Temporal Issues of Market Inefficiency in asset prices with an emphasis on commodities

By: Muhammad Farid Ahmed

Selwyn College

Faculty of Economics

Submitted for Examination in February, 2017

Declaration: This dissertation is submitted for the degree of Doctor of Philosophy in Economics

Temporal Issues of Market Inefficiency in asset prices with an emphasis on commodities

By: Muhammad Farid Ahmed

Summary:

This summary provides an overview of the contributions made in this thesis to the literature. No references are included in the summary; these can be found in the Bibliography on page 156. This dissertation consists of 6 chapters. The first chapter acts as an introduction to the thesis and discusses the central theme of the dissertation along with providing a preview of what to expect in the following chapters. The contributions of the different chapters vary. Chapter 2 is a more introductory chapter and its contributions are perhaps less consequential than those in Chapters 3-5.

Chapter 2 makes contributions to the literature on testing for explosive roots or bubbles. By modifying the Bhargava test statistic, we show in Chapter 2 that the Bhargava test can address earlier criticisms that had been cited against it; namely that it has low power when multiple bubbles are present in a particular series. Through introducing a rolling window approach, we are able to address that criticism and show that the modified Bhargava test statistic achieves better power. We compare and contrast the power of the modified test with the popular GSADF test statistic which has recently become popular in bubble testing literature. Another contribution made in this chapter is the application of these tests to a data set comprising of 25 commodities. As at the writing of the chapter this was believed to be a first attempt to perform bubble testing on a comprehensive commodity data set. Since commodities are often deemed to be targets of speculative behaviour, they are a natural universe for testing notions of market efficiency as they tend to go through different regimes through natural economic processes. Using both tests we are able to detect bubbles in similar periods with most of them being concentrated around the two oil price crises (1972-73 and 1979-80) and the financial crisis (2005-2007). Our conclusion is that the modified Bhargava statistic works better than the original statistic and can be used to complement the results of other statistics.

The major contributions of Chapter 3 and 4 are the introduction of different methodologies that enable the user to assess how often asset markets are efficient. In Chapter 3 we argue that commodity prices can be estimated using switching-regression models including hidden

Markov state-switching models. Instead of estimating Markov transition matrices directly from the estimation procedure, we estimate the transition matrix separately using unit root tests. By restricting the transition matrix to our estimated matrix and then estimating a Markov state-switching regression we show that we get more accurate smoothed probabilities i.e. a high probability is assigned to explosive states when the price was actually explosive and a high probability is assigned to the random walk/efficient state when the price exhibited efficient behaviour. This methodology is then extended to the three state case and it is argued that the transition matrices estimated this way will inform us of how often commodity markets are efficient. The methodology is empirically applied to non-ferrous metals with particular attention to Copper; we believe this is an additional contribution of the article. Chapter 3 also presents a partial equilibrium model which leads to an estimable reduced form expression for commodities and thereby motivates estimation by Markov switching-regressions.

Three major contributions are made in Chapter 4. Firstly, we make a theoretical contribution to the literature on threshold auto-regressive models with exogenous triggers. Conditions for the existence of a mean and variance when a series follows a threshold auto-regressive (TAR) process with an exogenous trigger are derived. The second contribution is the use of TAR simulations to show that the tests which try to detect bubbles in asset prices lose a substantial amount of power when the asset price spends some time in the mean reverting state in addition to being in the explosive and random walk states. The third contribution of this article is the provision of a framework using TAR models which acts as a metric for market efficiency. By considering three states, an efficient/random walk state, a mean reverting state and an explosive state, we show that estimating asset prices as TARs with exogenous triggers can allow us to measure how often an asset market is efficient. This methodology uses a different class of models from those used in Chapter 4. The methodology is then applied to the S&P500 and FTSE100 process and it is shown that under the most general model specification, the indices primarily exhibit market efficiency.

Chapter 5 looks deeper into how commodity prices are determined and thereby the main contribution is to the literature on commodity market pricing. By making three important changes to the commodity storage model of William and Wright (1991), we are able to show that our model is able to capture essential features of commodity prices that have not been captured by previous iterations. The numerical solution for the model is obtained using the Parameterized Expectations Algorithm (PEA) and simulated series based on this solution are

able to reproduce some statistical features of real commodity price series including a high degree of first order auto-correlation, skewness and kurtosis. A second contribution is with regards to the application of the model; we calibrate the model to match five real commodities and show that the model's solution is able to match real life data. The model is also able to explain why we observe spikes (bubbles) in commodity prices and cites the impact of storage as a probable contributor. Chapter 6 provides concluding remarks on the dissertation.

Table of Contents

Chapter 4: What proportion of time is a particular market inefficient? A method for analysing the frequency of market efficiency when equity prices follow Threshold Auto-regressions

PREFACE:

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in this Preface and specified in the text. Chapter 4 of this dissertation was written in collaboration with my supervisor, Professor Satchell.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma, or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

Acknowledgements

I am grateful foremost to God for answering my prayers and for getting me to where I am today.

I am very thankful to my Supervisor, Professor Satchell. I could not have asked for a better mentor. It is through his support and guidance that I have made it this far and have learnt a great deal from him along the way. In particular, I would like to thank him for taking me on for supervision 6 years after I first applied to Cambridge for a PhD.

This dissertation and indeed any academic endeavour will not have been possible without the foresight and patience of my parents; my mother, Nahid and my father, Awais. They had the vision to send me to the best schools and invested a substantial amount of financial and emotional capital to first, send me to the best University in Pakistan and then to Cambridge. It is through their constant support and patience that I was able to leave a job in the city and concentrate fully on what I love to do. A special thanks to my siblings, Maria, Faria and AB who have always been there, through ups and downs and have made the downs bearable.

My sponsors, the Higher Education Commission of Pakistan, The Cambridge Commonwealth Trust and the Trustees of the Tudor and Wrenbury Scholarship have my gratitude. It would not have been possible to pursue this degree without their financial support. A special thanks to the Charles Wallace Trust for providing funding towards the end of my PhD.

I would also wish to acknowledge the helpful comments and support offered by my peers and the faculty of Economics at Cambridge. In particular, I would like to thank Dr. Cavlacanti and Dr. Geraats for their invaluable comments on my workshop presentations. Among my peers, I would like to thank Simon, David, River, Peng, Jeroen, Axel, Hassan, Jasmine, Margit, Sam, Katja, Cherry and Anil for all the good times in the DAE room and the faculty.

Last but not the least, I thank all my friends, spread all over the world, who have helped me during this time, who have made me believe in my ability when I have doubted myself and have shown me that it is occasionally okay to take a break (a short one) from research. If I started naming all my friends individually it would cover the length of this PhD so I'll thank the six who have stood by me throughout my time as a university student – so thank you Askari, Furqan, Sana, Sarah, Yumna and Zara.

CHAPTER 1: Introduction

Market efficiency is arguably the most talked about and divisive topic in both Economics and Finance. Since Fama first introduced the notion of market efficiency in 1965, experts in both academia and the professional world have invested a considerable amount of time either emphasizing its merits or discrediting it. Many Nobel laureates have been crowned for their contributions towards the efficient markets literature; the most recent among them being Fama and Shiller who offer sharply contrasting views on market efficiency. Sewell (2011) provides a chronological review of work carried out on the efficient market hypothesis.

This dissertation broadly falls within the market efficiency literature. The contributions made have direct implications for market efficiency and how it is generally talked about in the literature. The research conducted can neatly fit within 3 wider streams – time series econometrics, finance and macroeconomics. It is perhaps most appropriate to consider the location of this thesis at the intersection of the above mentioned subjects. The main contributions to the efficient markets literature in this dissertation are contained in Chapter 3 and 4 of this manuscript. Chapter 2 discusses efficient markets only briefly while Chapter 5 is an attempt to explain price movements that are typically associated with market inefficiency.

When I started my research, my main interest was to develop techniques that could assist in detecting 'bubble-like' or explosive behaviour in commodity markets, how that may impact income distribution within an economy and how that in turn could have macroeconomic consequences. However, it was clear within a few months that the 'bubble' literature in and off itself was too vast for one PhD. This explains why, even though the research can generally be classified within the market efficiency literature, a large amount of space is dedicated towards explosive, bubble-like regimes and the empirical applications often focus on commodities. At the outset, it is paramount that we explain what we mean by market efficiency and explosiveness as this is critical for understanding most of the research that follows.

The underlying notion of market efficiency is that the current market price of an asset reflects all publicly available information; thus, no additional returns can be made by relying on information that exists in the public domain (semi-strong market efficiency). Roberts (1967) introduced the idea of strong and weak form market efficiency. If a market is weakly efficient, no additional returns can be made by investors through trading on historical information as that information will already be reflected in the prevailing price. Semi-strong form market efficiency on the other hand precludes the possibility of additional gains based on all publicly available information such as press releases or recently released financial information. Strong-form market efficiency goes a step further; if a market is strong form efficient, investors cannot make excess returns even if they possess insider information.

As stated earlier, economists have dedicated their lives to either proving or disproving each form of market efficiency through a variety of methods including parametric and non-parametric tests.

This thesis primarily considers weak form market efficiency. A weakly efficient market can be represented as a random walk with drift (a Martingale is usually sufficient to define efficiency; random walk is a stronger assumption). Thus, if p_t represents the log price of a financial asset, it may be written as:

$$p_t = \alpha + p_{t-1} + \eta_t \qquad\qquad (\mathbf{1.1})$$

where α is a drift term and η_t represents an independently and identically distributed disturbance term (not necessarily normally distributed). Under these assumptions, it is clear that the geometric rate of return $(p_t - p_{t-1})$ is independent of all past price movements. Throughout this dissertation, any reference to market efficiency is to be understood with reference to equation (1.1).

Bubbles on the other hand are much harder to define. When an asset price experiences a bubble-like period, it is said to be deviating from its fundamental value (mostly increasing rapidly) as understood by Blanchard et al (1979). While understanding bubbles in this way is somewhat simpler, it raises an important question i.e. what is the fundamental value of an asset. Economists who have contributed to this literature have gone back and forth over the definition of fundamental prices; an Economist comes with a test for bubble detection and argues that it works well with a certain class of assets; a second Economist considers the test invalid by arguing that the definition of the fundamental value is incomplete and needs to be richer. As we can see, this is an endless and somewhat circular argument. Instead of delving into this debate, we classify bubbles or explosiveness (bubbles and explosiveness are used interchangeably) with reference to equation (1.1). Consider equation (1.2) below:

$$p_t = \alpha + \phi p_{t-1} + \eta_t \qquad\qquad (\mathbf{1.2})$$

An asset is said to exhibit bubble-like or explosive behaviour if $\phi > 1$ (statistically significantly greater than 1) for a period of time. Thus, if we conduct a right-sided unit root test and reject the null hypothesis, we would conclude the presence of a bubble or an explosive root and p_t would be non-stationary. Any mention of explosiveness made in this dissertation is to be understood as it has been defined above. From a theoretical viewpoint we contend that a richer definition of fundamentals than the one above will be able to explain explosiveness. However, this does not diminish the need for tests to identify when such periods may occur as they cause a number of distortions in the market and may highlight the need for structural reform. We only need to look back to the 2007-2009 financial crisis to understand the impact of bubbles in financial markets.

If a market is efficient as in (1.1) then it cannot exhibit bubble-like behaviour like (1.2); at least not long enough for the behaviour to be detected by a unit root test. Market efficiency has been widely

understood in a binary sense; a market is either efficient or inefficient. Econometric tests similarly treat market efficiency as an 'either or' question. We abstain from a binary classification of market efficiency and instead understand assets as going through various 'states' or 'regimes'; market efficiency is one such state or regime. In this setting we can observe states that are non-efficient; they may be mean reverting (in which case ϕ in (1.2) will be less than 1) or explosive. This classification immediately raises a number of questions – how can we objectively identify these states? Can a non-binary understanding of market efficiency help us in better estimating asset prices? What are the implications for tests of market efficiency or bubble detection? Can theoretical models help in explaining reduced-form state switching behaviour? The main contribution of this thesis, in addition to the introduction of a regime based definition of market efficiency, is to provide procedures to address some of these questions as well as some tentative answers.

In Chapter 2, we extensively discuss various methodologies for detecting bubble behaviour including one that has been developed recently i.e. the Philips, Shi and Yu (PSY hereafter) or the Generalized Supremum Augmented Dickey Fuller (GSADF hereafter) test (2013). We introduce a new test for detecting explosive states and argue that this new test is as good at identifying bubble behaviour as the GSADF test; specifically, the test we introduce helps in identification of periods in which the asset exhibits explosiveness. Chapter 3 builds on the results of Chapter 2; Chapter 3 begins with the introduction of a partial equilibrium model which justifies the use of a state-switching auto-regressive function as the reduced form of an asset price under rational expectations. The state-switching auto-regressive function is then estimated using a Markov state switching algorithm. However, instead of using a simple algorithm, we reinforce the algorithm and make it more accurate by estimating transition probabilities using the GSADF test.

Chapter 4 formally introduces the notion of market efficiency as being one of a finite number of states in a state-switching model. This Chapter was jointly written with my supervisor, Professor Satchell. In this chapter we derive formulae for the first two moments of an asset that may have a state-switching form; we outline the conditions that need to be met for the existence of these moments. We go on to show that only when these conditions are not satisfied is the power of bubble detection tests, such as the GSADF test, high. The most important contribution of this Chapter and perhaps this dissertation is the introduction of a methodology that helps us identify how often a market remains efficient. This allows us to not only say whether a particular asset market is mostly efficient or inefficient but also enables the identification of periods when inefficient states occur.

Chapter 5 takes the ideas of chapter 4 and applies them to the commodity market by building upon William and Wright's (1991) model of commodity storage. Specifically, the augmented model is able to capture features of real commodity data much more accurately than previous attempts. The augmented model also helps in explaining why and when we may observe different states in

commodities. The driving force behind state-switching behaviour at least from the supply side is storage. Each chapter contains a main contribution in addition to one or more minor contributions. Below, we briefly highlight the contributions made in each chapter.

A foray into an area such as bubbles requires a thorough understanding of the literature in order to appreciate the significant contributions that have been made and to identify areas where further research could be carried out. Chapter 2, entitled '*Detecting Multiple Collapsing Bubbles: A modification to the Bhargava statistic',* begins with an in-depth look into bubble detection literature. Building upon an earlier literature review carried out by Camerer in 1989, the Chapter provides a more contemporary update. While state-switching is mentioned only with regards to bubble testing, the chapter does lay the foundation for Chapters 3 and 4 which deal more explicitly with state-switching in assets and their implications for market efficiency.

The two main camps in the market efficiency debate also define the notion of 'bubbles' differently. On one side, we have Blanchard et al (1979, 1982) and Fama et al (1988) who contend that bubbles are a rational aspect of the market. As per their argument, asset prices contain two components, a fundamental component (think dividends for a stock market share) and a bubble component which collapses periodically with a given probability π, in each period. If the bubble component does not collapse in a particular period it grows at the rate $(1 + r)$, where 'r' is the nominal interest rate. Investors continue to price the asset with the bubble component as neglecting the bubble component would be considered irrational. Thus, the bubble component exists but investors can distinguish between the bubble and the fundamental and understand that the bubble component may collapse in any period.

Opposing the rational bubble theory is the Irrational or fad theory of bubbles chiefly attributed to Shiller (1981, 1984, 2005). According to Shiller, asset prices contain an irrational component (instead of a rational bubble component) which he refers to as a fad. The fad is not determined by market fundamentals as it is not known to arbitrageurs or investors. It arises due to speculation and intrinsic evaluation of investors and may thus be attributed to 'animal spirits.' The rate at which the 'fad' component grows depends on agent behaviour and the institutional structure of the asset market.

A minor contribution of this chapter is an updated analysis of bubble detection tests. As at the writing of the chapter, parametric tests of bubble detection could neatly be characterized into 3 categories. Most parametric tests try to model bubbles as rational and the tests are carried out on the bubble component. The earliest bubble detection tests were based on volatility. The underlying principle of volatility based tests is that if a bubble is not present in an asset, then the volatility of the asset price cannot exceed the volatility of its fundamentals. These were initially constructed as tests of market efficiency and thus became relevant to this literature.

A related category of bubble detection tests relied on testing the stationarity properties of asset prices. By testing for co-integration, these tests argued that if fundamentals could be made stationary through differencing but not the underlying asset price, it would indicate the presence of a bubble component which caused non-stationarity. The first two categories of bubble detection or explosive root tests were criticized by Evans (1991) who proved that if there were multiple bubbles that kept collapsing, it would be very difficult to statistically detect bubbles.

The third category of bubble detection tests relies on switching-regression analysis and explosive root tests. While explosive root tests fail to answer Evans' criticism, switching-regressions tend to do relatively better. It is this switching-regression framework that we build upon to provide a metric for market efficiency in Chapter 4. Most recently, the statistic provided by Philips, Shi and Yu (2013) also referred to as the GSADF statistic claims to overcome Evans' criticism. The GSADF statistic is an algorithm which uses two rolling and expanding sub-samples to calculate right-sided Dickey Fuller statistics. The methodology is able to detect bubbles as well as time stamp periods where these bubbles occur. The statistic proves extremely useful for the Markov-state estimation strategy that is introduced in Chapter 3.

Our main contribution in this chapter is the introduction of an alternative statistic, one that significantly modifies Alok Bhargava's (1986) right-sided unit root test. The Bhargava statistic has been proven to be more powerful than the Dickey Fuller statistic for the detection of right-sided unit roots for small samples (less than 200). However, most asset price series are long and as noted above, bubble detection becomes difficult when there are multiple collapsing bubbles in a series; a modification to the procedure became necessary. We provide a rolling window algorithm that requires the calculation of the Bhargava statistic on each window and argue that this algorithm is as useful for detecting bubbles as the GSADF tests.

In order to show that our test has merit, we numerically calculate the power of the modified Bhargava test; our numerical results show that the new test performs just as well as the GSADF statistic. We go a step further and empirically conduct the GSADF as well as the modified Bhargava test on a set of 25 commodities. The empirical application of bubble detection tests on an extensive commodity data set is also a significant contribution of this chapter as no in-depth study of bubbles in commodities had been carried out when the chapter was written. The modified Bhargava test also performs well empirically and is able to detect bubbles in similar periods as the GSADF test. The oil price crises in the 70's and the financial crisis of 2007-2009 are periods where most bubble episodes are seen.

While our test does not date stamp bubbles as precisely as the GSADF test, it does enable the identification of periods where the commodity price may not have been efficient; thus, it is a fairly useful test of market efficiency. A major benefit of using our test is that it can be applied to small samples; in comparison, the GSADF test loses a lot of power over small samples. We believe that it

could be particularly useful in detecting explosive roots in high-frequency data. Thus, Chapter 2 lays the foundation for a state based understanding of market efficiency arguing that explosiveness is one possible state for an asset price which may otherwise be considered efficient.

As mentioned above, Markov-state switching has been used to identify explosiveness in markets and has therefore been adopted as a tool for detecting market efficiency. As the title suggests, Chapter 3: '*Understanding commodity markets using Markovian state-switching,*' uses this technique to identify episodes of explosiveness. However, the main contribution of this chapter is methodological. We introduce an estimation strategy that uses Markov-state switching and allows us to assign probabilities to efficient and non-efficient states.

The chapter begins with the introduction of a partial equilibrium commodity storage model. While simple, the model uses Muth's (1961) commodity storage setting to derive conditions under which a non-linear auto-regressive process can be used as a reduced form expression for commodity dynamics. We also derive conditions that need to be satisfied by model parameters for the commodity price to exhibit explosive as well as mean reverting behaviour. The critical parameter is the speed with which investors change their inventory holdings. This approach is different from the one adopted in Chapter 5 as we are able to derive all results analytically while in Chapter 5, the model adds several layers of complexity and requires an approximate numerical solution.

Despite being simple, the model nevertheless provides a motivation for estimating storage based commodities through state-switching techniques. The rest of chapter 3 is dedicated towards explaining the Markov state-switching technique; a maximum likelihood technique introduced by Kim et al (2012) and adopted by Perlin (2014) forms the basis of our numerical solution. The original technique allows the user to measure not just state specific parameters but also a transition matrix which contains probabilities of switching from one state to other states every period. We noted that a direct application of the algorithm on commodity price data estimated high probabilities for explosive and mean reverting states during time periods when prices were stable. The estimated parameters were often found to be in inadmissible ranges.

We improve the estimation technique and reduce the numerical burden on the algorithm by first estimating the transition probabilities before finding the parameters through the likelihood approach. The transition probabilities are estimated through the application of the GSADF test mentioned above. Since the GSADF test time stamps periods during which a commodity is in an explosive state, we can find the transition probability by using the time stamps. In addition to an explosive state we also included a mean reverting state. The computed transition probabilities are used as weights in the likelihood function; the estimated parameters were found to be in the correct range and we also noted that high probabilities for explosive states were found during the oil and financial crises and thus

agreed with the results we found in Chapter 2. Details of the numerical algorithm as well as the estimation of the transition matrix have been reserved for the Chapter.

An in-depth empirical application is provided for the copper market although we also apply the technique to other non-ferrous metals. Our results indicate that the copper price (and other non-ferrous metals) tends to follow an explosive path when stock levels are low and a mean reverting path when there is a glut in the market. However, most metal markets are primarily in the efficient state. The procedure allows us to assign probabilities to each of the 3 different states in each period and thereby allows us to analyse whether an asset market is efficient at a particular point in time through assigning a probability to each state. Thus, our results are intuitively plausible and support not only the partial equilibrium model in the chapter but also the augmented commodity storage model in chapter 5.

