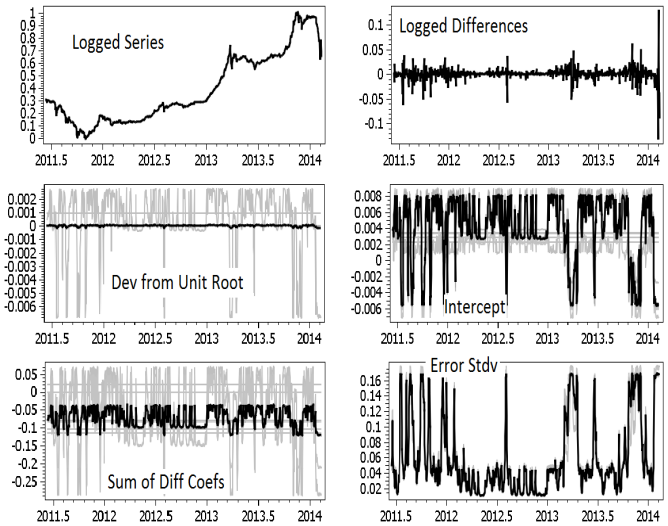
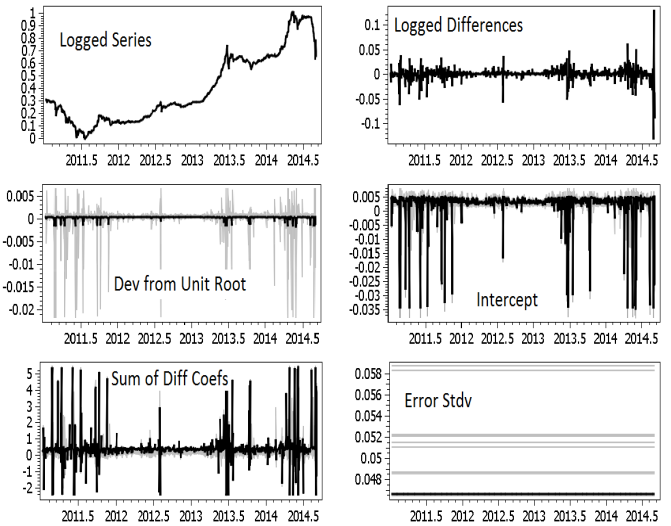
Fig 18. BITC (Constant Variance)