A restriction of the technique is that it is realistically limited to 3 states. Each state has to be imposed on the model i.e. it does not allow for two mean-reverting or two explosive states. Our procedure of estimating transition probabilities and the use of an AR(1) limits the type of states that can be detected. The chapter estimates probabilities for an efficient state, a mean-reverting state and an explosive state. Estimating transition probabilities for more than 3 states will prove to be considerably difficult. Additionally, the numerical algorithm will take considerably longer to estimate the parameters. Whereas this chapter introduces the notion of efficient and non-efficient states, a more general technique is introduced in chapter 4.

Chapter 4: '*What Proportion of Time is a particular Market inefficient?...A Method for analysing the frequency of market efficiency when equity prices follow Threshold Autoregressions*' substantially improves upon the results from Chapter 2 and 3. We make 3 major contributions in this Chapter including a detailed algorithm for a market efficiency metric which arguably is the most important contribution in this dissertation. The chapter revolves around the central theme of the dissertation i.e. that asset prices go through different states or regimes and one of these states is efficient (usually characterized by a random walk).

Our first contribution in this chapter is the derivation and analysis of analytical expressions for the first two moments of an asset price's distribution when the asset can be estimated as a first order threshold auto-regression, if these moments exist. A first order threshold auto-regressive model with an exogenous trigger can be written as:

$$P_t = \psi_{t-1} + \phi_{t-1}P_{t-1} + \eta_t \text{ where } \eta_t \sim N\left(0, \sigma_\eta^2\right) \qquad \textbf{(1.3)}$$

The state switching behaviour is based on the values taken by an exogenous trigger variable, Z_{t-1}. Thus, for a 3-state case, where Z_{t-1} takes values between $-\infty$ and ∞, (1.3) can be written as:

$$\psi_{t-1} = \alpha_1, \phi_{t-1} = \beta_1 \; if -\infty < Z_{t-1} < c_1$$

$$\psi_{t-1} = \alpha_2, \phi_{t-1} = \beta_2 \; if \; c_1 \leq Z_{t-1} < c_2$$

$$\psi_{t-1} = \alpha_3, \phi_{t-1} = \beta_3 \; if \; c_2 \leq Z_{t-1} < \infty$$

where c_1 and c_2 are threshold levels that need to be estimated. The unconditional probability that Z_{t-1} will be between any two thresholds c_j and c_{j+1} is given by π_{j+1}. It is these thresholds that tell us how efficient the market is. The values taken by parameters under different states are not necessarily consistent with a stationary model; thus, if a price series is non-stationary for a substantial length of time the overall price process may become non-stationary and thus, a steady state distribution for the process would not exist.

We derive the conditions that need to be satisfied by a process so that its steady state distribution and its first two moments exist. The conditions and expressions are derived for 3 different cases: a case with no drift, a case with constant drift in all states and a case with a switching drift. We note that when a process has a switching drift in addition to switching slope parameters the moments of the process depend not only on the variance of the slope parameters but also the variance of the drift term. This explains why model specification is important when estimating a threshold model. Inclusion of a switching drift term instead of a constant drift term implies that a lot of the variation in the process is captured by the switching drift term; theoretically too, a switching drift term provides a different explanation for explosiveness than the case with a constant drift term or a case where no drift is assumed.

Derivation of conditions for the existence of a mean and variance for a threshold auto-regressive model raised a number of issues. We realized that when Evans pointed out the inability of bubble tests to detect non-stationary behaviour, what he found was that if series were efficient (or had a unit root) for a substantially long period of time, conventional tests would have very low power to detect an explosive root (considering the threshold model above that would imply $\phi_t > 1$). While Evans considered 2 states, we consider a 3 state case; in addition to a weakly efficient state with $\phi_t = 1$ and an explosive state $\phi_t > 1$, we also consider a mean reverting state $\phi_t < 1$. The motivation for considering a mean reverting state comes from our results in chapter 3. Thus, the second contribution made in this chapter is with regards to the power of bubble detection tests when a process follows a threshold auto-regressive model with a mean reverting state in addition to an explosive and an efficient state.

We consider processes similar to 1.3 and specify a range of probabilities for our numerical analysis. Some of these probabilities imply that the simulated process does not meet the criterion for the existence of a mean and variance. In addition to considering exogenous triggers we also consider a

trigger variable that follows a Markov process. The Markov-state process is the more realistic scenario as the type of triggers we consider tend to be auto-correlated. The test we consider for the power test is the GSADF statistic which we have used in this dissertation previously. We use the GSADF test since the authors have shown that the test has higher power than other contemporary tests. Our results show that the power of the GSADF also reduces considerably when a mean reverting state is present. Whenever the parameters and probabilities are such that steady state moments exist, we find that the power of the tests is low. Thus, we make an important contribution in the bubble detection literature: ignoring a mean reverting state can artificially inflate the power of such tests. When applied to real data, these tests may fail to detect bubbles in series which spend a substantial proportion of time in a mean reverting state (and as we have noted in Chapter 3, a number of non-ferrous metals do exhibit mean reverting behaviour).

The final and arguably the most important contribution made in this chapter is a methodology for measuring efficiency in an asset market. Estimating equation (1.3) by restricting one state to be efficient (i.e. restricting ϕ_t to be 1) and identifying thresholds measured through non-linear least squares (or through an alternative estimation methodology), we can characterize different states. Not only does this provide us estimates of parameters, thereby enabling us to say whether the asset shows explosive or mean reverting behaviour, but also allows us to measure the length of time an asset spends in the efficient state. We also note, that if the parameters in the non-efficient states are not statistically different from the parameters in the efficient state, we may argue that the market is always efficient.

In contrast to the methodology outlined in Chapter 3, we can consider more than 3 states. The algorithm can be programmed easily even though the numerical burden may increase considerably when more than 3 states are considered. Importantly, we do not need to place any restrictions on the inefficient parameters. We may find only explosive, only mean-reverting or only efficient states. However, the algorithm also requires finding an appropriate trigger variable. While the trigger variable may itself be Markovian, we have to ensure that it is independent of the error term of the underlying process. Thus, the main challenge for using the algorithm is the availability of a suitable variable that may employed as a trigger.

To demonstrate how the metric may be used in practice, we provide an illustrative example using the S&P500 and the FTSE100 stock market indices. The trigger variable used is the Michigan State University's Consumer Sentiment Index (or MCSI). We justify the use of the trigger variable in the chapter and use our methodology to estimate the indices in a 3-state setting. Since the example is illustrative we consider all 3 cases outlined in the chapter, i.e. a case with no drift, a case with a constant drift and a case with switching drift. As we have previously mentioned, the specification of

the drift term is important; it changes our estimates as well as the interpretation of our results. A detailed commentary of the results is reserved for the chapter.

Our setting raises an interesting question for economists and financial analysts alike; what causes the state-switching behaviour in commodities and what determines how long you may be in a particular state? Since our main class of assets in this thesis has been commodities, we try to answer this question for storable commodities in Chapter 5.

Chapter 5, '*Making the Commodity Storage Model Empirically relevant,*' resurrects a commodity storage model first introduced by Williams and Wright in 1991. The original model has been through various iterations over the years but had failed to effectively capture what it had set out to do i.e. capture features of real world commodities. Specifically, the model tried to replicate the correlation structure of agricultural commodities as well the tendency of commodity prices to jump occasionally (what we would refer to as a bubble or an explosive state in the context of this thesis).

The model is a partial equilibrium model with demand, supply and a storage sector. Storage is the main driving force in the model and depends upon the availability of the commodity; availability in the model's setting is defined as current period production and previous period's storage. Current period production depends upon the prevailing weather conditions. Periods with low production resulting from adverse weather conditions often lead to a state where no storage is held; this substantially increases equilibrium price in the period and also leads to non-linearity in the model solution which has to be solved numerically. On the other hand periods where bumper crops are observed are characterized by high storage and low prices. These conditions would help us explain why different states occur. When prices are high due to low crop output we enter explosive states and when storage is high we enter mean reverting states.

Our main contribution in this chapter is making this model empirically relevant; we introduce 3 changes in the model which give us a solution that allows us to simulate prices that have similar distributional properties to real commodity data. We also use the GSADF statistic to test for bubbles in real data and find that while nominal prices show statistically significant evidence of an explosive root, real prices show no such evidence even though there is evidence of sharp spikes and downturns. Our solution and simulated prices are able to replicate both features.

Specifically, we make 3 simultaneous changes in the model, 2 of which have been used in this literature previously but have not been used together. Firstly, we use an iso-elastic demand curve as opposed to a linear demand curve (Gouel, 2013). Secondly, we use a multiplicative convenience yield first used by Ng et al (2000). Finally, we introduce saving behaviour in the demand curve by making the demand curve a function of not just current price but also future period consumption and hence, storage. The final change allows us to better capture the change in demand when storage switches.

The main benefit of our approach as opposed to alternative approaches that also claim to replicate features of real data is that we retain the simplicity of the model and the numerical solution approaches that were first used by Williams and Wright. Our focus also stays on storage. We abstract from using a general equilibrium model as it adds significant complexity and numerical burden.

Our research also has other features which have been missing in other articles on the subject. We carry out a thorough sensitivity analysis to understand what parameters influence the correlation structure that we find and whether these parameters make economic sense. Price elasticities of demand and supply along with the variation in weather are the primary determinants of price and in the chapter we provide a detailed argument supporting our findings.

Finally, we also conduct empirical analysis using a novel identification strategy recently introduced by Roberts and Schlenker (2013) which relies on geographical and weather data to correctly identify price elasticities. We empirically estimate price elasticities of demand for agricultural commodities using their approach and calibrate our model accordingly. Even though we calibrate only one parameter, we note that results from our calibrated simulation capture features of real commodity prices reasonably accurately and are a substantial improvement compared to previous work carried out in the field. Thus, we make significant improvements to a model which allows us to capture features of real commodity data through the state-switching mechanism.

Throughout this dissertation we emphasize the importance of thinking of markets as going through different states. This allows us to understand and estimate prices processes better, enabling us to identify whether state switches are temporary or permanent. We believe that thinking about financial assets in this way opens up the door to a lot of other avenues of research, both theoretical and empirical. We have focussed substantially on explosive states and most empirical applications have relied on commodity data; however, we strongly believe that the methodologies we have provided (particularly in Chapters 3 and 4) are applicable to a wide range of asset classes.

Each Chapter in the dissertation is self-contained and Chapters may be read independently of each other although we do make cross-references across chapters to highlight how the dissertation is connected.

CHAPTER 2: *Detecting multiple collapsing bubbles: A modification to the Bhargava Statistic*

Chapter 2 applies a significant modification to the Bhargava-Sargan (1986) statistic which enables detection of multiple collapsing bubbles in asset prices thereby addressing Evans' 1991 criticism. We compare the statistic to the recently introduced GSADF statistic by Philips et al (2013) and show that the modified Bhargava statistic compares favourably with the Philips et al statistic. Using data on 25 commodities we show that the Bhargava statistic is able to detect multiple bubbles in this setting and is able to identify periods in which these may have occurred.

2.1   Introduction

The financial crisis of 2007-2009 revealed the vulnerability of financial markets to changes in expectations. Volatility in credit, stock and commodity markets saw assets attain very high prices before collapsing. All major financial crises in the past two centuries have been preceded by, what has been termed, bubbles in asset prices and rapid credit growth. Given the consequences of such crises, econometricians have taken an interest in devising techniques to detect such episodes. While these techniques can be used to improve our understanding of historical episodes of asset bubbles and credit growth, they may also be used as a mechanism for early detection of over-exuberance in financial markets.

As mentioned in the introduction, bubbles are to be understood from an econometric perspective i.e. instances of explosive behaviour or non-stationarity, usually triggered by speculative activity or through a trigger event (such as a stock-out for commodities) with the explosive behaviour sustaining for a few time periods and then usually followed by a precipitous decline. While a number of articles do come up with theoretical models of bubbles by separating the price of an asset into a fundamental and a bubble component, application of the tests to real data does not make this distinction obvious. Some attempts have been made to separate fundamentals and bubbles to give a more precise quantitative estimate of bubbles. The interested reader is referred to Alessandri (2006) who uses a Kalman filter approach to separate bubbles and fundamentals in US Stock market data.

At the outset we would like to make a distinction between explosive bubbles that periodically collapse and bubbles that are found in other economic literature. The type of bubbles that we consider here are

perhaps best understood in the context in which Evans (1991) explains them i.e. bubbles that see periods of explosiveness or non-stationarity which are then followed by a very rapid decline. They are feature of asset markets. Such bubbles arise over short horizons, usually a few months, and then collapse equally quickly. Chapters 3 and 4 consider bubbles one of many states an asset goes through. In this chapter, the treatment is slightly different; instead of treating a bubble-like episode as a state we regard it as a temporary deviation from the efficient state.

Bubbles considered in Macroeconomic literature on the other hand, such as those considered by Tirole (1985) Carvalho et al (2012) and Ventura et al (2012), are different. Bubbles in their research grow much slower and do not collapse as often. Bubbles are modelled as a feature of the asset market and they keep growing over long periods of time and their collapse leads to business cycles. A lot of progress has been made recently on bubble detection techniques with some of these techniques also allowing for dating of exuberant episodes. The most significant recent contribution has come from Philips et al (2013, 2010) who use a recursive Augmented Dickey Fuller test to both date and detect bubbles.

The current chapter looks at an alternative to the technique developed by Philips et al. We propose a new methodology using the Sargan-Bhargava (1986) statistic to detect bubbles in asset price series. The modification also allows us to identify specific periods that show evidence of exuberant behaviour and thus, allows an alternative dating mechanism. We show that this modified Bhargava test procedure can match the Philips et al test in terms of power when large samples are considered.

For smaller sample sizes, the modified Bhargava test does better than the Philips et al test. The test also has the advantage of not requiring the estimation of any parameters; thus, in the domain of elliptical distributions, the conclusions derived from this test will not differ significantly regardless of the distribution of the error process. The statistic also has the advantage of having an exact distribution, i.e. the critical values are not approximate unlike those used by Philips et al. We compare and contrast the two tests to show that the tests are more complementary in nature and lead to the same conclusions.

An alternative way to think about this test would be that it detects periods when the asset price moves to an explosive state. In Chapter 4 we outline conditions that are required for a variable to have a steady-state distribution. Tests such as the one introduced by Philip et al assume that such a steady-state distribution exists and thus, rely on stationarity of the underlying process. The modified Bhargava statistic on the other hand, due to its local nature, requires no such assumptions to be made and may be applied on subsamples.

The second contribution of this chapter is with regards to the application of these tests. Tests of explosiveness have largely been carried out on stock market data using a case study approach. A

typical article in this literature usually considers the S&P 500 or the NASDAQ indices with some currencies and commodities added on. Our research focuses on commodities; we apply the modified Bhargava statistic to 25 commodities. We also applied the GSADF test to the same commodities for comparison purposes; these results have not been reported but are available upon request.

In light of the recent financial crisis, we believe that commodities have become increasingly important and they exhibit exuberant behaviour similar to that found in stock markets. Apart from buyers and sellers, there is a substantial market for commodity storage which tends to drive prices. Additionally, the advent of exchange based trading has allowed investors to enter the market without physically storing the commodity and thus, there is scope for speculation. Leverage levels are often high resulting in significant speculative activity. Climate change has further added volatility to commodity markets. There is a large futures market for commodities and fund managers have started paying increased attention to commodity returns when optimizing asset allocation. The London Metal Exchange Traded over 156 million contracts in 2016 (as opposed to 85 million in 2007) for metals alone, with contract sizes ranging from 5 tonnes to 25 tonnes (London Metal Exchange, 2016). Thus, we believe, that commodities, as an asset class, is ideal for bubble analysis.

With large metal exchanges in Chicago, New York, London and Shanghai, trading activity in commodities has been on the rise. A number of Exchange Traded Funds (ETFs) also follow commodities, enabling investors to enter this market. Commodities lend themselves readily to this type of analysis given that commodity markets often go through peaks and troughs. Looking at a large data set of commodities enables us to see how bubbles in commodities tend to emerge and collapse together. We show that the modified Bhargava statistic is particularly attuned to detecting such behaviour.

In this chapter we test bubbles in 25 commodities including agricultural, mining and energy commodities. The article closest to our research which uses 32 commodities is Gorton et al (2013); however, the focus of that article is to estimate the risk premia in commodities using data on futures. A more recent article by Etienne et al (2014) looks at bubbles in 12 agricultural commodities. They use daily futures data and use the GSADF test. However, we use a more comprehensive data set which allows us to comment on contagion in commodity markets. Thus, in addition to the new test and dating technique our empirical application is also unique.

Our results indicate three main periods for commodity price bubbles; these being the two oil crises in the 70's and 80's and the 2007-08 financial crisis. Both the modified Bhargava and the Philips et al tests detected bubbles in a multitude of commodities during these periods, vindicating the proposition that financial crises may be preceded by exuberant behaviour in asset prices.

The Chapter is organized as follows: Section 2.2 presents a chronological literature review of the different techniques that have been applied over time to detect bubbles. This provides a context to the bubble literature and a motivation for our modification to the Bhargava statistic and our selection of the Philips et al. statistic as a comparative statistic. Section 2.3 explains the proposed modification to the Bhargava statistic and also provides an outline of the Generalized Supremum Augmented Dickey Fuller statistic. Power comparison between the two tests is carried out in Section 2.4. Section 2.5 looks at the commodity price data that has been used in the current study. Section 2.6 presents our empirical application using commodity data and section 2.7 concludes.

2.2   Testing for Explosiveness – literature review:

The backdrop of testing for bubbles is the motivation to either confirm or refute the efficient market hypothesis. Testing for bubbles has evolved along with new developments in Time Series and Financial Econometrics. With new tools and more powerful computers econometricians have adopted more sophisticated methodologies to test for unit roots and explosive roots in order to identify bubble behaviour. The pivotal point in bubble literature came in 1991 when Evans in his seminal article discussed the notion of periodically collapsing bubbles and the difficulty of detecting a bubble in the presence of such behaviour. Almost all tests of bubbles are reduced form tests which specify an autoregressive model for the price process of the asset under question. The current section takes a chronological view in discussing the vast literature that relates to testing for bubbles.

2.2.1 Volatility based tests:

Early tests for bubbles centred on finding excessive volatility than was theoretically possible. Leroy and Porter (1981) were among the first to use volatility to test whether stock prices violated the efficient market principle, thereby indicating the presence of a bubble. They specified an ARMA model and constructed upper and lower bounds on the variance of asset prices. They showed that the upper bound on the theoretical variance of a price series should be the sum of the variance of the observed price and the variance in deviation from the fundamental price. The theoretical variance is derived solely from the variance of fundamentals which in their case was data on dividends. Blanchard et al (1982) and Mankiw et al (1985) extend these variance tests to overcome some of their limitations.

Mankiw et al try to address criticisms levelled against volatility tests including the low power resulting from finite samples i.e. small sample variance for stock prices can be biased and be expected to be higher than the fundamental stock price due to the use of sample means instead of population means. Secondly, even in large samples variance could be biased if the dividend process were non stationary and instead followed a random walk. They posit a test that does not bias this variance. Instead of finding variances around the mean price, they calculate the test statistic around a

hypothesized Naïve Forecast Price described as $P_t^0 = \sum_{k=0}^{\infty} \beta^{k+1} F_t D_{t+k}$, where the naïve forecast is known to the market. Their test statistic is based on the following relationships:

$$E(P_t^* - P_t^0)^2 = E(P_t^* - P_t)^2 + E(P_t - P_t^0)^2;$$
$$E(P_t^* - P_t^0)^2 \geq E(P_t^* - P_t)^2 \qquad (2.1)$$
$$E(P_t^* - P_t^0)^2 \geq E(P_t - P_t^0)^2$$

P_t^* is the perfect foresight price P_t is the observed stock price. The perfect foresight and observed prices differ by an error term v_t. Assuming the naïve forecast to be a fixed process of past dividends and plausible values for the real interest rate, Mankiw et al test for these relationships in similar vein to Shiller and LeRoy et al. Though the differences they find are less marked than earlier authors, they continue to find violation of the inequalities specified above, concluding that there continued to be evidence of fads.

Tests based on theoretical volatility were criticized by various authors on the ground that the specification of fundamentals may be inherently flawed and that part of the high frequency can be explained by including unobserved components of fundamental price. Hamilton and Whiteman (1985) were among the critics of the variance test approach. In their article they contend that a significant deviation of variance from the variance of the dividend processes can also be an indication that fundamentals are not properly specified, rather than the process being in a bubble. Hamilton suggests that portfolio managers use more information to form their fundamental price and part of this information is not available to the econometrician.

This debate, however, does appear to be circular as one can always find alternative specifications for fundamentals that may attempt to explain the variance of a particular series during an explosive period. New fundamental specifications would require rigorous theoretical foundations so that fundamentals and bubbles can be separated. Summers (1986), provides a different criticism of these tests, namely that these tests have very low power to reject the hypothesis of market efficiency.

2.2.2 Tests of Co-integration and mean reversion:

Asset prices showing bubble tendencies tend to show evidence of non-stationarity and are likely to contain explosive roots. Diba and Grossman (1984) exploited this to outline a strategy for detecting bubbles in the price of gold and other assets. They argued that explosive components of a price process cannot be made stationary through differencing. Thus, if differencing a price process and its fundamentals makes fundamentals stationary but not the price process, a bubble may be present. Diba et al also use this methodology in their 1988 paper and apply the methodology to stock prices instead of gold.

Campbel et al (1987) formulate a simple VAR which requires that the spread on an asset be equal to its theoretical value (in the case of stocks, the expected future values of dividends); if these values differ, there will be evidence of a non-predictable term or a bubble. West (1987, 1988) also postulates tests using the stationarity properties of the price process. He proposes that the estimates for parameters in the present value model should be the same when estimated using two different methodologies. He measures the price series using a simple regression and an ARIMA equation for the dividends process allowing the parameters to be estimated implicitly. Intuitively, the information set used to forecast a price has less information compared to the full information used to determine the real price. This lack of information leads to the variance of the forecast being theoretically larger than the variance of the price series.

Fama and French (1988) and Summer and Poterba (1988) use variance ratio tests to find evidence that stock prices show mean reversion. They distinguish between temporary and permanent components of stock prices arguing that the long term component would show negative correlation thereby showing evidence of mean reversion. Violation of these trends will indicate the presence of bubbles although the results noted were generally mixed. Bubbles are not periodic and may arise at uncertain intervals and collapse at later dates. This criticism considerably lowers the efficacy of the tests discussed above. Evans (1991) in his seminal paper first highlighted this issue and hinted at a strategy that could be used for further research in this area.

2.2.3 Tests of explosive roots and switching regressions:

Evans (1991) argued that if rational bubbles were periodically collapsing they will tend to appear like a random walk and fall out of scope of the stationarity and co-integration tests being performed by his contemporaries. Evans' article marked a critical point in bubble testing literature and led researchers to focus their attention on periodically collapsing behaviour. Evans used the original Bhargava statistic to show how even the most powerful test for detecting bubble behaviour was unable to detect explosiveness in a series with multiple collapsing bubbles.

Hall et al (1999) tried to address the issue by specifying a Markov-switching model for the auto-regressive parameter and using a more generalized form of the Augmented Dickey Fuller test in order to detect periodically collapsing bubbles. Tests of unit roots and co-integration have low power in the presence of periodically collapsing bubbles and a series containing a bubble may appear stationary if the bubble collapses within the series. The specification used by Hall et al for their Markov-Switching Unit root test is as follows:

$$\Delta p_t = \mu_0(1 - s_t) + \mu_1 s_t + [\varphi_0(1 - s_t) + \varphi_1 s_t]p_{t-1}$$
$$+ \sum_{j=1}^{k} [\psi_{0j}(1 - s_t) + \psi_{1j} s_t]\Delta p_{t-j} + \sigma_e e_t \qquad (2.2)$$

's_t' is specified to be a homogenous Markov chain on the state space {0,1} with specified transition probabilities. Thus, when s = 1, the price series is in a bubble and when s = 0, the bubble collapses and the price reverts to fundamentals. Their tests show a marked improvement compared to Evans' test and they are better able to detect bubbles. We use a similar but significantly modified methodology to detect not just bubble states but also mean reverting states in Chapter 3. It is also worth noting that estimating a price process this way is only possible if the conditions for the existence of a distribution are satisfied

Van Norden et al (2002) separately test fad and bubble specifications within a switching-regression framework. They aim to identify which one of the two specifications was more in line with data. Other authors who have used similar switching approaches to test for bubbles include Enders et al (2001) who aimed at a generalization of the Momentum Threshold Auto Regressive approach (MTAR) for multivariate models and Bohl (2003) who used the Enders et al tests to test for the presence of periodically collapsing bubbles in Stock Market indices. Philips et al (2010, 2012) use a recursive regression methodology to date bubbles and analyse a number of asset prices during the financial crisis.

Knight et al (2014) derive a steady state distribution for a model that includes bubbles in the context of regime switching where the bubble process is modelled as an exogenous sunspot. They use results derived in Knight and Satchell (2011) and examine conditions under which a steady state distribution of price exists. 3 different distributions for the error process are considered and necessary and sufficient conditions derived that would ensure the existence of a steady-state distribution.

Using the results obtained they calculate conditions which would ensure that moments are finite and exist. This enables them to show why tests of co-integration that try to test for bubbles through standard and DF unit root tests fail to detect the presence of bubbles and have low power. The model used is an extension of the Blanchard and Watson (1982) model where the price process is specified as: $p_t = p^* + c_t$ where p^* represents the fundamental price and c_t represents the bubble term. Given this price process the generalized Knight and Satchell model takes the following form:

$$p_t - p^* = \beta_1(p_{t-1} - p^*) + \varepsilon_t \; if \; I_{t-1} = 0 \tag{2.3}$$
$$p_t - p^* = \beta_2(p_{t-1} - p^*) + \varepsilon_t \; if \; I_{t-1} = 1 \tag{2.4}$$

And I_{t-1} is defined as:

$$I_{t-1} = 0 \; if \; Z_t \leq c \; with \; probability \; 1 - \pi \tag{2.5}$$
$$I_{t-1} = 1 \; if \; Z_t \geq c \; with \; probability \; \pi \tag{2.6}$$

Z_t is the forcing variable in the model. Letting $\beta_1 = 0$ and $\beta_2 = \beta$ $with$ $X = p_t - p^*$. This leads to the simplified model $X_t = \beta I_{t-1} X_{t-1} + \varepsilon_t$. Backward substitution leads to $X_t = \sum_{j=0}^{\infty} B^j \prod_{k=1}^{j} I_{t-k} \varepsilon_{t-j}$. This model allows Knight and Satchell to derive the conditions required for strong stationarity. We build upon these results in Chapter 4 which also provides more details on this methodology.

Philips, Shi and Yu's recent contribution has been the most recent development in bubble literature and we talk about their methodology at length in the next section. Since we employ the GSADF statistic throughout this dissertation, we provide a detailed overview of the statistic.

2.3   Methodology:

In light of the above discussion, it is evident that any serious research on bubble detection will need to address Evans' criticism. His criticism also suggests that markets can appear to be efficient even if significant inefficiency is present; this is something we build up on in Chapter 4. Our proposed modified Bhargava test attempts to address this criticism. The test is also applicable to a wide range of situations as we will show and the fact that the test is designed to be a local test allows it to be used in small samples. Our empirical application of the test also fills a gap in literature as most empirical bubble testing has been focussed on stock price indices. While interest in commodities has increased, a coherent study looking at commodities as an asset class has been lacking.

2.3.1 Modified Bhargava Test:

The Bhargava statistic, also referred to as the Bhargava-Sargan statistic (1986), is the locally most powerful invariant statistic to test for a unit null against an explosive alternative hypothesis. It does not rely on any parameter estimates from the underlying process under question and thus allows one to test for explosive behaviour in small samples reasonably accurately. Consider the reduced AR(1) form for an asset price series:

$$p_t = \phi p_{t-1} + \epsilon_t \qquad\qquad (2.7)$$

where ϵ_t is assumed to be a white noise process and ϕ_t is the autoregressive parameter and the parameter of interest which drives bubble behaviour. The null hypothesis under consideration is:

$$H_0: \quad \phi_t = 1 \quad for\ t = 1,2 \ldots \ldots .. T \qquad\qquad (2.8)$$

The alternative hypothesis tests for an explosive root in the series, i.e.

$$H_1: \quad \phi_t > 1 \ for\ t = 1,2 \ldots \ldots \ldots .. T \qquad\qquad (2.9)$$

The original Bhargava (1986) statistic is:

$$B = \frac{\sum_{t=2}^{T}(p_t - p_{t-1})^2}{\sum_{t=2}^{T}(p_t - p_0)^2} \qquad (2.10)$$

If the series under consideration has a drift component, the Bhargava statistic is:

$$B = \frac{\sum_{t=2}^{T}(p_t - p_{t-1})^2 - \frac{1}{T-1}(p_T - p_0)^2}{\frac{1}{(T-1)^2}\sum_{t=1}^{T}[(T-1)p_t - (t-1)p_T - (T-t)p_0]^2} \qquad (2.11)$$

We can see from the statistic that the test is not designed for large samples. With large sample sizes, the statistic approaches zero and thus, the test becomes trivial. Our methodology will not only allow us to apply the test to larger samples but will also try to address Evans' criticism. The original Bhargava statistic did not incorporate a structural break in the series; thus, if one had to test for a bubble in the series, it was assumed that the series will be in an explosive state for ever or at least explosive enough for the right-sided unit root test to yield statistically significant evidence. Some researchers have applied modifications to the Bhargava statistic. Breitung et al (2010) alter the statistic so that it is able to detect a structural break in the series. They use a version of the inverse of this statistic for their testing procedures. The shortcoming of their modification is that the statistic is still unable to detect multiple bubbles in a long time series as it identifies only 1 structural break.

We use the original form of the statistic but significantly alter how it is applied to data. Bhargava himself and Evans, when testing his collapsing bubble hypothesis, applied the Bhargava statistic only once for the full series. With availability of longer and more frequent data, it is possible to apply the statistic to sub-samples. Since the test is meant to be a local test with limited observations available, the statistic can readily be applied to truncated sub-samples. The statistic is particularly useful to determine short term exuberance which characterizes behaviour of assets in general and commodities in particular during periods of crises. Bhargava noted the exact critical values for his test statistic which are contained in table 1 of his 1986 article. He used different sample sizes varying from a size of 20 up to 100. The critical values are different for series with a drift term.

When Evans identified the need for high persistence in bubble terms to enable detection, he used simulations of 100 periods and each simulation contained multiple bubbles. Given the high number of collapses in his series, the test is unable to distinguish the series from a random walk. However, bubbles are a short or medium term phenomenon and with the availability of longer time series it is imperative that we will observe periods of collapse.

There appear to be multiple bubbles in the commodity series under consideration here; however, our methodology will help us isolate these instances. Our modification is as follows. Instead of

calculating the statistic over the whole series, we split the full sample into shorter sub-samples. Using a rolling window over the sub-samples we compute the statistic for each sub-sample. This allows us to isolate bubbles and identify periods where explosive behaviour is observed. Since the time series data we have is for a maximum of 687 periods, our analysis below uses 700 observations.

For illustrative purposes suppose that we divide our sample into sub-samples of 100 periods and use a rolling window of 50 periods. The rolling window is not necessary and is more a convenience tool to help in detection of explosive periods when testing multiple series. We start at the beginning of the series and calculate the Bhargava statistic for the first 100 observations. Next, using our rolling window, we calculate the Bhargava statistic for observations 50 to 150 and keep using the rolling window of 50 observations till the end of the sample.

Once the rolling methodology described above has been applied and the Bhargava statistic calculated for each sub-sample, we obtain a series of Bhargava statistics. Each entry in the series can then be compared to the critical values from Bhargava's original table at conventional significance levels. If a particular entry in the series is rejected, we conclude that there is evidence of explosive behaviour in the series for that time period. Otherwise, the null of a random walk appears more appropriate. Note that this statistic works differently from more conventional tests. The hypotheses are rejected for low values of the test statistic. The lower the value of the statistic, the more likely the series is to contain an explosive root. This also explains why this is more useful as a local test, as with longer series, the critical values will approach zero and we will no longer be able to distinguish between a random walk and an explosive series. The figure below, illustrates the methodology.

t = 0   t = T

Sub-sample 1

Sub-sample 2
Sub-sample 3

Overlap
window

Sub-sample 4
Sub-sample 5
Sub-sample 6

The Bhargava statistic is calculated at each sub-sample. The sub-sample size and the overlap window are decided by the researcher
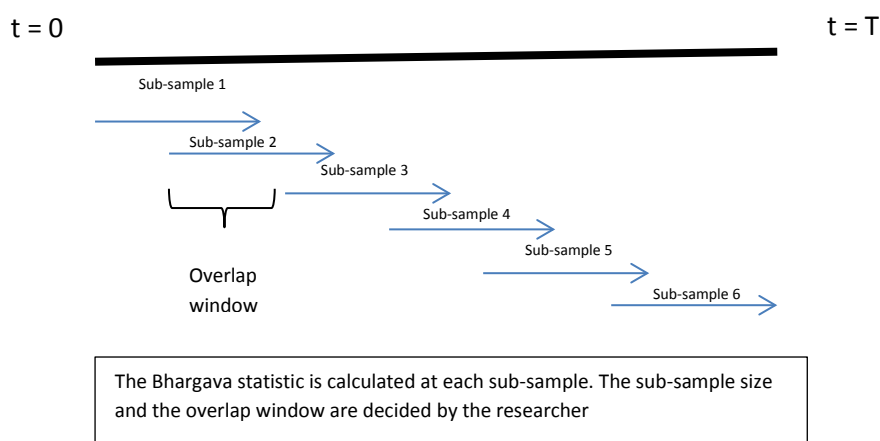
The null of a random walk allows us to use this procedure. Under the null we can assume that successive windows are independent as the difference between two consecutive periods is determined

by the error term and not be a deterministic term. The procedure would not have been valid for any value of ϕ other than 1.

In each window, the statistic will continue to follow the exact distribution as derived by Bhargava et al. If the exact cumulative distribution of the statistic is given by $F(r) = P(B < r)$, the distribution of the modified Bhargava statistic, $Bh = \min(B_1, B_2, \dots B_n)$, can also be derived (where $B_1, B_2, \dots B_n$, represent the Bhargava statistic in each window).

$$Bh = P(\min(B_1, B_2, \dots B_n) < r = 1 - P(B_1 > r, B_2 > r \dots B_n > r)$$

Under the null hypothesis of random walk, B_i is uncorrelated with B_j if $i \neq j$. Thus,

$P(B_i > r) = 1 - F(r)$, which implies:

$Bh = 1 - \left(1 - F(r)\right)^n$, where n is the total number of windows.

$F(r)$ does not have a closed form solution; Since the Bhargava statistic is a ratio of quadratic forms in normal variables, the exact limit for a test size α is calculated using the Imhof procedure (Bhargava, 1986):

$$P[\sum_{i=1}^{T-n} (1 - r\delta_k)z_i^2 \leq 0] = \alpha$$

Where z_i are i.i.d (0,1) and δ_k represent the (T-n) smallest eigenvalues of a symmetric $T * T$ matrix G, given by

$$G_{ij} = (T + 1 - j) \qquad j \geq i \quad (i, j = 1, \dots, T)$$

For a test of non-stationariy, the Eigen-values are of the form:

$$\delta_k = \frac{1}{\left[4 \sin^2\left(\frac{2k-1}{2T-1}\left(\frac{\pi}{2}\right)\right)\right]} \qquad (k = 1, \dots, T - 1)$$

Further details can be found in Section 3 of Bhargava (1986). We used simulation analysis to derive critical values for Bh, and have included our code in Appendix 2 below. We find that the exact critical value for a given window size are the same as derived in the original article. The critical values depend on the size of the window with smaller window sizes requiring a higher critical value. We report our simulated critical values in Table 2.4.

As the series moves farther away from its initial sub-sample value, the denominator increases faster than the numerator even though, the numerator is also increasing due to the presence of the bubble. y_0

is adjusted for each sub-sample for which the statistic is calculated. If the null hypothesis of a unit root is rejected in favour of the alternative of an explosive root, we can say there is evidence of bubble like behaviour in the series which occurs at some point during the sub-sample under consideration. Thus, in a 700 observation series containing 13 sub-samples, each 100 observations long with an overlapping window of 50, we would say that a bubble is present in the series if for any of the sub-samples our calculated statistic exceeds the critical value of 0.022. The sub-sample in which this statistic is found can be classified as an explosive period.

This helps us address the multiple collapsing bubble issue highlighted by Evans. A caveat to be noted is that even though we can detect multiple bubbles in the full series using this methodology, Evans' critique will still remain valid for each sub-sample. Thus, if there are multiple bubbles within a sub-sample, the modified test procedure may fail to detect them. However, this may be resolved using a different sub-sample size. Reducing the sub-sample size may enable us to detect these bubbles. The question of what sample size and overlap size to choose is discussed in Section 2.4.

2.3.2 Philips, Shi and Yu test

Philips, Shi and Yu (2013, PSY henceforth) built upon the earlier work of Philips, Wu and Yu (2010, PWY) and devised the Generalized Supremum Augmented Dickey Fuller statistic (GSADF) in order to detect multiple collapsing bubbles. They also suggest a methodology for dating bubbles. For a random walk process with an asymptotically negligible drift they devised a recursive regression methodology to detect the presence of a bubble in a particular price series. Specifically, the data generating process used is:

$$p_t = dT^{-\eta} + \theta p_{t-1} + \varepsilon_t, \qquad (2.12)$$

where $\theta = 1$, d is a constant, T is the sample size and $\eta > 1/2$. The recursive right-tailed Augmented Dickey-Fuller statistic is then calculated by specifying an initial window r_w. The empirical regression begins with the first value in the series r_0 and computes the Dickey-Fuller statistic for the first r_w terms in the series. Recursions are then carried out by increasing the number of terms used to run the regression and calculate the Dickey-Fuller statistic 1 additional observation at a time. The last regression uses all observations in the series and the window spans the whole series. The Sup ADF statistic is then found as the supremum of the series of ADF statistics computed using the recursive regressions above. Their empirical regression model is

$$\Delta p_t = \alpha_{r0,rw} + \beta_{ro,rw} p_{t-1} + \sum_{i=1}^{k} \psi_{r0,rw}^i \Delta p_{t-i} + \varepsilon_t \qquad (2.13)$$

In this model, r_0 is fixed while r_w increases by 1 observation for every recursion. k represents the lag order and ε_t is an iid disturbance term. The Sup ADF statistic (SADF) indicates if bubbles are present in a particular series. Date stamping a bubble requires the use of the series of ADF statistics generated by the recursive regression. According to PWY's methodology the start of a bubble is identified as the first point in the series when the ADF statistic exceeds the 5% right-tailed critical value based on the simulations. Similarly, the collapse of the bubble is identified as the first point following identification of a bubble when the ADF statistic falls below the 5% right-tailed critical value. Although this methodology could identify multiple bubbles, its power is relatively low compared to the GSADF statistic.

In contrast to the SADF statistic which uses an expanding sample with a fixed starting point, the GSADF statistic, re-initializes the recursive regressions by changing the starting point of the series. In essence this represents a double recursion to calculate right-tailed Dickey Fuller statistics. To illustrate this point further, suppose the first empirical regression comprises the first r_w observations and runs from r_1 to r_2, so that for the first regression $r_w = r_2 - r_1$. The following iteration in the series would keep r_1 fixed and increase r_2 by including one additional observation. Recursive regressions continually increase r_2 by one observation and calculate the ADF statistic until r_2 is the last observation in the series. The SADF test would terminate here.

The GSADF test on the other hand will re-initialize the recursion by moving r_1 from the first observation to the second observation and again using r_w observations to initiate the recursive process. The recursion will then start again with the ADF statistic being calculated, the sample being increased by 1 observation at a time and so on. The double recursion is repeated until $r_1 + r_w$ gives us the last observation in the series at which point the double recursion cannot proceed any further and the last recursion will represent the SADF for the last date in the series. The GSADF statistic is then defined as the supremum of the ADF statistics calculated from this double recursion (or the suprememum of the sequence of SADF statistics), i.e.:

$$GSADF = \sup_{\substack{r_2 \in [r_w, T] \\ r_1 \in [r_0, T - r_w]}} \left\{ ADF_{r_1}^{r_2} \right\} \qquad (2.14)$$

where r_0 is the first observation in the series. The limiting distribution of the GSADF statistic is:

$$\sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \left\{ \frac{\frac{1}{2} r_w [W(r_2)^2 - W(r_1)^2 - r_w] - \int_{r_1}^{r_2} W(r) dr [W(r_2) - W(r_1)]}{r_w^{\frac{1}{2}} \{ r_w \int_{r_1}^{r_2} W(r)^2 dr - [\int_{r_1}^{r_2} W(r) dr]^2 \}^{\frac{1}{2}}} \right\} \qquad (2.15)$$

where $r_w = r_2 - r_1$ and W is a standard Weiner process

The following diagram, taken from PSY, illustrates this point further.

SADF Test

GSADF Test
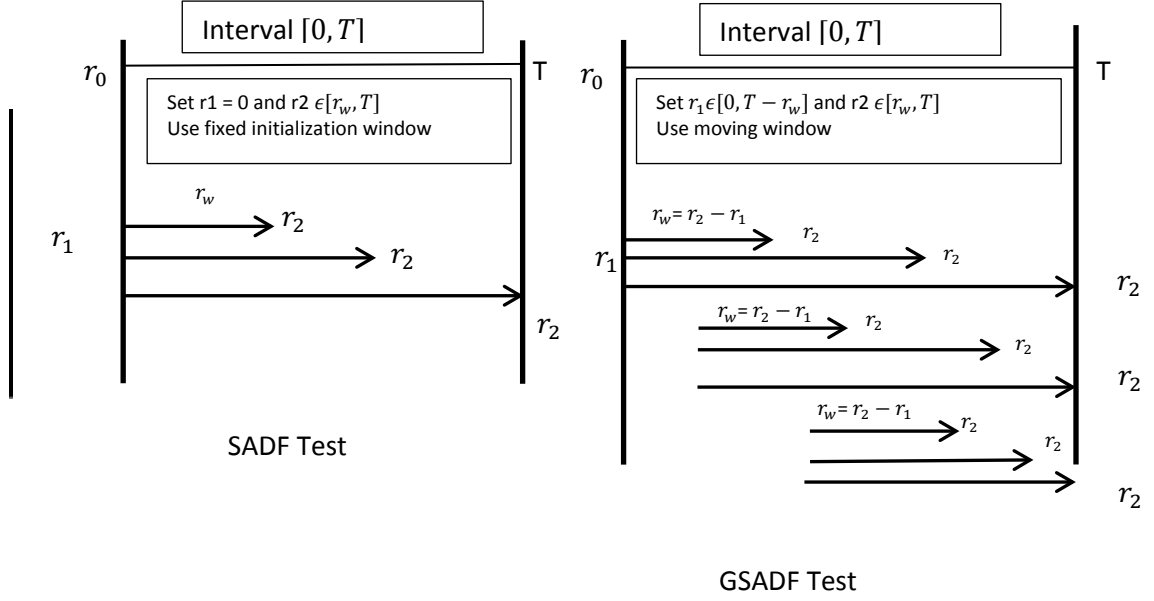
This double recursion leads to a sequence of sup ADF statistics which is used to date bubbles. Bubbles are dated by comparing a series' SADF sequence to the 95% critical SADF sequence which PSY obtained through simulating equation (2.15) and (2.16) above. While the PWY test allowed detection of multiple bubbles as well, the PSY test allows detection as well as dating of multiple bubbles. The PSY test also has higher power than the PWY test. Moving the starting point r_1, allows this statistic to overcome the Evans' criticism. Identification of bubbles is based on the following rules:

$$r_e = \underbrace{\inf}_{r_2 \epsilon [r_w, T]} \{r_2 : SADF_{r_2} > Critical\ Value\} \qquad (2.16)$$

$$r_f = \underbrace{\inf}_{r_2 \epsilon [r_e + r_b, T]} \{r_2 : SADF_{r_2} < Critical\ Value\} \qquad (2.17)$$

r_e is the start of the bubble, r_f is the collapse date of the bubble and r_b represents the length of the bubble. The SADF sequence for the concerned series is then plotted and compared to the 95% SADF sequence. When the SADF sequence crosses and is above the 95% sequence, PSY conclude that the series is in an explosive state. When the calculated SADF sequence falls below the 95% SADF sequence, the series is a random walk. Multiple bubbles may be identified and dated in this manner.

Philips et al classify an episode of exuberance a bubble if its length is at least log(T). We follow their suggestion and any explosive episodes shorter than log(700) = 3 months, are not considered in our results. We use the GSADF statistic to formulate our methodology in Chapter 3. In Chapter 4 we have more to say about the statistic and identify scenarios where the GSADF statistic may fall short.

As mentioned previously, PSY used the data generating process in (2.12) to obtain critical values using 5000 simulations and calibrating $d = \eta = 1$, T = 1680 and r_w= 36 observations. Instead of a critical value, this technique generates a critical '*sequence*' from the simulation SADF sequences. Thus, PSY base their inferences on the 95% SADF sequence of critical values. PSY were carrying out this test for historical US Stock price data and thus used a sample size of 1680. Our sample size is considerably smaller so we need to calculate critical values more attuned to our data.

We simulate the same data generating process that we used for the modified Bhargava test with 10000 replications but use T = 687 and r_w= 100 observations in order to compare the results to our modified Bhargava statistic results. We also simulated the data generating process by including a time trend with a small slope. Note, that the power of the test will be affected when the sample size is changed. A bigger sample size is more likely to detect bubbles than a smaller sample using this methodology. The critical values obtained when the time trend was included were not significantly different from the critical values without trend and were generally lower; thus, we conducted our test with the more prudent estimate. The critical values were also not very sensitive to the initial window size of a 100. Smaller window sizes did not yield very different critical values. The critical values obtained from simulation results are reported in table 2.3 below.

2.4   Power Comparison

In this section we compare the power of the two tests. Our first comparison considers a relatively large sample. Later in this section we also show that the modified Bhargava statistic performs better than the PSY test when we decrease the number of observations. This technique may also have implications for threshold autoregressive models and may help the researcher in identifying an explosive state or a regime change. While we do not touch upon regime-switching models in this Chapter, Chapter 3 and 4 will deal with such models in detail.

PSY showed that their GSADF statistic had very high power, particularly for larger sample sizes. In their article they use a range of sample sizes to calculate the power. The results of this analysis are contained in table 4 of their paper. The sample sizes they consider are monthly and of sizes 100, 200 and 400 respectively. The sample sizes that we are primarily concerned with, however, are slightly larger. We recalculate the power of both statistics in light of our commodity price data set.

In order to test the power of our Modified Bhargava Statistic we utilize the process originally used by Evans and subsequently built upon by West (1988). Another possible alternative we could have used was the one proposed in Knight and Satchell (2011). The main difference between the Knight and Satchell model and the Evans model is that the former model uses an exogenous trigger while, the Evans model uses an endogenous trigger in the form of a bubble process that together with the

fundamental process forms the prevailing price of an asset. The processes thus obtained look similar and the only reason for using the Evans model was to aid direct comparison with the power tests conducted by PSY for the GSADF statistic. We follow the notation used in Phillips et al.

The asset price consists of a fundamental component and a bubble component:

$$P_t = P_t^f + \kappa B_t \qquad (2.18)$$

where P_t^f is the fundamental component, B_t is the bubble component and κ is a constant that adjusts the magnitude of the bubble component and influences the impact of the bubble to the price process.

$$P_t^f = \frac{\mu\rho}{(1-\rho)^2} + \frac{\rho}{(1-\rho)}D_t \qquad (2.19)$$

with

$$D_t = \mu + D_{t-1} + \varepsilon_{D,t}, \quad \varepsilon_D \sim N(0, \sigma_D^2)$$

The term D_t is the driver of the fundamental price and can be thought of as a cash flow from an asset or dividends if one considers stocks. μ is the drift of the random walk process, ρ represents the discount factor (i.e. $\frac{1}{(1+r)}$). The bubble process on the other hand has a trigger value 'b'. The bubble process slowly approaches this trigger value at the rate $(1 + r)$ as per West (1988).

When the trigger value is reached the bubble process switches to an explosive process. The duration of this explosive process is determined by a Bernoulli process ϑ which takes on the value 1 with probability π. To clarify, when the process becomes explosive, the probability of the explosiveness continuing is π. Once the explosive process collapses (i.e. the Bernoulli process takes a value of 0), the value of the bubble process collapses to ζ after which the bubble starts slowly rising again until the trigger is breached.

$$B_{t+1} = \rho^{-1}B_t\varepsilon_{B,t+1}, \quad if\ B_t < b \qquad (2.20a)$$

$$B_{t+1} = [\zeta + (\pi\rho)^{-1}\vartheta_{t+1}(B_t - \rho\zeta)]\varepsilon_{B,t+1}, \quad if\ B_t \geq b \qquad (2.20b)$$

$\varepsilon_{B,t}$ is a log normally distributed exponential error term with a mean of 1 and a standard deviation of 0.05. The initial parameter settings are identical to the monthly parameter settings used by PSY and are contained in the table below. We note that the simulated process may not satisfy the conditions for a steady state distribution to exist (which have been outlined in Chapter 4).

Initial parameter values

μ	σ_D^2	D0	ρ	b	B0	π	ζ	κ
0.0024	0.0010	1.0	0.985	1	0.50	0.85	0.50	50

The random process generated using the above parameter values contains a number of collapsing bubbles as shown in the figure below.
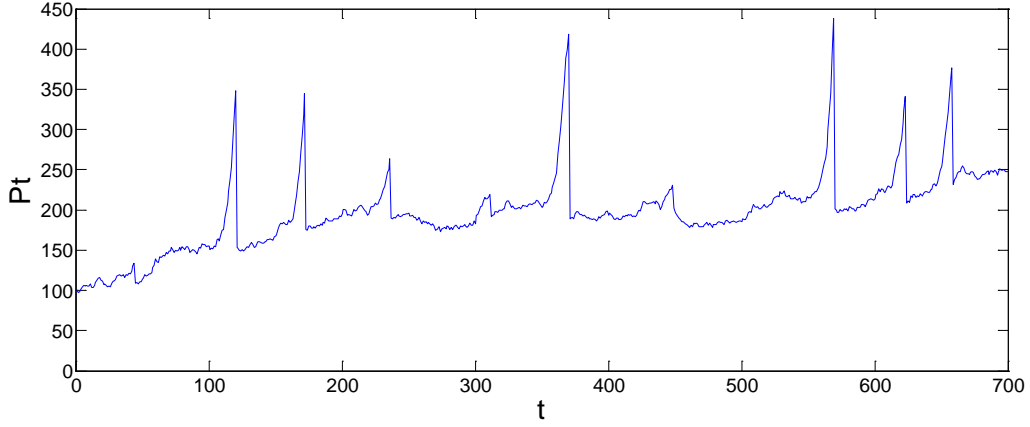
As seen above, the length of each bubble is shorter and there are many more bubbles in this process than the number we observe in our data. Both statistics under consideration are applied differently. While the GSASDF statistic uses a recursive procedure, the Bhargava procedure uses a rolling window; thus, the best comparison will require a small sub-sample size for the Bhargava statistic. We conduct our power test using different sub-sample sizes. In empirical application the size of the sub-sample will be dictated by the data available and whether the bubble to be detected occurred in the past or is currently expected to be present in a series.

Table 2.2 (in the appendix) shows results of power tests for both the Bhargava Statistic and the GSADF statistic. We use the data generating processes (2.18) – (2.20) to simulate series containing explosive behaviour and apply both the modified Bhargava Statistic and the GSADF statistic. Each series has 700 observations. Initial calibration is based on the parameter values above. For the GSADF statistic we use an initial window of 100. Using a smaller window does not change the power significantly, so we keep the initial window at 100 as we use the same window size for our empirical application. For the modified Bhargava test we use different sample sizes and employ an overlapping window that is half the size of the sub-sample. Critical values for the GSADF statistic are contained in table 2.3 and for the modified Bhargava test in table 2.4. We say that the series contains a bubble if the infimum of the modified Bhargava statistic series is less than the critical value as explained in section 2.4 above.

From Table 2.2 we see that the power of the modified Bhargava test varies according to the sub-sample size. When we consider a sub-sample size of 20 with an overlapping window of 10, the power of the statistic is 99% compared to 97% for the PSY test. Thus, we can clearly see that the modified Bhargava statistic holds its own in a power comparison against the PSY test; particularly when the sub-sample size is low which is the most relevant comparison given the recursive nature of the PSY test. For our empirical application, we use the sub-sample size of 100 and an overlap window of 50, primarily to make the analysis of data more tractable and interesting.

Our power test suggests optimal window-sizes given the size of the data-set. If the objective is to date explosiveness, smaller sub-sample sizes will be more appropriate for series with a low number of observations, or when the objective is to determine whether a series may currently be in an explosive state. For a smaller sample size (<100) a sub-sample size of 20 with a rolling window of 10 will be appropriate. For larger samples, a sub-sample size of 100 with a small rolling window (such as 20) will yield optimal results. This will keep the analysis manageable while not compromising on statistical power. We use a sub-sample size of 100 and a rolling window of 50 in order to be consistent with our empirical results and to keep the analysis tractable (while a smaller rolling window will have been more appropriate, using a rolling window of 50 allows us to make our results presentable; results with a smaller rolling window were not significantly different). Our results indicate that even with a relatively large sub-sample size, we are able to detect a number of explosive episodes in our data. On the other hand, if the user is not interested in dating explosiveness and is primarily concerned with its presence in a series, the smallest sample size (20) will be optimal given the results in Table 2.2.

Next, we show how the power of the Bhargava statistic evolves as we change the initial parameter values. We use a sub-sample size of 100 and an overlapping window of 50 observations. Again, we use the infimum of the modified Bhargava Statistic in order to test power. If the infimum of the modified Bhargava statistic series is less than the 5% critical value (0.022), the null hypothesis is correctly rejected in favour of the alternative. As shown above, we obtained a power of 39% for the initial parameter values. These are much higher than those observed by Evans (6.5%) for the same parameter values.

For the data at hand, the number of bubbles appears to be lower and of slightly longer duration, lasting between 1 and 2 years; thus, in order to conduct a more relevant power test for the commodity data we have, we adjust the parameters in order to generate series that have fewer bubbles but a higher duration. More precisely, the parameters we alter are π, the probability of a bubble continuing in the next period, which leads to longer bubbles and 'b', the trigger value, which determines how often explosiveness is triggered in the process. The table below shows the results of our power test for a

range of values of π and '*b.*' Each result was obtained using 5000 replications of the data generating process outlined above.

Power test for the modified Bhargava statistic

'b'	π			
	0.85	**0.90**	**0.95**	**0.99**
1.0	39%	52%	77%	97%
2.0	48%	61%	84%	97%
3.0	60%	70%	88%	98%

As seen above, the frequency of bubbles and the duration of each bubble appear to have a significant impact on the power of the modified Bhargava Statistic. Thus, the statistic, even with a high sub-sample size is more likely to detect bubbles in data with bubble episodes separated by a number of years. This strengthens the argument for using the modified Bhargava statistic for data containing multiple bubbles lasting between 1 and 2 years, such as the ones we observe in the commodity price data. The figure below was generated from the sample corresponding to $\pi = 0.95$ and $b = 2.0$. As we can see, there are between 2 and 3 significant deviations from the random walk process which is in line with what we observe for most commodities.
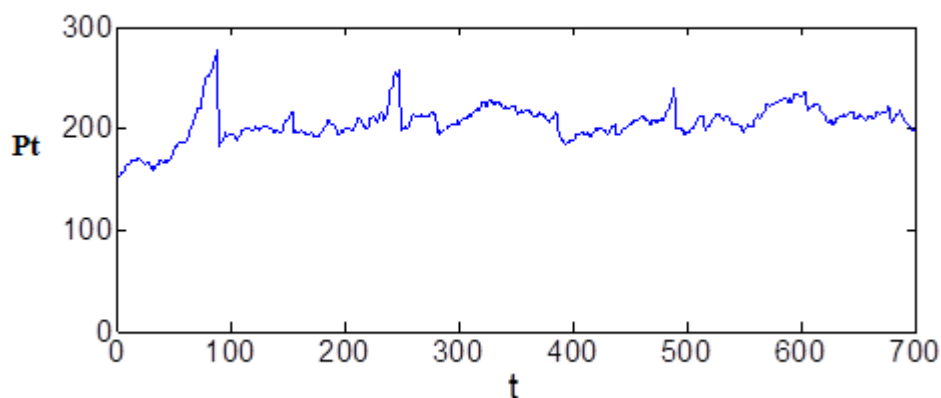
The above analysis uses an exogenous, fixed probability value for the binomial random variable which triggers the bubble. A more theoretically appealing alternative would make the probability endogenous to the data generating process. In a rational bubble framework, rational investors will keep investing in the market expecting the bubble to keep growing; however, the possibility of the bubble collapsing is always present. The further away the prevailing price moves from the

fundamentals, the more likely the collapse becomes; thus, the probability of a bubble continuing from one period to the next will be endogenous to the data generating process.

Taking the above analysis into account, we evaluate the modified Bhargava statistic's power using endogenous probabilities. Specifically, the parameter for the binomial random variable determining continuation of the bubble is determined as follows:

$$\pi = 2\left(\frac{1}{1 + (b_t)^a}\right)$$

As the size of the bubble increases, the probability of the bubble continuing goes down. The scalar *a* controls how quickly this probability goes down. Multiplying the expression by 2 ensures that when the bubble size increases to 1 unit, the probability of the bubble continuing is close to 1. We use the same data generating process as (2.18) and (2.19) in order to test the power of our statistic using this endogenous probability. Parameter values are identical to our initial power test. The scalar 'a' is initially set at 0.1. Numerically, this is equivalent to having a probability around 0.99 when the bubble size is slightly greater than 1.

Simulated prices look fairly similar to those generated by the original data generating process. The power of the test given endogenous probability is 70% which compares favourably with our previous results. Thus, even with endogenous probabilities the test has significant power to detect bubbles that last for periods of 1 or more years. Again, if we use shorter sub-samples the power will be much higher. As we decrease the scalar 'a' further, the power of the test goes up further. The following diagram shows a typical simulated price series using an endogenous probability.
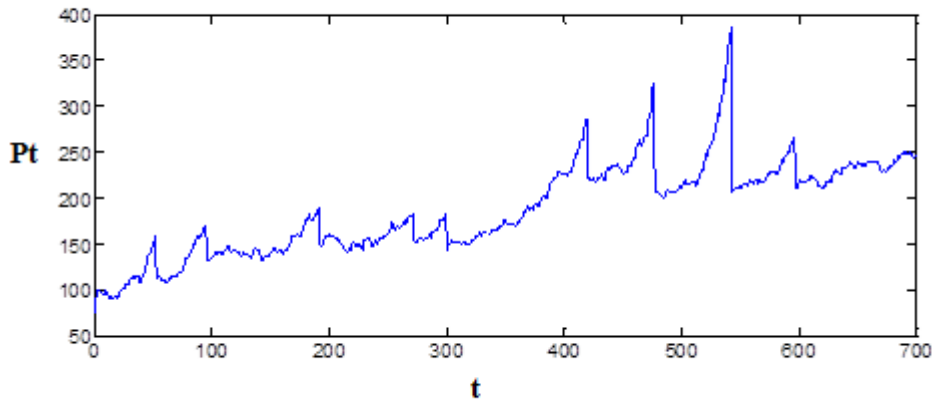
Finally, we compare the small sample properties of both the PSY and modified Bhargava statistics. For this purpose we used the same data generating process as in (2.18)-(2.20) but instead of generating a sample size of 700, we generate a sample of 100 observations. Next we test for bubbles

using the modified Bhargava statistic with a sub-sample size of 20 and an overlap window of 10. Note that Philips et al have already performed the test for 100 observations in their 2012 article where they use an initial window of 40 observations. 5000 such simulations were performed. Philips et al find that the power of their GSADF statistic for a sample size of 100 is 55.6% (as reported in their article). The modified Bhargava statistic on the other hand has a power of 63%. If we decrease the overlap window to 5, the power increases further to 78%. Thus, as the sample size decreases the modified Bhargava statistic starts to outperform the PSY test.

The analysis above shows that the modified Bhargava statistic is a powerful tool for detecting bubbles and identifying periods where bubbles may be present. The GSADF statistic has the advantage that it can date bubbles more precisely; however, it is computationally more demanding and is less accurate in smaller samples. It should also be mentioned that the GSADF statistic uses the Augmented Dickey Fuller test which has an asymptotic distribution so it may not be fully reliable with limited data. The Bhargava statistic on the other hand has an exact distribution as shown by Bhargava so the test can be applied to small samples as well. This will be particularly important when one wants to determine if a continuing series is in a bubble contemporaneously. We note further shortcomings of the GSADF statistics in Chapter 4.

As we have shown above, the Bhargava statistic is more flexible and by changing the sub-sample size and the overlap window we can modify the procedure to suit the question we are trying to answer. For historical analysis and large samples, a sub-sample size of 100 will be more appropriate and the GSADF and Bhargava statistics should yield the same result. If on the other hand our interest lies in detecting whether the series is in a bubble currently, we may rely on a smaller sub-sample size, such as the size 20 we apply above. The modified Bhargava procedure will work much better at the end of a sample compared to the GSADF statistic as is evident from our power test with smaller samples.

Another advantage is that the modified Bhargava test can be used for any elliptical distributions without changing the critical values significantly. Thus, if we change the distribution of the error term in the data generating process above from a normal to a student's t-distribution, our critical values will not change significantly implying that application of the test will be robust to the distribution of the error term. The test does not rely on estimating any parameter values, which is another benefit it enjoys of the PSY statistic which requires estimation of the Auto-regressive parameter using OLS which can be biased in small samples.

We have argued above that the modified Bhargava statistic is at least as good at detecting bubbles as the GSADF statistic. While the GSADSF statistic may be able to date historical bubbles more precisely, the modified Bhargava procedure will be much better at detecting bubbles towards the end of a sample and for smaller samples. The new strategy continues performing adequately even with relatively large samples. The modified Bhargava procedure is also more flexible and has an exact

distribution thereby being more reliable econometrically. Next, we employ the modified Bhargava statistic to detect bubbles in commodity price data. For comparison we also performed the GSADF test on the same data; the results for GSADF testing are available upon request.

2.5   Commodity price series and Data:

Given the turbulence we observed in financial markets from 2007-2009, economists have increasingly started considering other asset classes as candidates for bubble testing. For our empirical application, we consider 25 commodity price series. Being used as inputs in manufacturing and energy sectors, commodity prices can have repercussions for the real economy which makes them interesting. In what follows we provide a brief discussion of our data.

The International Monetary Fund's (IMF) International Financial Statistics database was used to access data on commodity prices. The database contains monthly data on a number of commodities starting from 1957. Not all commodity price series begin from 1957 and in fact a number of monthly series start from 1971 during the first OPEC oil crisis period. All price series are in nominal US dollars. Each price series represents an average spot price calculated for the period in question.

The price series for a number of metal commodities such as Aluminium, Nickel and Iron Ore initially start as yearly series. Contract negotiations for these commodities were carried out on a yearly basis so the spot price stayed fixed throughout the year. Thus, contracts were undertaken a year in advance which fixed the average price for the importer for a year even though the spot price at a particular location was more likely to vary due to transportation costs and inventory considerations. Gradually with more liquidity in the market and the development of financial and transportation sectors, more regular contract negotiations were possible so that prices were determined on a quarterly and eventually on a monthly basis. Currently, pricing data is available on an intra-day basis and thus, there is much more variability seen in commodity price series for recent years. The one exception to this general trend is Iron Ore.

Up to the late 2000's, iron ore pricing took place in private negotiations between iron ore miners and Japanese shipping companies who were the primary consumers of raw iron ore. As demand for iron ore in China grew and with the emergence of swap contracts for refined Iron Ore, the pricing mechanism changed to a more regular monthly frequency. The swap contract itself is priced on a regular basis (Mining Journal). Annual contracting reduces the potential of volatility and thereby may also preclude the existence of a bubble.

In addition to individual price series, data on various price indices, prepared by the IMF, was also used to get a broader picture across different categories. Group indices are constructed based on individual price indices; the weights used are based on the trade value of individual commodities

compared to world commodity trade values. The weights are updated every 5 years and 2005 is used as the base year to index all commodities. Group commodity indices are likely to be influenced by aggregation effects but nevertheless may provide useful information when it comes to understanding particular bubble episodes. In most cases the results for these indices corroborate our results for individual commodities or the relevant commodity group.

Table 2.1 shows the commodities used for the tests, including the time period for which monthly data were available and the group they have been included in for construction of commodity price indices. Precious metals are not the focus of this article and have been included for comparison purposes only (they have thus been included in the Miscellaneous group along with price indices).

Preliminary analysis of the price series' in levels shows most commodity prices going up around the time of the first and second oil price shocks. This is followed by a period of stability in the 90's and the early 2000's. From 2005 onwards, commodity prices start rising again until 2008 when most commodities reach a peak which is fuelled in part by increasing demand from developing countries but may also be due to the presence of bubble elements. We would expect our test to detect explosiveness in these periods. It should be noted that the real values of commodities have been decreasing over time as technology has improved productivity in agriculture and mining. We apply our tests to nominal prices; in Chapter 5 we investigate explosive episodes in real prices.

Following the financial crisis there is a marked decline in most commodity prices; however, the decline is not uniform and some commodity prices have risen substantially since the commodity price crash in 2009. We aim to provide some context to the explosive episodes we detect to aid the analysis. Most explosive periods tend to be triggered by, what might be termed, a productivity shock. Following this shock commodity prices tend to exhibit the type of explosive behaviour that is akin to the kind of bubble econometric tests are designed to detect.

2.6   Results

We only report the results of the modified Bhargava test below but do note the results obtained from the GSADF statistic for comparison purposes. As mentioned before, the GSADF testing results are available upon request.

For the purposes of our test and given the length of the commodity price series we conducted the test with a sub-sample size of 100 and an overlapping window of 50. We believe that a sub-sample size of 100 will make the analysis more tractable and will allow us to detect all major episodes of explosive behaviour. Since the purpose of this exercise is to compare the explosive periods detected in the Bhargava and the GSADF tests, a sub-sample of 100 will suffice. Whereas having a shorter sub-sample will detect more bubbles, it will make the Bhargava statistic series for each commodity longer

and make the comparison much more difficult. However, we will still advocate that a researcher or investor, primarily using this procedure to make economic or financial decisions, use a smaller sub-sample.

In the commodities under consideration, bubbles have different durations in each series, although most periods exhibiting explosive behaviour do not exceed a year. We take each commodity in turn and apply the Bhargava statistic with a sub-sample size of 100 and a rolling window of 50. Table 2.5 shows how the 100 month windows are constructed including the overlap period. Data for all commodity series does not start from January 1957; hence, the number of times the statistic is calculated is shorter for such series. To give an illustrative example, Window 1 refers to the time period January 1957- May 1965. Window 2 starts from March 1961 and goes up to July 1969. The period from March 1961 – May 1965 is common to both Windows and is referred to as the overlap period. Table 2.6 shows the calculated value of the statistic. The critical values are obtained from table 2.4 and correspond to a sub-sample size of 100.

We fitted a random walk with drift model to each series to identify if the series' trend component was significant. The STAMP software was used for this purpose which uses the Kalman Filter in order to estimate a model with unobserved components. Table 2.6 shows the results of our tests and also lists whether each series has a significant trend component. All commodity series had a statistically significant trend component with the exception of Tea, Coal, Gasoline and the Beverages index. Although the trends were statistically significant, their magnitude was small. Thus, the relevant critical value to use is the one with trend. It should be noted that the results do not vary greatly even if we employ the statistic without the trend and use de-trended critical values.

Table 2.6A shows the Bhargava statistic values for Food commodities. For a majority of the commodities, the statistic is able to detect an explosive root in periods where bubble like behaviour was observed. Particularly important are Windows 3, 4, 11 and 12 which correspond to the oil crises and the financial crisis period respectively. Additionally, the statistic is also able to capture instances of explosive behaviour in individual series. For instance, the increase in coffee prices between 1994 and 1996 results in the random walk hypothesis being rejected for coffee (window 8). A severe frost in Brazil in 1994 resulted in coffee prices considerably increasing giving rise to this period of explosiveness. All other food commodities on the other hand were stable during the period and so the random walk hypothesis was not rejected during the period.

If we interpret rejection of the hypothesis at the 5% level as providing stronger evidence for bubble like behaviour than rejection at the 10% level, we note that for periods which coincide with the highest prices, the null hypothesis is rejected at the 5% level while periods containing a build up to the bubble period are rejected at the 10% level. The first major episode of bubble like behaviour occurs

during the first oil price crisis. In the period between 1965 and 1977 (windows 3 and 4), the statistic detects explosiveness in Cocoa beans, Coffee, Lamb, Palm oil and Wheat.

Bubbles are also detected for the Lamb series between 1982 and 1990 and for Wheat in the windows corresponding with the 1980s. During the financial crisis (Window 11 and 12: 1998-2011) we note a spur of activity in food commodities and explosive roots are detected in Barley, Coffee, Palm Oil and Sugar. Although the random walk hypothesis is not rejected in this particular window for the remaining commodities, the value of the statistic is fairly close to the rejection region which could indicate that the bubble like activity has had an impact. It also highlights the effect of the crash that followed the bubble. Since the price of a number of food commodities reduced considerably after the crises, the prices came close to their starting values in the sub-sample which reduces the value of the denominator, inflating the statistic.

Metals and energy commodities follow a similar pattern. Bubbles are detected in a majority of metal and energy commodities during the oil crises as well as during the financial crisis. The strongest evidence for bubble behaviour is present during the 2nd oil price crisis between 1973 and 1982 (corresponding to windows 3, 4 and 5), when metal prices soared above their historic values. 1982 saw oil prices collapse and we see the value of the statistic reducing. An isolated incidence of exuberant behaviour is detected in Tin prices in the period between 1977 and 1986. This was the period when the Tin price was being controlled by the International Tin Council, a consortium of Tin producing countries. Their agreement collapsed in 1985 leading to the fall of tin prices from over £10,000 to around £3,500.

The 1998-2007window sees the most significant evidence of bubble like behaviour. The null hypothesis is rejected for all metal commodities except Copper. Thus, at the height of the financial crisis when metals recorded their peak nominal values, the Bhargava statistic is able to correctly reject the hypothesis of efficiency or prices following a random walk. The subsequent windows show very limited activity, even though these sub-samples also contain the bubble. As discussed before, this indicates the sensitivity of the statistic to the sub-sample. The statistic is better able to detect bubbles when they are either in the formation stage or when they are close to their peaks.

Evidence for energy commodities on the other hand is mixed. While the UK Brent series is seen to exhibit bubbles in both the oil crises as well as the financial crisis, the statistic cannot be rejected for Coal, Gasoline, nor West Texas oil during the financial crisis even though Coal, Gasoline and Natural Gas all peaked during the financial crisis. On the other hand, bubbles are detected during both the first and the second oil price shock in the Petrol series, with rejections at the 5% significance level. Energy prices have rebounded following the crash and oil has crossed the $100 a barrel threshold. The impact of this rebound is captured well by the Bhargava statistic and the statistic detects bubbles in West Texas petrol series between 2007 and 2014.

Raw material commodities have had more instances of commodity specific shocks. With the exception of the financial crisis in 2007, raw material commodities follow explosive paths in periods coinciding with supply or demand shocks for each commodity. It is possible that speculative activity during these instances drove up the price of raw materials above the value that would otherwise have been dictated by shocks to fundamentals. The statistic captures the Cotton price shock in the 80's, but fails to detect the Jute price shock in the mid 1980's. There is an additional bubble detected in cotton prices between 2002 and 2011. This is primarily indicative of the increase in cotton prices towards the end of this sub-sample when there was panic buying in the cotton futures market.

Indices follow a similar pattern to the group they represent with food, energy and metals exhibiting bubble like behaviour during the oil crises and the financial crisis while raw materials and beverages do not show significant evidence of bubble like behaviour.

This shows that the modified Bhargava statistic is very good at detecting explosive behaviour even with a high sub-sample size. A smaller sub-sample may be used by investors as a warning mechanism to highlight a potential bubble arising. We also note that detection of explosive behaviour is dependent on the period being considered. Using a case study approach for the detection of specific episodes of explosive behaviour may be more prudent and accurate for the concerned individual. As mentioned in Section 2.3.2, within a sub-sample Evans' criticism will still hold. Thus, the ideal way to apply this test will be to divide the full series into regions of explosiveness and collapses, and then run the modified Bhargava algorithm we have suggested. Our analysis of commodities reveals that bubbles are more likely to be detected when the explosive period occurs towards the end of the sub-sample.

Using the same windows for all commodities also allows us to see the number of coincident bubbles in each period. As suspected, the incidence of bubbles is the highest during the two oil crises and the financial crisis. During the first oil crisis, which corresponds to the 1969-1977 window we observe that 13 out of 25 commodities were in a bubble. Similarly 11 out of 25 commodities are in an explosive state between 1973 and 1982, roughly the period around the second oil crisis. The financial crisis on the other hand had the highest incidence with 15 out of 25 commodities being in a bubble during both the 1998-2007 and 2002-2011 windows. The explosive behaviour had started emerging around 2005 which explains why we observe such a high number during the 1998-2007 window as well. Thus, what may have started as a shock to fundamentals in a particular commodity may have spilled over to other commodities resulting in bubbles in other commodities. A more thorough investigation would warrant a case study approach as we have stated before. The bar-chart below shows the number of commodity bubbles the modified Bhargava statistic detects in each window.

Number of Bubbles in each time period across all 25 commodities
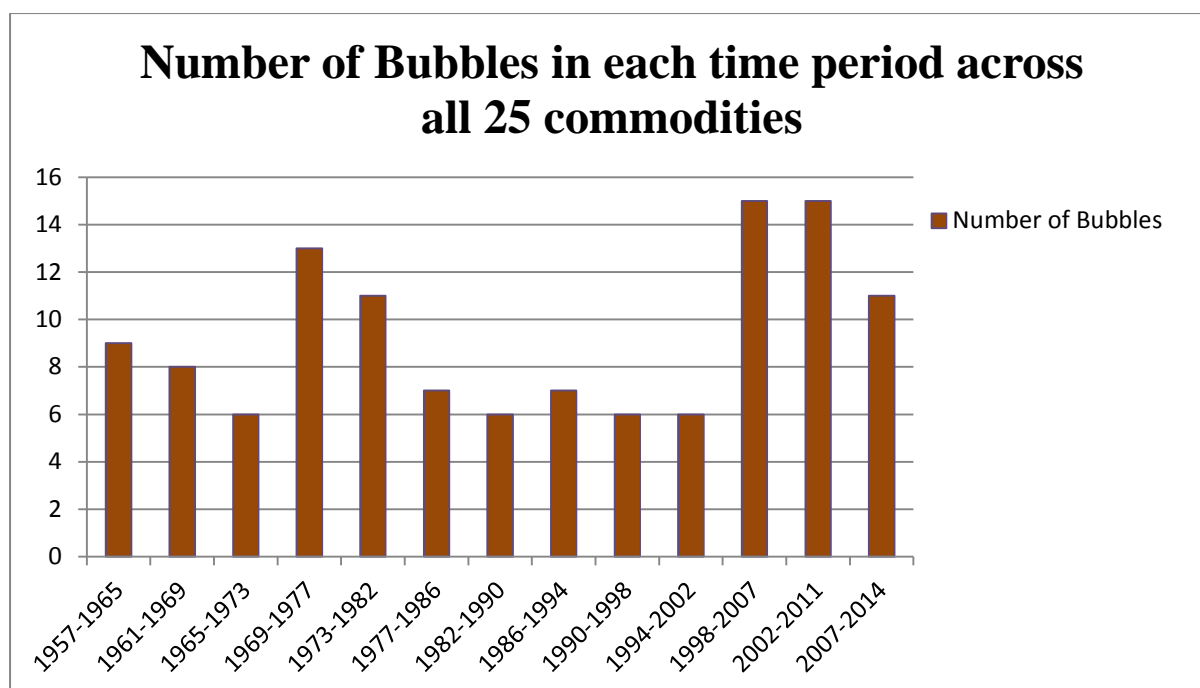
Thus, our modified Bhargava statistic can prove to be particularly useful for emerging bubbles which will concern policy makers and financial managers alike. Applying the Bhargava statistic only once on the whole sample would have led to the rejection of the bubble hypothesis due to the existence of multiple bubbles. Our methodology allows us to detect multiple bubbles and the rolling window has enabled us to understand how best to use the Bhargava statistic in applied work. For more research oriented work, this flexibility will also be able to act as a robustness check.

For each commodity series we also performed the GSADF test to check if the GSADF value indicated the presence of bubbles in the series. Using the SADF sequence of values obtained from the procedure described in section 2.3.3, we time stamped the bubble episodes. The test statistic and date stamps for each commodity are contained in table 2.7 below. The columns indicate the dates for each bubble episode. Very few series had more than 3 bubble episodes so we have reduced the table to include only the 3 most significant bubbles, in chronological order. Bubbles are detected at the 1% level of significance for all commodity series with the exception of Poultry and Gasoline for which the random walk hypothesis is only rejected at the 5% level of significance. Our results were broadly in line with the modified Bhargava statistic and we note that bubbles detected for each commodity are in similar periods.

2.7   Conclusion

We set out by describing a new methodology for using the Bhargava statistic to detect explosive bubbles in asset prices. We have shown that by dividing a series into shorter sub-samples and using a rolling window, we are able to not only detect bubbles but can also identify periods in which they

occur. This addresses Evans' criticism of detecting multiple collapsing bubbles in a price series. Within the context of multiple collapsing bubbles, the new test also has much higher power than the original Bhargava statistic. Additionally, we have also compared the test against the GSADF test devised by Philips, Shi and Yu. While the PSY statistic is able to date bubbles better, the power of both statistics is high. We have also shown that the modified Bhargava test procedure performs better and attains higher power when the sample size is smaller.

In addition, the modified Bhargava statistic has the advantage of having an exact distribution instead of an asymptotic distribution. The modified Bhargava statistic does not require the computation of any parameter while the PSY statistic is based on the calculation of an auto-regressive parameter. For small samples, the modified Bhargava test will be more appropriate given its design as a local test; thus, the Bhargava statistic may be used even if the sample size is as low as 50 or 100 observations while the properties of the PSY statistic may not hold for smaller samples. The Bhargava test also takes less computational time than the PSY statistic.

The modified Bhargava statistic offers a lot of flexibility as the size of the sub-sample and the overlapping window can be varied. For historical analysis we recommend using a larger sub-sample size; 50 or 100 with an overlapping window that is half the size of the sub-sample. To assess whether a particular asset is currently in a bubble, our recommendation will be to use more recent data and use a smaller sub-sample with the modified Bhargava statistic. The main advantage of using the PSY test is that it is able to date bubbles more accurately.

Our second contribution is the empirical application of the modified Bhargava statistic and the PSY statistic to commodity price data. This helps us in two ways. Firstly, it helps us gain an understanding of explosive episodes in commodity prices and secondly, it allows us to compare the results obtained from both statistics. Critical values for both tests were evaluated using simulated data for price series that are 687 observations in length.

Both PSY and modified Bhargava statistic were able to identify all well-known instances of explosiveness in commodity prices. Bubbles occurred in most of the commodity series under consideration and a number of them had multiple bubble episodes. The first oil price shock and the 2007-08 financial crisis saw the most bubble instances. The 1990's and the early 2000's were relatively calmer periods with only sporadic bubble episodes. Using the modified Bhargava statistic, for the 25 commodities used, no bubbles were detected for Poultry, Rice, Tea, Coal and Gasoline at the 5% level of significance. For the remaining 20 commodities we detected 41 sub-samples where bubble behaviour is present at the 5% level. The PSY statistic on the other hand detected significant evidence of explosive behaviour in all commodities except Poultry.

We also note that the duration of most bubble episodes is short and often less than 1 year which implies that detecting bubbles with quarterly or yearly data can be more difficult as the data may not contain the variability required for accurate detection of bubble episodes. In yearly or quarterly data, bubbles would collapse within the span of 1 or 2 periods making detection more difficult. For food, metal and energy commodities, we also note that bubble episodes tend to cluster together and bubbles tend to occur for whole groups of commodities. In most instances the results from the Bhargava statistic tended to corroborate those of the PSY statistic.

Our modification to the Bhargava statistic has provided a useful alternative to bubble detection in asset prices. While we have applied the test to commodities, we believe it will work just as well with other asset classes. This Chapter also motivates the use of switching regime regression to estimate asset prices. The period that we have referred to as a bubble can be thought of as a deviation from a random walk for a certain period of time; we can detect such periods using the tests outlined in this chapter. In fact the following two chapters will build upon the notion of treating explosiveness as one possible state of nature.

APPENDIX 2 – Tables and figures

Table 2.1: Commodity price series

Commodity	Time Period	Units	Commodity Group
Barley	Jan 1975 - Mar 2014	$ per metric tonne	Food
Cocoa beans	Jan 1957 - Mar 2014	$ per metric tonne	Food
Coffee	Jan 1957 - Mar 2014	cents per pound	Food
Lamb	Jan 1957 - Mar 2014	cents per pound	Food
Palm oil	Jan 1957 - Mar 2014	$ per metric tonne	Food
Poultry	Jan 1980 - Mar 2014	cents per pound	Food
Rice	Jan 1957 - Mar 2014	$ per metric tonne	Food
Sugar	Jan 1957 - Mar 2014	cents per pound	Food
Tea	Jan 1957 - Mar 2014	cents per KG	Food
Wheat	Jan 1957 - Mar 2014	$ per metric tonne	Food
Aluminum	Jan 1957 - Mar 2014	$ per metric tonne	Metals
Copper	Jan 1957 - Mar 2014	$ per metric tonne	Metals
Iron Ore	Jan 1957 - Mar 2014	$ per metric tonne	Metals
Lead	Jan 1957 - Mar 2014	$ per metric tonne	Metals
Nickel	Jan 1957 - Mar 2014	$ per metric tonne	Metals
Tin	Jan 1964 - Mar 2014	$ per metric tonne	Metals
Zinc	Jan 1957 - Mar 2014	$ per metric tonne	Metals
Coal	Jan 1979 - Mar 2014	$ per metric tonne	Energy
Gasoline	Jan 1979 - Mar 2014	cents per gallon	Energy
Natural gas	Jan 1985 - Mar 2014	$(000s) per million BTU	Energy
Petroleum UK Brent	Jan 1957 - Mar 2014	$ per barrel	Energy
Petroleum West Texas	Jan 1959 - Mar 2014	$ per barrel	Energy
Cotton	Jan 1957 - Mar 2014	cents per pound	Raw Materials
Jute	Jan 1957 - Mar 2014	$ per metric tonne	Raw Materials
Rubber	Jan 1957 - Mar 2014	cents per pound	Raw Materials
Gold	Jan 1964 - Mar 2014	$ per Troy ounce	Misc
Silver	Jan 1969 - Mar 2014	Unspecified units	Misc

Source: International Monetary Fund, International Financial Statistics (May 2014)

Table 2.2

Test	Power
PSY ($r_w = 100, n = 700$)	97%
Bhargava (sub-sample = 100, overlap = 50, n = 700)	39%
Bhargava (sub-sample = 50, overlap = 25, n = 700)	86%
Bhargava (sub-sample = 20, overlap = 10, n = 700)	99%

Table 2.3 - Critical Values for PSY Test	
Confidence Interval	Critical Value
90%	1.812
95%	2.082
99%	2.547

The critical values through simulation of the following data generating process:

$$y_t = dT^{-\eta} + \theta y_{t-1} + \varepsilon_t,$$

with d = $\eta = \theta = 1$ and $\varepsilon_t \sim N(0,1)$. 5,000 replication were performed

Table 2.4: Simulated Critical Values for the Bhargava Statistic

	Without Drift			**With Drift**		
Critical Values	**10%**	**5%**	**1%**	**10%**	**5%**	**1%**
Sample Size						
n = 20	0.041	0.032	0.021	0.149	0.113	0.078
n = 30	0.029	0.021	0.013	0.101	0.078	0.05
n = 50	0.017	0.012	0.008	0.057	0.043	0.028
n = 75	0.011	0.008	0.005	0.038	0.029	0.019
n = 100	0.008	0.006	0.004	0.029	0.022	0.014
n = 200	0.004	0.003	0.002	0.015	0.011	0.007
n = 700	0.001	0.001	0.000	0.004	0.003	0.002

The critical values were obtained using 5000 replications. The data generating process is a random walk with and without drift

Table 2.5

Window Number	Time Period Covered	Overlap period with previous window
1	Jan 1957 - May 1965	N/A
2	Mar 1961 - Jul 1969	Mar 1961 - May 1965
3	May 1965 - Sep 1973	May 1965 - Jul 1969
4	Jul 1969 - Nov 1977	Jul 1969 - Sep 1973
5	Sep 1973 - Jan 1982	Sep 1973 - Nov 1977
6	Nov 1977 - Mar 1986	Nov 1977 - Jan 1982
7	Jan 1982 - May 1990	Jan 1982 - Mar 1986
8	Mar 1986 - Jul 1994	Mar 1986 - May 1990
9	May 1990 - Sep 1998	May 1990 - Jul 1994
10	Jul 1994 - Nov 2002	Jul 1994 - Sep 1998
11	Sep 1998 - Jan 2007	Sep 1998 - Nov 2002
12	Nov 2002 - Mar 2011	Nov 2002 - Jan 2007
13	Jan 2007 - Mar 2014	Jan 2007 - Mar 2011

Table 2.6A - Bhargava Statistic for Food commodities

Trend	Series	1	2	3	4	5	6	7	8	9	10	11	12	13
Yes	Barley	N/a	N/a	N/a	N/a	0.2209	0.1667	0.1752	0.0584	0.0344	0.0302	0.0159*	0.0824	0.0416
Yes	Cocoa beans	0.0263**	0.0382	0.0359	0.0216*	0.0305	0.0799	0.0961	0.0294**	0.0578	0.0301	0.0531	0.0539	0.0627
Yes	coffee	0.0088*	0.0556	0.0137*	0.0593	0.0518	0.0284**	0.1229	0.0085*	0.0792	0.1061	0.0322	0.0187*	0.061
Yes	Lamb	0.085	0.4021	0.0166*	0.125	0.1352	0.0626	0.0517*	0.2201	0.0696	0.0558	0.1389	0.1781	0.0533
Yes	Palm oil	0.0077*	0.0421	0.0249**	0.0909	0.1026	0.0359	0.1076	0.0992	0.168	0.058	0.0124*	0.0538	0.0788
Yes	Poultry	N/a	N/a	N/a	N/a	N/a	0.1206	0.0654	0.0849	0.0317	0.0684	0.0245**	0.0319	
Yes	Rice	N/a	N/a	N/a	N/a	0.0936	0.2364	0.1549	0.4124	0.2255	0.2458	0.0301	0.2376	0.1672
No	Sugar	0.3662	0.1031	0.0584	0.041	0.1609	0.0586	0.5736	0.2618	0.0922	0.0788	0.0431	0.0148*	0.469
No	Tea	0.2797	0.2509	0.4657	0.2496	0.1371	0.6867	0.1111	0.3691	0.2561	0.1182	0.1868	0.4491	0.063
Yes	Wheat	0.1067	0.0976	0.0153*	0.0603	0.0453	0.0213*	0.0339	0.0892	0.0445	0.0469	0.1874	0.0902	0.1288

Table 2.6B - Bhargava Statistic Metals

Trend	Series	1	2	3	4	5	6	7	8	9	10	11	12	13
Yes	Aluminum	0.0266**	0.0321	0.1122	0.0179*	0.0196*	0.0512	0.1087	0.084	0.0964	0.1421	0.0249**	0.0704	0.1118
Yes	Copper	0.0189*	0.1718	0.0551	0.089	0.1029	0.0546	0.0711	0.0653	0.0774	0.1088	0.0351	0.0999	0.0841
Yes	Iron Ore	0.0359	0.0828	0.0454	0.0958	0.0377	0.0527	0.0177*	0.0336	0.0556	0.0612	0.0253**	0.0294	0.059
Yes	Lead	0.0119*	0.058	0.0152*	0.0833	0.0476	0.0376	0.043	0.0695	0.0362	0.0763	0.0142*	0.1228	0.1267
Yes	Nickel	0.0938	0.0201*	0.1404	0.2042	0.0516	0.0569	0.0850	0.0504	0.0830	0.1010	0.0111*	0.0691	0.0789
Yes	Tin	0.0197*	0.0539	0.0076*	0.0229**	0.0764	0.0281*	0.0538	0.1101	0.1668	0.0746	0.0177*	0.032	0.1125
Yes	Zinc	0.0166*	0.0495	0.0050*	0.0456	0.1031	0.0359	0.0377	0.0294	0.0347	0.0669	0.0112*	0.044	0.0278**

Table 2.6C - Bhargava Statistic Energy

Trend	Series	1	2	3	4	5	6	7	8	9	10	11	12	13
No	Coal	N/a	N/a	N/a	N/a	0.0149	0.0217	0.0193	0.0515	0.0123	0.0215	0.1028	0.1381	0.5888
No	Gasoline	N/a	N/a	N/a	N/a	0.0343	0.1089	0.2645	0.0872	0.1963	0.1167	0.2333	0.1697	0.919
Yes	Natural gas	N/a	N/a	N/a	N/a	N/a	N/a	N/a	0.0381	0.0798	0.0377	0.0168*	0.0515	0.0775
Yes	Petroleum UK Brent	0.1607	0.0392	0.0162*	0.2378	0.0922	0.0112*	0.0857	0.1684	0.1271	0.0952	0.0151*	0.0194*	0.1000
Yes	Petroleum West Texas	0.0298	0.0418	0.0104*	0.0156*	0.1017	0.0939	0.1838	0.0225*	0.0389	0.041	0.1376	0.0796	0.0134*

Table 2.6D - Bhargava Statistic Raw Materials

Trend	Series	1	2	3	4	5	6	7	8	9	10	11	12	13
Yes	Cotton	0.0536	0.0652	0.0055*	0.0450	0.0426	0.0125*	0.0407	0.0647	0.0622	0.0999	0.0590	0.0059*	0.0774
Yes	Jute	0.1191	0.0184*	0.2011	0.0993	0.0638	0.0566	0.0543	0.1037	0.0785	0.0658	0.1532	0.1479	0.0938
Yes	Rubber	0.1298	0.0238**	0.0177*	0.0604	0.0446	0.0312	0.0775	0.0332	0.0228**	0.0477	0.0572	0.0333	0.0490

Table 2.6E

Trend	Series	1	2	3	4	5	6	7	8	9	10	11	12	13
Yes	Gold	N/a	N/a	0.0140*	0.0374	0.0212*	0.0425	0.0182*	0.0416	0.1200	0.0383	0.0163*	0.0372	0.0218*
Yes	Silver	N/a	N/a	N/a	0.0854	0.1879	0.1253	0.1858	0.0988	0.0439	0.3069	0.0214*	0.1293	0.0503
Yes	All commodities index	N/a	N/a	N/a	N/a	N/a	N/a	N/a	N/a	N/a	0.0467	0.0226**	0.0321	0.0991
Yes	Agriculture index	0.0935	0.1708	0.0064*	0.0593	0.0488	0.0426	0.0607	0.0518	0.0380	0.0838	0.0792	0.0129*	0.0684
No	Beverages index	0.0709	0.2160	0.0335	0.0274	0.0236	0.0304	0.0504	0.0081**	0.0981	0.0456	0.0167	0.0306	0.0576
Yes	Energy Index	N/a	N/a	N/a	N/a	N/a	N/a	N/a	N/a	0.0355	0.0471	0.0384	0.0181*	0.0655
Yes	Food Index	0.0236**	0.1447	0.0095*	0.0375	0.0901	0.0199*	0.0781	0.1090	0.0487	0.0527	0.0385	0.0593	0.0970
Yes	Metals Index	0.0096*	0.1271	0.0235**	0.0721	0.0321	0.0394	0.0779	0.0438	0.0655	0.0979	0.0099*	0.0771	0.0793
Yes	Non fuel index	N/a	N/a	N/a	N/a	N/a	N/a	0.0195*	0.0331	0.0372	0.0180*	0.0071*	0.0971	0.0534

*Table 2.6 – Results of the Bhargava Statistic calculated between 1957 and 2014. Index numbers denote windows described in Table 3. * - rejection of the random walk in favor of non-stationarity at the 5% level. ** - rejection of the random walk in favor of non-stationarity at the 10% level of significance. 5% significance level with trend is 0.022 and without trend is 0.006. 10% significant level is 0.0290 with trend and 0.0090 without trend*

Table 2.7: GSADF Statistic

Commodity	GSADF Statistic	Bubble 1	Bubble 2	Bubble 3
Barley	2.8716***	Feb 2008-May 2008	-	-
Cocoa beans	7.1647***	Oct 1973- Nov 1974	Mar 1976 - Jun 1977	-
Coffee	13.3371***	Jan 1976 - Jun 1977	Feb 2011 - Jun 2011	-
Lamb	2.9942***	Apr 1980 - Aug 1980	N/A	-
Palm oil	4.6608***	Jul 1973 - Mar 1974	Dec 2007 - May 2008	-
Poultry	2.1067**	-	-	-
Rice	3.0433***	May 2008 - Jul 2008	-	-
Sugar	6.2118***	Oct 1974 - Feb 1975	Mar 2011 - May 2011	-
Tea	5.3398***	Feb 1977 - May 1977	-	-
Wheat	5.8162***	Dec 1972 - Mar 1974	Aug 2007 - Apr 2008	-
Aluminum	7.0251***	Dec 1974 - Jun 1975	Jan 1977 - May 1980	Apr 2006 - Jun 2006
Copper	8.5244***	Oct 2005 - Nov 2006	-	-
Iron Ore	12.5672***	Jun 2008 - Oct 2011	-	-
Lead	10.1832***	Nov 1973 - May 1974	Feb 1979 - Nov 1979	Nov 2006 - Mar 2008
Nickel	9.3304***	Aug 1974 - Jun 1977	Mar 1988 - Jul 1988	Jul 2006 - Jul 2007
Tin	11.9976***	Oct 1973 - Oct 1974	Sep 1977 - Sep 1980	Feb 2007 - Sep 2008
Zinc	16.8465***	Mar 1973 - Jun 1974	Jan 1989 - Apr 1989	Jan 2006 - Jan 2007
Coal	9.1621***	Apr 2004 - Sep 2004	Oct 2007 - Oct 2008	-
Gasoline	2.4618**	Sep 2005 - Oct 2005	-	-
Natural gas	5.8602***	Nov 2005 - Jan 2007	Dec 2007 - Mar 2009	-
Petroleum UK Brent	6.9214***	Jul 1973 - Mar 1974	Apr 1979 - Aug 1980	Nov 2007 - Sep 2008
Petroleum West Texas	8.6355***	May 1979 - Nov 1980	Nov 2007 - Sep 2008	-
Cotton	17.0883***	Sep 1971 - Mar 1972	Mar 1973 - Apr 1974	Oct 2010 - May 2011
Jute	5.9007***	Sep 1984 - May 1985	Dec 2009 - Jul 2010	-
Rubber	6.6105***	Feb 1995 - May 1995	Feb 2006 - Aug 2006	Nov 2010 - June 2011
Gold	11.4261***	May 1972 - Sep 1973	Jun 1979 - Apr 1980	Oct 2009 - Mar 2013
Silver	10.1277***	Sep 1979 - Mar 1980	Feb 2008 - Jun 2008	Nov 2010 - Oct 2011
All commodities index	6.9956***	Mar 2006 - Sep 2006	Oct 2007 - Sep 2008	N/A
Agriculture index	11.9535***	Dec 1972 - Jun 1974	Feb 1979 - Mar 1980	Feb 2011 - May 2011
Beverages index	17.1833***	Feb 1974 - Jul 1974	Apr 1976 - Nov 1977	-
Energy Index	6.2130***	Nov 2007 - Sep 2008	-	-
Food Index	13.2790***	Nov 1972 - May 1974	Dec 2007 - Sep 2008	-
Metals Index	7.7511***	Jan 1974 - Jul 1974	May 1988 - Jul 1988	Dec 2005 - Dec 2007

The table represents the values of the Philips, Shi and Yu statistic using the GSADF procedure outlined by PSY (2013)   *- null hypothesis rejected at the 10% level; ** - null hypothesis rejected at the 5% level; *** - null hypothesis rejected at the 1% level; N/A – No more bubbles in the sample. Column 2 is the sup GSADF statistic which indicates whether bubbles are present in the full sample.

Figure 2.1 – Commodity Price Series in Levels: Metals(others available upon request)

ALUMINUM (Units: US Dollars per Metric Ton)

COPPER (Units: US Dollars per Metric Ton)

IRON ORE:CHINA (US DOLLARS/DMTU) (Units: US Dollars per Metric

LEAD (Units: US Dollars per Metric Ton)

NICKEL (Units: US Dollars per Metric Ton)

TIN (Units: US Dollars per Metric Ton)

ZINC (Units: US Dollars per Metric Ton)
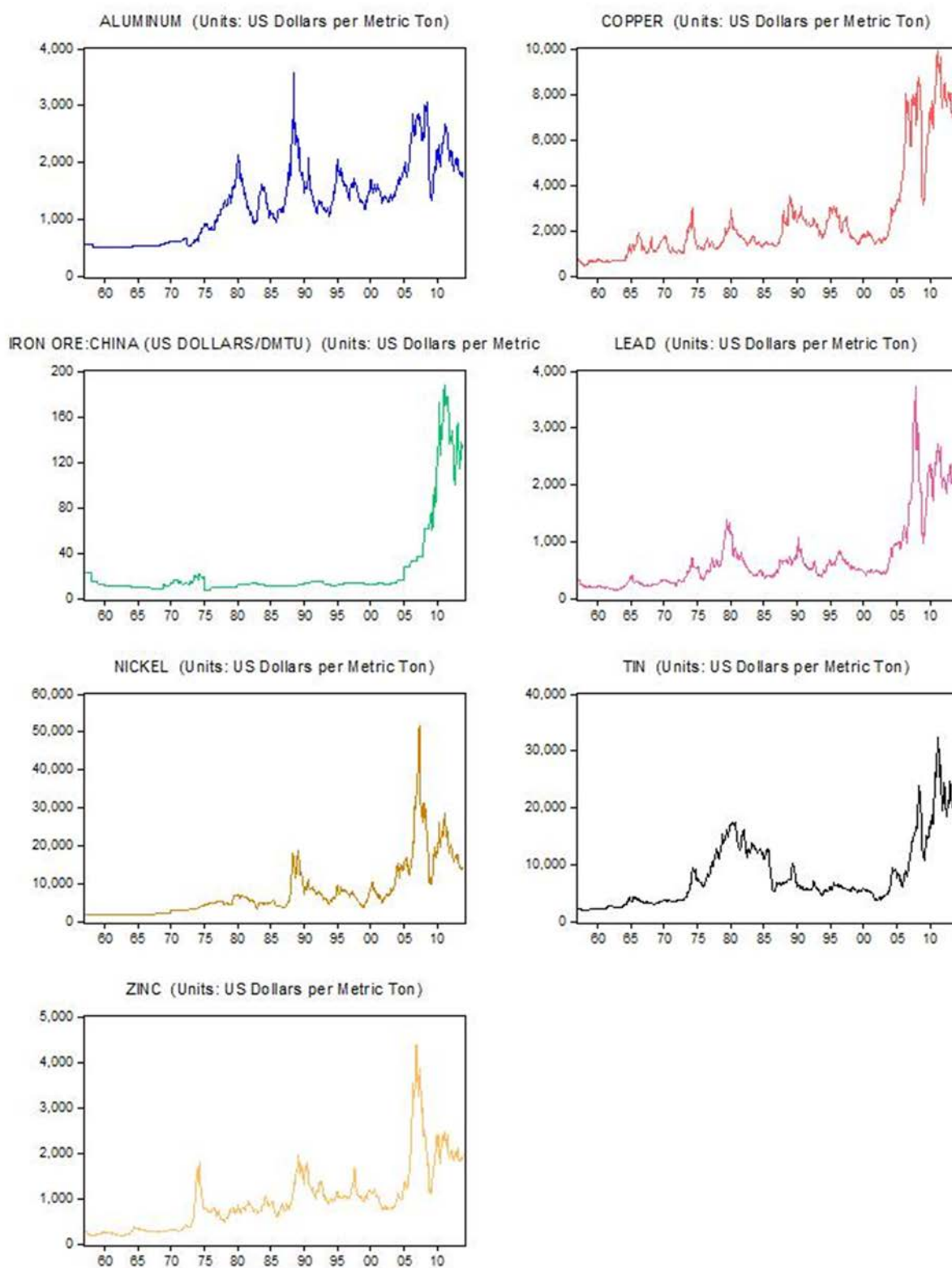
MATLAB code for simulating critical values of the modified Bhargava Statistic

```matlab
count = 1;
simulations = 10000; %number of simulations to be performed
stat = zeros(simulations,1); %this variable stores the statistic value for
each simulation
qe = [0.10 0.05 0.01]; %defined quantiles
window = 700; %size of the full sample
trend = 1;
eta = 1;
samplesize = 100;

distribution = 0;

while count < simulations + 1
    if distribution == 0 %this variable determines whether the statistic is
a Normal or t
        e = randn(window,1);
    elseif distribution == 1
        e = trnd(2,window,1);
    end

y = zeros(window,1);
T = window;
for i = 1:size(y) %the following loop simulates a random walk series
    if i == 1
        y(i) = (trend*(T^(-eta))) + e(i);
    else
        y(i) = (trend*(T^(-eta)))+ y(i-1) + e(i);
    end
end
    %the code below calculates the bhargava statistic as described in the
    %chapter; the statistic calculation is called from a different MATLAB
    %procedure.
    stat(count,1) = min(borigsim2(y,samplesize,samplesize/5,1));
    count = count + 1;

end
    stat;
    quantile(stat,qe)
```

CHAPTER 3: *Understanding commodity markets using Markovian state-switching*

Chapter 3 presents a partial equilibrium model which gives us an estimable reduced form expression for commodities. The form thus obtained can be estimated using Markov Switching Auto-Regressions. We show that estimating this form naively leads to incoherent results. An alternative approach which uses the GSADF statistic to restrict the Markov-switching probabilities improves matters. The suggested methodology is applied to non-ferrous metals. A two-state regression is first used to distinguish between a mean reverting state and an explosive state. The results are extended to a three-state framework which allows us to comment on how efficient commodity markets are. We show that estimating state probabilities directly through the GSADF test leads to better results.

3.1 Introduction:

Anyone familiar with asset pricing literature will be aware of the plethora of approaches that have been adapted to model and understand asset prices. Since Gustafson (1958) first introduced dynamic programming to the rational expectations literature, the field of asset pricing in general and commodity pricing in particular has seen an explosion in interest. Commodity markets add another layer of complexity to the picture. More recently, the behavioural school has taken a keen interest in asset pricing issues (e.g. Sherfin, 2010).

A separate but related strand of literature considers bubbles in asset prices which we have discussed at length in chapter 2. This chapter elicits from the different approaches that have been used previously to come up with a simple yet intuitive form for the price process of a storable commodity. We build on Muth's 1961 framework to formulate a partial equilibrium model for commodity prices. Our model gives a central role to inventories as the presence of inventories differentiates commodities from other asset classes. More specifically inventories can act as both a stabilizing and a destabilizing mechanism in our model which can lead to commodity prices behaving in an explosive manner. Though we start off with a partial equilibrium model, we obtain an intuitive reduced form expression. The reduced form thus obtained is similar to the reduced form employed in bubble literature. In chapter 5 we take a more structural approach but arrive at similar conclusions.

The second and main contribution made in this chapter is with respect to the estimation of the proposed reduced form. Hall et al (1999) suggest that a two-state Markov switching regression will be appropriate for an asset which has a random walk or efficient state and an explosive state. Our analysis contends that using hidden Markov switching regressions without using series specific information leads to estimates that do not help us classify the different states. Thus, we propose an alternative method to estimate Markov switching regressions which relies on estimating transition probabilities directly from unit root tests instead of estimating the transition probabilities from the Markov switching regression.

The GSADF statistic proposed by Philips et al (2013), introduced in chapter 2, is used to date bubbles in commodities. Using the dates obtained from the GSADF test we estimate the state transition probabilities directly. Restricting the Markov switching regression to these transition probabilities leads to better and more intuitive results. The approach helps us clearly distinguish a stationary and a non-stationary state making the interpretation of smoothed probabilities more comprehensible.

We contrast the results obtained through our methodology to results obtained from the unrestricted switching regression. The comparison relies on using smoothed Markov-state probabilities. We note that the unrestricted switching regression often overestimates the explosive or alternative state while the restricted regression assigns high probabilities to the explosive state only during periods when the commodity in question was actually going through an explosive period. To illustrate our procedure we use data on Copper from 1957 to 2014. This procedure is then extended to build a framework that will enable us to evaluate whether a particular asset market is efficient (where efficiency is to be understand as per the description in Chapter 1). We use a 3 state Markov switching regression in order to evaluate the efficiency of Copper and other non-ferrous metals.

The three states we consider are a random walk or efficient state, an explosive or bubble state and a stationary or mean reverting state. Using the GSADF and the its implied counterpart a Generalized infimum Augmented Dickey Fuller Test for mean reversion, we are able to estimate steady state probabilities for each state. Our analysis shows that estimating the steady state probabilities using this method provides better and more reasonable estimates and smoothed-probabilities than if the steady state probabilities were estimated through an unrestricted Markov-switching regression. We find that the market for non-ferrous metals is mostly efficient although there are deviations from efficiency. The duration of these deviations from efficiency vary across the different metals under consideration. While Tin and Zinc stay efficient almost 90% of the time, metals like Nickel are only efficient around 55% of the time; the discussion in section 3.7 expands upon the results. We believe that this approach will help improve our understanding of efficiency in commodity markets better. This can act as a metric for market efficiency. A different metric for market efficiency is developed in chapter 4 using threshold autoregressions.

In Section 3.2 we review the literature on the dynamics of commodity prices which lays the foundation of our commodity price partial equilibrium model in Section 3.3. Section 3.4 outlines the Markov-switching regression methodology and the GSADF test which is used to estimate transition probabilities. Section 3.5 presents a brief overview of Copper over the period under consideration. Section 3.6 discusses results for Copper in detail. Section 3.7 presents results for our efficient markets steady state probabilities and Section 3.8 concludes.

3.2 Modelling the dynamics of commodity prices

The current chapter derives from two different strands of literature, theoretical models of commodity prices and techniques to accurately estimate relationships that govern commodity prices. Each strand continues to be popular with economists. We outline relevant literature for modelling commodity prices. Econometric methods for estimating our model will be discussed in the research and methodology section.

Commodities are fundamentally different from other assets as they can be stored. The market for storage is completely absent from conventional financial assets and this additional market necessitates the use of dynamic models in commodities as storage is primarily used to carry commodities across periods. One of the objectives of coming up with a simple model for commodities is to use the market for storage as an additional facet in explaining temporary explosive behaviour exhibited by commodity prices.

Storage markets serve two primary functions; they can be used for precautionary measures and allow commodity owners the opportunity to meet any unforeseen demand for the commodity or an increase in uncertainty due to supply concerns. Secondly, they can be used as a speculative tool, particularly in the presence of excessive volatility in the financial asset market (not just the market for that particular commodity). Thus, a model incorporating the speculative motive will need to have a dynamic element which warrants the use of dynamic programming. Gustafson (1958) was the first to employ dynamic programming as a methodology to solve a model which assumed rational expectations. Muth (1961) provided a framework for commodity pricing that incorporated rational expectations. Our partial equilibrium model is partly influenced by the approach that Muth introduced. Samuelson's (1971) article is a seminal contribution as it considers uncertainty in output in the context of commodity markets.

While the speculative function can be fulfilled by the derivatives market, the precautionary function is best served by the storage market as having inventory to hand allows storage providers to meet demand immediately should it be required. This additional benefit due to having the commodity to hand is referred to as the convenience yield once the cost of storage has been accounted for. While some authors have propagated the notion that convenience yields act the same way as dividends do

for equities, they do not play a large part in explaining the emergence of bubbles in commodity prices. Kaldor (1939) is credited with coining the term, convenience yield, and for being the first to explicitly use convenience yields in explaining commodity storage. A vast literature has developed over the years using elements from these seminal articles.

Both structural and reduced form models are paramount in the commodity price literature. Structural models incorporate storage as an essential component. Storage acts as a market stabilizing mechanism and controls volatility in commodity markets Within storage models, the supply of the commodity can be inelastic as modelled by Williams et al, (1991) or it can be determined within the model by optimizing agents. Newberry & Stiglitz (1979) in particular have used both risk neutral and risk averse producers in their model. Demand on the other hand has conventionally been assumed in the model as a standard function ($D_t = a - bP_t$); this has primarily been assumed to facilitate analytical solutions as more complex demand functions do not aggregate well.

While most structural models have used discrete time, there are some continuous time structural models as well (e.g. Ribeiro, 2004). Ribeiro provides a continuous time version of a model based on Williams et al's earlier model and derives a numerical solution which emulates pricing behaviour for the commodities in her study. The main benefit of using continuous time models is the ease with which volatility can be incorporated within the model. With discrete time models, this has to be done indirectly through assuming risk averse agents, selecting particular utility functions and then using Taylor approximations. Using Taylor approximations does not provide a sufficient explanation for the presence of bubbles in the market since these bubbles cause a much larger deviation and the error due to a Taylor approximation in such instances would become too large. We look at William et al's structural model in detail in chapter 5 and try to make it more empirically relevant. This chapter contends with models which allow for an analytical reduced form expression for commodity prices.

Reduced form models on the other hand directly model the spot price process in continuous time. A majority of these models tend to be mean-reverting, which is a feature of commodity markets in particular. These are typically modelled as single factor or multi-factor models with the other factors being either the convenience yield or interest rates. Ito calculus is typically employed to simplify these models although analytical results are rare. Ribeiro provides a good summary of reduced form models and also uses 2 factor and 3 factor models to explain commodity behaviour in the presence of storage. Other reduced form methodologies have explicitly modelled long run and short run dynamics of commodity prices (Wets and Rios, 2012). Another approach has seen continuous time volatility models (Brooks et al 2013, 2015) whereby the volatility process is specified as part of a system of continuous time differential equations with the volatility process evolving dynamically. This provides a good foundation on which to build a continuous time reduced form model for commodities that can incorporate bubbles.

Analytical results for both structural as well as reduced form models are few and far between and economists have primarily relied on numerical procedures. Testing for bubbles in commodities often relies on a reduced autoregressive form which does not necessarily follow from the models discussed above. Reduced form models, on the other hand, do not have sufficient theoretical underpinnings which allow us to explain how commodity prices behave when they are in a bubble state. Literature on bubble testing has used reduced form models in discrete time which are quite distinct from the continuous time models mentioned above.

When testing for explosiveness in asset prices, either through Markov-switching regressions ala Hall et al (1999) or tests such as the GSADF test introduced in the previous chapter, researchers rely on the reduced autoregressive form. While the reduced Auto-regressive form can be derived for equity markets, e.g. see Blanchard et al (1982), it is not clear why the same model can be used for testing explosiveness in commodities. As outlined above, commodities are fundamentally different from other assets due to the presence of a storage market. Investor behaviour tends to drive inventory levels which in turn influences price. For instance, speculation in commodity markets can lead to very high inventory levels, leading to a shortage for consumers and resulting in high prices.

Our simple model bridges the gap between structural and reduced form models of commodities while also attempting to explain the role of bubbles in commodity pricing in the presence of storage. We are able to obtain an autoregressive form for commodities which thereby allows the use of existing procedures for testing explosiveness. At the same time it also highlights the limitations and restrictions that have to be placed in order to test explosiveness in commodities using procedures that are conducted on the autoregressive form. The model also lays the foundation of multi-state analysis which we later use to test commodity market efficiency.

3.3 A Partial Equilibrium Model of commodity prices

The foundations of our partial equilibrium model are based on Muth's (1961) model. As alluded to above, our model aims to arrive at a reduced form for commodity prices that is readily employed in bubble testing. A number of microeconomic studies have attempted to explain rational bubbles (e.g. Blanchard 1979, Abreu et al 2003) which make a distinction between fundamentals and bubbles.

As pointed out by West (1987, 1988) such a distinction is arbitrary and depends on how one models fundamentals. If we restrict our definition of fundamentals then we are sure to find evidence of bubbles. A broader definition of fundamentals on the other hand that takes into account individual decisions incorporating trade volumes and beliefs making it difficult to define what a bubble is; is it a deviation from each individual's decision rule or has some fundamental element that is part of individual decision making been neglected.

We restrict our attention (initially) to two states of nature; a mean reverting state and a non-stationary state. As we point out in the appendix, if the two states of nature were assumed to be a random walk and an explosive state, the process will not have a steady state distribution; thus, for the process to have a stationary distribution, the price needs to be mean reverting for some length of time (albeit small). The mean reverting state nests the efficient state and is the default state for a commodity; market agents make use of all available information to make decisions which is already reflected in the price (although at times the process may become mean reverting). No other systematic gains can be made. Inventory owners act to stabilize markets i.e. they sell inventory when prices are expected to be high and build up inventory when prices are expected to be lower in the future. Our model in Chapter 5 primarily models this state.

The second state on the other hand is the bubble state; this may be triggered due to a variety of reasons including low inventories, a large productivity shock or through increasing uncertainty (due to a financial crisis for instance). The storage market plays a destabilising role in this case; we observe hoarding behaviour from investors. In the bubble state, the investors change behaviour markedly. They build up inventories when they expect future prices to compensate them for storage and current period price. Plausible reasons for this include anticipation of further increases in future periods as well as inventory managers being extra cautious in reaction to uncertainty. Re-building sufficient inventory for future periods (where prices are expected to stay high) is another plausible motive for this behaviour. We describe our partial equilibrium model below. All variables are market aggregates.

3.3.1 Demand, Supply and Storage:

We use a linear aggregate demand function as is conventional in the commodity literature.

$$D_t = -\beta P_t \qquad\qquad \beta > 0 \qquad\qquad (3.1)$$

The exclusion of the intercept term does not substantially change the analysis that follows[1]. We consider the impact of the intercept term (including a switching intercept term) in chapter 4. For this analysis we can assume that the demand is a deviation from a maximum level of demand.

The supply function is also assumed to be a linear function of price. This is different from the storage model used in chapter 5 where it is assumed that supply is a stochastic process dependant on weather. Farmers/miners react to higher prices by supplying more of the commodity. Using aggregate relationships allows us to assume away any individual capacity or productivity constraints. Productivity constraints can be incorporated in the model but we have not included them to keep the focus on storage driven dynamics. Supply behaves like a flow where new production comes in

---

[1] Inclusion of a constant requires a switching constant term in the reduced form. The switching constant does not alter our results or our analysis so for simplicity the analysis is presented without a constant term.

response to changes in price. No supply shocks are considered for this model; these are dealt with in Chapter 5.

$$S_t = \gamma P_t \qquad\qquad \gamma > 0 \qquad\qquad\qquad (3.2)$$

We simplify the storage market by reducing the decision from a stock decision to a flow decision. Instead of deciding how much to sell and how much to store, the decision investors make over time is how much inventory to carry relative to previous periods. In essence, we are modelling the investment decision faced by storage owners. Between two periods, they can increase inventory or decrease inventory. This decision is taken keeping in mind the future prospects of the market i.e. expected future price. We are implicitly assuming that storage costs stay fixed (as they do not feature in the decision) and that there are no aggregate capacity constraints.

While assuming no capacity constraints may have been a restrictive assumption if we considered individual investors, it is not restrictive when aggregate storage demand is considered. If a particular set of commodities is facing a boom it is likely that the market will dedicate more storage space to storing that commodity as opposed to commodities with lower margins.

$$I_{t+1} - I_t = \alpha_t(P_{t+1}^e - P_t) \qquad\qquad\qquad (3.3)$$

where $P_{t+1}^e = E_t(P_{t+1})$. I_t represents stock levels in period t and I_{t+1} is storage carried over to next period. α_t represents the responsiveness of investors to changes in price. The parameter α_t switches values across states and determines what state the market is in. It measures the impact of a change in expected price on inventory investment. Our assumptions on expectations coupled with restrictions placed on α_t determine the state that the commodity market is in at a particular point in time. As seen in equation (3.3), the storage decision is taken based on the expectation of future price and how it compares to the prevailing price. We specify the values α_t can take in the following sub-section and analyse the information it could contain.

3.3.2 Market Equilibrium

Market equilibrium requires that the quantity supplied and the storage carried over from previous periods be equal to the demand in the current period and any additional storage demand. Thus, equilibrium in this commodity market may be characterized as follows:

$$D_t + I_{t+1} = S_t + I_t \qquad\qquad\qquad (3.4)$$

Expectations play a critical role in equilibrium as expectation formation influences the storage decision. We assume rational expectations in our model i.e. investors take all available information

into account in making their storage decision. Every period the price expected by investors deviates from actual price by a random term.

$$P^e_{t+1} = P_{t+1} + \epsilon_{t+1} \qquad\qquad (3.5)$$

We assume that the error in expectations is ϵ_{t+1} which has a normal distribution with a mean 0 and variance σ^2. This implies that on average the investors get the price right but period on period they err in their forecast.

Substituting (3.1) (3.2) (3.3) and (3.5) into (3.4) we get:

$$-\beta P_t + \alpha_t(P_{t+1} + \epsilon_{t+1} - P_t) = \gamma P_t$$

This relationship simplifies to the following expression:

$$P_{t+1} = \frac{\alpha_t + \beta + \gamma}{\alpha_t} P_t + \eta_{t+1} \qquad\qquad (3.6a) \quad \text{or more succinctly}$$

$$P_{t+1} = \phi_t P_t + \eta_{t+1} \qquad\qquad (3.6b)^2$$

A two state model (in logs or levels) can thus be represents as follows:

$$\Delta P_{t+1} = [\phi_0(1 - s_t) + \phi_1 s_t]P_t + \eta_{t+1} \qquad (3.6c)$$

where $\epsilon_{t+1} = -\eta_{t+1}$ (we do not lose any generality but the representation with a positive error term is more conventional); s_t represents the state variable and takes a value of 0 in the mean reverting state and 1 in the explosive state. We assume that the state variable is an unobserved Markov process (Chapter 4 considers exogenously triggered states). Section 3.4 elaborates on Markov processes, the model and model estimation. Appendix 3A lays out the conditions required for the price process in 3.6 to have a steady-state distribution.

In equilibrium, the sensitivity of investors to expected changes in prices is what determines how prices behave i.e. the value of α_t. It also depends on the direction in which prices are expected to go. α_t may be dependent on a number of factors that could cause a change in investor behaviour even though expectations continue to be rational. These factors may be purely behavioural; we can think of these behavioural changes as the fads that Shiller refers to (2000). Alternatively, we could think of α_t as containing additional informational content on the dynamics of the market. For example, we may posit that α_t is dependent upon investment horizon i.e. investors' expectations on how long an

---

[2] The reduced form, had a constant term been included, would have been: $P_{t+1} = \phi_t P_t + \psi_t + \eta_{t+1}$, where ω_t is a switching constant term and depends on the value taken by α_t. Given the range of values for α_t discussed above, the constant term is zero if α_t is large or the process is in a random walk state; it takes a positive value in the explosive state when α_t is a positive constant. Refer to chapter 4 for further comments when a drift is present.

increase or decrease in price will continue. When investment horizon is short, i.e. the price change is expected to last only one period, investors react immediately. If $P_{t+1}^e > P_t$, the price increase is expected to last only one period. Thus, in order to maximize profits, investors divest and run down their inventories. Since all pricing information is already reflected in the price in this setting, we can think of this state as the efficient or mean reverting state. In this situation, $\alpha_t < 0$, any new information regarding the price, will immediately be reflected in the price through the movement of inventories.

On the other hand, if investment horizon is long as price changes are expected to last beyond one period, investors' behave differently. If $P_{t+1}^e > P_t$ and the directional change is expected to last beyond one period, investors' realize that price is on an upward path. Thus, they can continually increase their profits by building up their inventory and selling off some of it in the following period. This implies that $\alpha_t > 0$. Blanchard et al (1982) provide a similar explanation for the rise of a bubble. Investors continue to buy inventory at ever higher prices because they expect this rise to continue even though they are aware that each period there is a probability that the 'bubble' may collapse. Thus, $\alpha_t > 0$ corresponds to the explosive state. Below, we discuss how these different behavioural patterns affect the reduced form parameter, ϕ_t. Note, that we have only provided one possible explanation for the informational content of α_t. Other possible explanations may be favoured by behavioural or experimental economics. We wish to keep the interpretation of α_t open for future research on the subject.

We posit values for α_t for which the market behaves like a mean reverting state or a bubble state, respectively. Let $\alpha_t = b_t - \bar{\alpha}$, where b_t takes two values which determine the state of the market and $\bar{\alpha} > 0$ such that when $b_t = 0$, $\frac{\alpha_t + \beta + \gamma}{\alpha_t} \leq 1$ i.e. α_t is large and negative relative to β and γ i.e. $\alpha_t + \beta + \gamma < 0$ Under these conditions, the reduced form becomes a random walk or a stationary process depending on the magnitude of α_t, as explained above. The behaviour of the asset price in this state is akin to the behaviour we observe in our model in Chapter 5.

In the bubble state b_t takes on a value greater than $\bar{\alpha}$ which implies $\frac{\alpha_t + \beta + \gamma}{\alpha_t} > 1$ as all parameters are positive. This represents a two-state switching form for the commodity price in equilibrium. A data-generating process for b_t could also be stated, making the parameter a continuous variable; however, restricting the parameter to a finite number of values, such as the case of a Bernoulli random variable or a Markov process, keeps the analysis tractable and simpler to comprehend. Importantly, it also allows us to talk about the efficient market hypothesis in this context. As per the results in Appendix 3, we will consider a mean reverting and an explosive state. If we have a random walk state instead of a mean reverting state, the two-state processes will not have a steady-state distribution (although that is not an entirely implausible situation).

While this simple model is able to capture behavioural aspects, it does not capture fundamental supply and demand shocks; the impact of supply shocks is analysed in Chapter 5. Investors react when they foresee uncertainty in their expectations. In the mean reverting or efficient state (i.e. when $\phi_t \leq 1$), investors have very limited additional inventory demand and sell all their excess inventory which causes the process to follow a mean reverting form (i.e. a large and negative α_t). However, this may eventually lead to a shortage of the commodity if the demand for the commodity stays strong and could cause a switch from the random walk state to the bubble state (after a few periods). This would explain why in the run up to peak prices, inventories reach historic lows.

In the bubble state, inventories are already low and the price is expected to stay high for the foreseeable future. Expectations of high prices in the future coupled with already low inventories encourage investors to start rebuilding their inventories leading to an increase in inventory demand. As investors re-build their inventories, this adds further pressure on the price to grow as they crowd out private consumption, until inventories reach a point where the market corrects its expectations and jumps back to the mean reverting or efficient state. This explanation is in line with proponents of the rational expectations theory of the bubble (e.g. Blanchard, 1979). Thus, the investors hoard inventory, anticipating even higher returns in the future.

The build-up of inventories can eventually lead to a collapse of the bubble state, which could alternatively also be triggered by a slowing economy (which decreases the demand of the commodity leading to a glut). The market realizes that sufficient inventory has been built and the high price is no longer warranted. In reality, the transition to or from a random walk is more smooth. This would require a specification of b_t, such that 3.6 becomes a smooth transition autoregressive model (STAR). We do not consider this class of models in this dissertation but it is worth considering for future research.

The above analysis is particularly relevant for mining commodities. While the inventories for agricultural commodities can be built up or sold within a year, building capacity for mining commodities takes substantially longer. An increase in price needs to sustain for a period of time in order for suppliers to build additional capacity or to make existing capacity operational. Similarly when additional capacity has been brought on board it is difficult to take it offline within a few months. This explains why commodity prices keep rising or falling even though demand may adjust. When a price collapses initially, producers are often unable to take capacity offline immediately as they have to fulfil existing contracts. Since production does not reduce, the price continues to fall down until production levels adjust.

The fad theory of bubbles accredited to Shiller (1981) can also be justified by thinking of the stochastic bubble parameter b_t as a measure of uncertainty or paradoxically an expectation of very

high prices by investors. When uncertainty is high and b_t is high and greater than 0 investors become more cautious. The fad explanation would not require an explicit data generating process for b_t which could be representative of what Shiller terms animal spirits or irrational exuberance. Thus, our framework accommodates both rational expectations and fads narratives based on how an economic theorist looks at the informational content of b_t.

The final form of our model has dynamic and stochastic features which can capture changing volatility or a stochastic parameter. As mentioned before, this model is restrictive and only captures shocks to investor sentiments; Chapter 5 looks at the impact of Supply shocks. We treat this reduced form as a hidden Markov process and proceed to estimate equation (3.6) using Markov-switching methodologies. In chapter 4 we use the same specification but estimate the model using threshold autoregressions with an exogenous trigger.

3.4 Estimation Methodology

Equation (3.6) is in the form of a switching auto-regression with the parameter ϕ_t being the switching parameter. The switching regression methodology has developed in leaps and bounds since the seminal articles on the subject by Goldfelt and Quant (1973) and Quant (1972). Hamilton (1989) introduced dynamics within the Markov Switching framework by applying Markov Switching to GNP data, where GNP grows at different rates during booms and recessions. Other authors who have attempted similar approaches for bubble detection include Van Norden et al (2002), Enders et al (2001) who use a generalized Momentum Threshold Auto Regression (MTAR) and Bohl (2003).

Hidden Markov switching models lend themselves readily to estimation through the application of Kalman Filtering techniques introduced in econometrics by Harvey (1981). Models with switching parameters can be expressed in a State-Space setting which allows the application of the Kalman filter thereby enabling the estimation of time varying parameters as well as providing the approximate maximum likelihood function for the data generating process under question.

Kim (1993a, 1993b, 1994) has used state-space algorithms for estimating Markov switching Auto Regressions also allowing for varying transition probabilities. His series of articles is also accredited with introducing heteroscedastic disturbances in the Markov-switching regression setting. It is important to consider heteroscadasticity in this context as it may be plausible to suggest that being in an explosive state causes exuberance among investors and they may react more to shocks in the bubble regime than in the random walk or stationary regime, leading to heteroscedastic disturbances. Indeed we do note evidence of heteroscedasticity in our estimates below.

Alternative methods for estimating switching regressions have also been developed in the Bayesian setting. Chib et al (1993) introduced switching regression using Gibbs Sampling in a Bayesian setting.

Kim and Nelson (1998, 1999) have further developed this methodology through an algorithm that uses a Bayesian Gibbs Sampling approach in a State Space setting to estimate a Markov Switching model. More recently Kim et al (2012) have employed a Bayesian Markov Chain Monte Carlo (MCMC) methodology to estimate evolving regime-specific parameters i.e. the parameters act like random variables instead of taking on fixed values. Thus, a variety of methodologies are available to us to estimate equation (3.6).

Our contribution to this literature is to use information from recursive unit root or bubble tests in order to find better estimates of smoothed regime probabilities. Other switching-regression methodologies such as those proposed by Bohl are not feasible for commodities as data on a fundamental process is often not available at the same frequency as pricing data. Inventory or production data is usually annual and may not be as reliable as pricing data as it is based upon estimates.

We use the test developed by Philips et al (2013), which was introduced in Chapter 2, in order to date bubbles in commodities[3]. The detected bubble dates are then used to estimate Markov-state transition probabilities. A relatively simple approach is to run an auto-regression using the bubble dates as a separate dummy variable. This will provide an estimate of the auto-regressive parameter for the two states. However, the underlying assumption for using this approach is that bubbles are known for certain and we know when they arise. Forecasting on the basis of such a regression will require a conjecture on the state the commodity price is currently in. Nevertheless we do carry out the dummy variable auto-regression in order to compare the forecasting results with those obtained from the Markov-switching approach. The GSADF test has gained significant traction since it was first introduced in 2013.

Within the Markov-switching regression framework, we use bubble dates to arrive at an estimate for the expected duration of the two regimes. Estimates of expected duration then allow us to derive the matrix of transition probabilities. The sub-section below explains what we mean by transition probabilities. The transition probabilities thus obtained can be used in different ways. An obvious way to use these probabilities is to constrain the Markov-switching algorithm such that the two-states occur with the same transition probabilities.

A second methodology is to use these as initial probabilities and to compare them to the probabilities once the switching-parameters have been estimated. Another alternative is to use these transition probabilities as initial values while allowing the transition probabilities to be random variables. This will be the most prudent approach given that bubbles tend not to occur at specific intervals and the duration of each bubble state tends to vary. A caveat that needs to be kept in mind is that using

---

[3] While the modified Bhargava statistic introduced in Chapter 2 does an adequate job of detecting explosiveness in price series, it is less precise with dating periods of explosiveness; hence, we use the GSADF statistic since it allows us to date periods of explosiveness and thereby estimate a transition matrix as noted in this section

transition probabilities derived from Philips et al's test does not guarantee that the algorithm will match the bubble dates exactly. The two states detected may be markedly different from the bubble dates suggested by the Philips et al test.

In what follows we begin by briefly describing the Markov-switching maximum likelihood approach that we employ. We use the MATLAB toolbox developed by Perlin (2014) to estimate these models which employs the maximum likelihood methodology instead of the Bayesian approach as mentioned above. The second sub-section describes our use of the GSADF test.

3.4.1 Markov-Switching regimes:

The following section is based on Hall et al (1999) as it matches the aims and objectives of this chapter. For a general treatment of Markov-Switching methods using State-space methods and/or Bayesian Methods the interested user is referred to Kim and Nelson (1999). If our objective was to carry out a right-sided unit root test (a crude test for explosive behaviour), we could express equation (3.6) in the following Augmented-Dickey Fuller form:

$$\Delta P_{t+1} = \phi_t P_t + \sum_{i=0}^{k} \psi_i \Delta P_{t-i} + v_t \qquad (3.7)$$

where Δ is the first-difference operator and v_t is a white-noise disturbance term. In a single regime-framework we could estimate the auto-regressive parameter ϕ_t and test its significance to classify if the series has a unit root. However, with regime-switching we need to take the switching parameter into account. As stated above, we assume that there are two states of nature, a mean reverting/random walk state, $s_t = 0$ and a bubble state $s_t = 1$. Thus, equation (3.7) becomes:

$$\Delta P_{t+1} = [\phi_0(1 - s_t) + \phi_1(s_t)]P_t + \sum_{i=1}^{k}[\psi_{0i}(1 - s_t) + \psi_{1i}(s_t)]\Delta P_{t-i} + \sigma_e e_t \quad (3.8)$$

here e_t is an i.i.d random variable with zero mean and unit variance. The state, s_t, is selected by nature at each time t. However, the probability of the states occurring is not independent; they are Markovian i.e. the probability that state j occurs tomorrow depends on what state the process is in today. In essence this is the main benefit of using Markovian state switching to estimate state-specific parameters. The process $\{s_t\}$ forms a Markov Chain on the space $\{0,1\}$ with the following transition matrix:

$$A = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \qquad (3.9)$$

$$0 < p < 1$$
$$0 < q < 1$$

where the elements a_{ij} of the matrix A represent $Pr(s_t = i | s_{t-1} = j)$. Thus, the first element in the matrix is the probability that we stay in state 0 tomorrow given that we are in state 0 today or if we refer to the states in our model, it is the probability that we will be in the stationary state tomorrow given that we are in the stationary state today. Independence of the disturbance term from the state variables at all time periods and across all states is also required although state-specific disturbance terms can be accommodated as discussed in Kim et al (1999).

To estimate the parameters of the model as well the probabilities we need to use the maximum likelihood of the price process. The following is based on Kim et al (1999). Let the log-likelihood function of the price process be $f(P_t | I_{t-1}, s_t, s_{t-1})$ where I_{t-1} is the information set up to time period $t-1$ i.e. it contains the information on both states up to time period (t-1). s_t is the state of nature in period t and s_{t-1} is the state of nature in period $t-1$. If we know s_t apriori, the problem simplifies greatly and becomes a dummy variable problem. The objective then is to simply maximize the following log-likelihood function:

$$\ln L = \sum_{t=1}^{T} \ln(f(P_t | I_{t-1}, s_t, s_{t-1}) \tag{3.10}$$

For an AR(1) with a Markov-switching parameter as in 3.6, the probability density function, for the two state case, is given by:

$$f(P_t | I_{t-1}, s_t, s_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_{s_t}^2}} \exp(-\frac{\{(P_t - [\phi_0(1-s_t) + \phi_1 s_t]P_{t-1}\}^2}{2\sigma_{s_t}^2}$$

where the states are known with certainty.

With unobserved states however, the problem becomes sufficiently more involved and complicated. With unobserved states, instead of deriving the likelihood of the price process, we will need the joint likelihood of the prices process, the state in the current period and the state in the previous period conditional on the information up to the previous period; i.e.

$$f(P_t, s_t, s_{t-1} | I_{t-1}) = f(P_t | s_t, s_{t-1}, I_{t-1}) Pr[s_t, s_{t-1} | I_{t-1}] \tag{3.11}$$

where $f(P_t | s_t, s_{t-1}, I_{t-1})$ is given in the specification above. In order to get $f(P_t | I_{t-1})$ or the marginal density of the price process, we need to integrate out the states, s_t and s_{t-1} out of the joint density function (3.11). Thus for the two state case, we have,

$$f(P_t | I_{t-1}) = \sum_{s_t=0}^{1} \sum_{s_{t-1}=0}^{1} f(P_t, s_t, s_{t-1} | I_{t-1})$$

$$= \sum_{s_t=0}^{1} \sum_{s_{t-1}=0}^{1} f(P_t|s_t, s_{t-1}, I_{t-1})Pr(s_t, s_{t-1}|I_{t-1})$$

In our example the marginal density is a weighted average of 2^2=4 weights. With 'M' states this becomes an M^2 dimensional problem at **each** time period. The weights are given by the conditional state probabilities. Thus, in order to estimate switching parameters with unobserved states we need to maximize the following likelihood function:

$$\ln L = \sum_{t=1}^{T} \ln \left\{ \sum_{s_t=0}^{1} \sum_{s_{t-1}=0}^{1} f(P_t|s_t, s_{t-1}, I_{t-1})Pr(s_t, s_{t-1}|I_{t-1}) \right\} \quad (3.12)$$

In order to maximize the above likelihood we still need to find the weights. We refer to Kim et al (1999) which explains a recursive procedure for estimating the probability weights and also provides an algorithm for reducing the dimensionality of the problem. The recursive procedure requires stating initial values for the transition probabilities. The main difference between our approach and other approaches is that we do not use a naïve or un-informative state probability. We use information from the GSADF statistic to come up with an estimate of the transition probabilities. Once we fix the values of the transition probabilities, the marginal density of P_t can be easily obtained and the likelihood function maximized as we can use the transition probabilities as weights.

Since we estimate $Pr(s_t, s_{t-1}|I_{t-1})$ directly from unit root tests, (3.12) can be maximized easily. Assuming a probability distribution for P_t, the only element in (3.12) that needs to be maximized is $f(P_t|s_t, s_{t-1}, I_{t-1})$ with respect to the state parameters ϕ_t and $\sigma_{s_t}^2$. In most cases, 3.12 cannot be estimated analytically and thus a numerical procedure is required. The likelihood can easily incorporate the case for heteroscedastic disturbances as well (See Kim et al, p. 146)

To find $Pr(s_t, s_{t-1}|I_{t-1})$, we rely on the expected duration, D, of each state. The expected duration D of each of the two states is given by:

$$E(D) = \sum_{j=0}^{\infty} jPr[D = j]$$

$$= 1Pr[s_{t+1} \neq j|s_t = j] + 2Pr[s_{t+1} = j, s_{t+2} \neq j|s_t = j]$$
$$+ 3Pr[s_{t+1} = j, s_{t+2} = j, s_{t+3} \neq j|s_t = j] + \cdots$$

Given transition probabilities for each state, p_{jj} the duration equation becomes:

$$E(D) = 1 \times (1 - p_{jj}) + 2 \times p_{jj}(1 - p_{jj}) + 3 \times p_{jj}^2(1 - p_{jj}) + \cdots = \frac{1}{1-p_{jj}} \quad (3.13)$$

Thus, once we have the expected duration of each state using this method of moments, the transition probabilities follow from equation (3.13). In section 3.6 we consider results from this regression with the two states being an explosive state and a mean reverting state respectively.

Once the model has been estimated and parameter estimates found, we can find smoothed state probabilities based on Kim et al's methodology (1999, p. 68-69):

$$\Pr[s_t = j | I_T] = \sum_{k=0}^{1} \Pr[s_t = j, s_{t+1} = k | I_T]$$

where I_T represents all information in the sample (T denotes the whole sample as opposed to 't' which denotes the current time period).

Note that the analysis and discussion that follow primarily talk about smoothed probabilities of different states rather than the magnitude of the parameter. Thus, the procedure outlined in this chapter is best seen as a metric for market efficiency. The following chapter provides a more robust methodology for estimating parameters as well as measuring efficiency.

3.4.2 Estimating transition probabilities

We briefly revisit the GSADF statistic which we discussed in detail in Chapter 2. Given the model in (3.7), an initial window r_0 is stipulated. Given this window, the empirical regression in (3.7) is estimated and the GSADF statistic calculated. Caspi (2014) has coded the GSADF statistic within E-views which allows a user to calculate not only the GSADF statistic and the ADF sequence but also allows the generation of the critical value sequence. We have used Caspi's Eviews program to date bubbles in commodities in this chapter.

Once the test has been run and a price series dated for explosiveness, we can estimate the expected duration of the explosive state by dividing the number of periods for which the price was in the bubble state by the total number of observations. Using (3.13) we can then estimate the transition probability for the explosive as well the non-explosive state respectively. The underlying assumption is that our measure is an unbiased estimator of expected duration. Thus, $E(D) = \frac{number\ of\ periods\ in\ state, s_t}{number\ of\ switches}$. Using the transition probabilities we estimate parameters from (3.12) by applying Perlin's (2014) algorithm.

3.5 Copper: A Brief Overview

Copper is among the most widely traded and consumed metals in the world. With estimated reserves of over 7,000 Million Tonnes (MT) (ICSG, 2014), it is also available in great quantity although not all of it is currently economically viable. It is a malleable and ductile metal and is used in a variety of industrial equipment, construction, power generation and transmission, transportation and consumer goods, particularly electrical goods and wiring. It is essential in information and communication technology. In terms of usage, the world refined copper demand stood at 21.2 MT in 2013 with Asia accounting for nearly 13.8 MT of that demand. The demand for copper products has been growing at a steady rate though supply has usually kept pace with this demand. There have been periods when demand has outstripped supply which has resulted in shrinking inventories and rapid price increases as we discuss below.

While no metals are available in infinite quantity, it is unlikely that the world will face a shortage of copper in the near future. In addition to mining, copper is also widely recycled and the recycled usage has been growing year on year. According to the International Copper Study Group, mining production up till 2014 had reached 18.7 MT per year with Chile being responsible for nearly one-third of that production (5.8 MT). Recycling accounts for around one-third of copper production every year. Refined Copper production on the other hand has been increasingly concentrated in Asia with China producing more than a third of refined copper at (6.5 MT) (USGS, 2015). Despite being the largest producer of refined Copper, China remains the top importer of both Copper ores and refined Copper. This may change in the near future as the Chinese economy moves from an investment based economy to a consumption based economy and other emerging economies see a jump in Copper demand.

The risk of disruption to copper supply is considered relatively low as copper production is dispersed worldwide and the metal is distributed in regions throughout the world. Though the world may not run out of copper any time soon there are a number of other risks that can impact the supply of copper. As copper extraction from a particular mine increases, the quality of ore from that mine starts to decline. This has been particularly true for some mines in Chile and the USA (Papp et al, 2014).

Like other metals, the production of copper requires a large capital investment over a substantial period of time. Episodes such as the financial crisis make it difficult for large copper producers to raise financial capital. Similarly, volatility of copper prices can potentially make production in some mines infeasible as we saw in 2015-16 with the plunge in metal prices driven by low oil prices.

Another major risk to copper production is posed by climate change. With environmental agencies increasingly scrutinizing risks posed to the atmosphere by industrial processes such as the extraction of copper, the regulation on such production is likely to increase in the coming years. As copper

production primarily uses coal as the main source of power in copper mines, it is likely that the cost of copper production will increase if companies are required to switch to other sources. In addition, there are also recurrent risks of labour strikes which can often be prolonged and can halt production from major mines.

Copper contracts are exchanged on three major commodity exchanges; the London Metal Exchange (LME), the New York Mercantile Exchange (NYMEX) and the Shanghai Futures Exchange (SHFE) with the latter of the three increasingly gaining a larger share of the commodity market in general and the copper market in particular. In addition to offering a facility for buying and selling copper contracts (including futures and options), the exchanges also provide storage facilities which enable physical exchange of copper (ICSG) and settlement of contracts. While producers of copper maintain the largest inventories of copper, exchanges also store a substantial amount of copper.

Our monthly data extends from 1957 through to 2014 – Figure 3.1 in the appendix shows a time series plot of the Copper price series over the period. There have been a number of events during this time period during which the price of copper has been volatile. As mentioned in the previous sections, these events may have acted as the trigger for the switch from the random walk state to the bubble state.

The first of these events occurred during the mid-1960's. The start of the Vietnam War coupled with strong growth and some of the longest strikes in copper mines in the US led to the increase in price of copper through to the end of the 1960's. The next major jumps are seen during the oil price crises from 1972-74 and 1979-80. These jumps are primarily triggered by an increasing cost of production which made metal production as a whole an expensive activity. The 1979-80 prices were also marked by a record jump in copper consumption and lower inventory levels (Edelstein, 1999).

We see another peak in the late 1980's which was caused by growing world consumption and historically low inventories of copper. This particular price jump was specific to the copper market as opposed to the oil price shock which impacted virtually all commodities. The boom in East Asia also caused a jump in copper prices until the market cooled down after the financial crisis hit East Asia. In 2005, we observe the steepest rise in copper prices with prices as high as $3000 per Metric Tonne. The price jumped to almost $9000 per Metric Tonne during the financial crisis before plummeting back to $3000. This marks the most volatile period in the data we consider.

While the initial price increase was caused by exponential production growth in the Chinese market, the jump to $9000 in mid-late 2008 was the result of investment flowing from the crumbling sub-prime mortgage crises into the non-ferrous metals market as investors sought relatively safe returns. Metal prices in general and Copper in particular saw another sharp increase in 2011 mainly due to economic recovery in the developed world and strong economic performance from emerging

economies such as China and India. The situation was exacerbated by low stockpiles which further provided an impetus for a stronger copper price. In the sections below we focus primarily on the copper market but do report results for other non-ferrous metals when we introduce our methodology for estimating market efficiency.

3.6 MSAR Results with two regimes:

For the results in the current and subsequent sections we used the International Monetary Fund's International Financial statistics database to access monthly data on commodity prices for 6 non-ferrous metals (Aluminium, Copper, Lead, Nickel, Tin and Zinc). We use an initial window size of 36 for the GSADF test on Copper. Log prices are used and the dependant variable is the growth rate in price; thus, a parameter value greater than 0 represents an explosive state and a value close to 0 represents the random walk or efficient state. We use a 5 percent critical sequence for our analysis.

As outlined in the section 3.4.1, we try to improve the MSAR algorithm by estimating the Markov state probabilities directly from our GSADF test. This requires datestamping bubbles using the GSADF test developed by Philips et al (2013). We carried out the GSADF test using the Eviews package developed by Caspi (2014). The results for the test are shown in Figure 3.1 (the blue line is the GSADF stastic sequence and the red line is the critical sequence). As seen in the graph there are 5 instances that qualify for the bubble state according to the criteria put forth by Philips et al (recall that when the GSADF sequence for the copper series is above the critical value sequence in red, the series is said to be in a bubble state). These instances run from July 1964 to January 1965 (7 months), July 1973 to October 1973 (4 months), October 1987 to January 1988 (4 months), December 2003 to April 2004 (5 months) and December 2005 to October 2006 (11 months).

Based on these dates the average duration of the bubble state is 6.2 months with a standard deviation of 2.95 months. Given the length of the data, we do not have sufficient information to reliably estimate the full distribution of bubble states so we primarily rely on the expected duration of the bubble state. Using E(duration of bubble state) = 6.2 months, our estimated Markov-state probability for the commodity price to stay in a bubble state given that it was in a bubble state in the previous period (p_{22}) is 0.8387 and the probability that the commodity price will behave like a mean reverting process given that it was mean reverting in the previous period (p_{11}) is 0.9910.

The calculation of the transition probabilities from the expected duration data is based on the method outlined in Section 3.4.2. The transition probabilities are in line with the literature on bubbles in commodity markets; long periods of stable prices are followed by some short periods of a bubble state during which the price of a commodity can behave in a volatile manner until inventories and supply adjust. As seen above, the longest such period was in the run up to the financial crisis.

We begin our analysis by running a basic regression and using the date stamps to generate a dummy variable (BubbleCopp in Table 3.1). The dummy takes a value of 1 whenever the GSADF statistic for copper is above the critical GSADF sequence of values. The regression is essentially a random walk including the bubble dummy. As seen in table 3.1, we fail to reject the null hypothesis for no unit root and the bubble dummy is significant and positive. Since the regression is in levels rather than logs, the coefficient (209.86) is a dollar value. Inclusion of a time trend does not make the coefficient insignificant although the coefficient reduces in magnitude. One may consider this regression a very simple form of a Markov Switching Auto-Regression where the states are known with certainty. Although the regression itself is very simple, its implications are important as it indicates the existence of a bubble state separate from a random walk during which the price process is likely to follow an explosive path.

The second stage of our analysis relies on unrestricted Markov State Auto Regressions. For the MSAR regressions we estimate equation 3.6c using monthly returns. Including a trend or more lags does not substantially alter our results so the analysis presented below uses the model as outlined in section 3.3. We use the MSAR code developed by Perlin (2014) for this part of the analysis. As a first step we estimate a Markov Switching Auto Regression without placing any restrictions on the coefficients. The transition matrix thus obtained for Copper is:

$$\begin{pmatrix} 0.97 & 0.03 \\ 0.08 & 0.92 \end{pmatrix}$$

The Smoothed Markov-State probabilities are shown in Figure 3.2 below. As seen from the results, the regime probabilities predict a much longer second state probability than generally believed. This procedure is unable to correctly identify an explosive state. In the run up to the financial crisis, the copper price appears to follow a steep trend, which influences the algorithm a great deal. The unrestricted model predicts an expected duration for the second bubble of almost 16 months which is much longer than what we observe in the data.

Figure 3.2 shows dates during which the smoothed probability of being in the second state is high. High probabilities for the second state occur in periods during the 1960's; in 1973, between 1988-1990 and from 2005 onwards. While the unrestricted model does admirably before 2005, it overestimates the probability of being in a bubble state after 2005. The results suggest that the copper price series has fundamentally changed after 2005 and has moved to the bubble state. It only reverts back to stationarity around 2012.

Another reason why these results may have been impacted is the behaviour of the copper price series towards the end of the sample. At the end of the sample, the series is very volatile, and deviates significantly from the earlier period. It is likely that the unrestricted MSAR detected a high variance

and a low variance state. The steady state probabilities suggest that the series stays in state 1 for 69% of the time and in state 2 for 31% of the time. This is clearly an overstatement as empirical as well as theoretical analysis suggest a much shorter duration for the explosive state.

The final regression in this section restricts the transition state probabilities to values obtained from the GSADF test. While this does influence the estimates and the smoothed state probabilities, it is not intended to exactly replicate the GSADF test (which the first regression with the dummy variable does). Instead, we use an element from the distribution of the bubble state and incorporate that information in our MSAR. The ideal scenario would require a distributional restriction on the MSAR algorithm. The restricted transition matrix for Copper based on the bubble and stationary state date stamps from the GSADF test is:

$$\begin{pmatrix} 0.99 & 0.01 \\ 0.16 & 0.84 \end{pmatrix}$$

This suggests a steady state probability of 95% for the stationary and 5% for the bubble state which is in line with theory and empirical observation. Figure 3.3 plots smoothed probabilities from the restricted regression. The restricted model detects shorter bubbles which often coincide with dates during which copper prices were on an increasing path. We also note a bubble from 2005 onwards; however, as opposed to the unrestricted case the mean reverting state resumes from 2010 instead of 2012. The periods during which the bubble probability is relatively high (i.e. >0.3) tend to coincide with important events in the copper market.

The first instance in which we observe a high bubble state probability is 1967-1968 which (as stated in section 3.5) corresponds to periods of long strikes in copper mines which saw copper prices jump significantly. This is followed by the 1973-1974 period during the oil price crisis when prices increased in wake of the oil embargo by OPEC. The situation was exacerbated by low stockpiles which further provided an impetus for a stronger copper price. The next high sequence of bubble state probabilities is seen towards the end of 1979 and the beginning of 1980, a period of low stocks and high demand for copper which led to copper prices rising.

During much of the 1980's prices stay stable and our results indicate as such (high probability of state 1). Towards the end of the 1980's and the beginning of the 1990's we see the probability of the bubble state rising again. This was a period of historically low inventories and high demand which led to substantial increases in copper price including historical peaks being achieved. While both the unrestricted and the restricted model show high probabilities for the explosive state here, the restricted model does caution us as there is a significant fall in the bubble state probability in 1990 when prices had begun to stabilize. The unrestricted model continues to maintain a high bubble state probability during this time period.

State 1 probability stays high for most of the 1990's and the early 2000's with a short shift to state 2 during the East Asian financial crisis. From July 2007, both the unrestricted as well as the restricted model have a high probability for state 2. This period coincides with the financial crisis and its aftermath during which copper prices have stayed much higher and more volatile than historical averages. As stated before, the unrestricted model suggest a high bubble state probability up to 2012 while the restricted model suggest a return to a random walk state in 2010.

While both models capture important state shifts in the copper price series and give indications of when copper may be in a bubble state, our evidence suggests that the restricted model does better and it is better able to capture the bubble state. Particularly towards the end of the sample, the restricted model indicates that copper may have switched back to the stationary state which the unrestricted model does not. The unrestricted model substantially overstates the duration of the explosive state. Thus, without placing any restrictions on parameters we are able to obtain relatively accurate smoothed probabilities for each of the two states with our proposed methodology.

One criticism that can be levelled against our approach is with regard to our selection of critical sequences. As opposed to the unrestricted transition probabilities, the transition probabilities obtained from the GSADF test depend on the critical value chosen and may vary. Had we used a 10% critical sequence our transition probabilities and hence, the steady state probabilities will have been different. If we re-estimate the restricted model using a 10% right-sided GSADF critical sequence we obtain a steady-state probability of 8% for the explosive state. Thus, the steady-state probabilities tend to change as the critical sequence is changed. One's choice of the critical sequence depends on the type I error one is willing to accept for this analysis and how one chooses to define a bubble.

Results from sensitivity analysis at the 5% and 10% levels did not lead to significant changes in steady state probabilities. Steady state probabilities for the explosive state increased when a 10% critical sequence was used; however, the increase was in the 2-5% range. To justify our choice of a 5% right-sided critical sequence we carried out further analysis on the matter. Using grid search we tried to match the steady-state probabilities obtained from the unrestricted model in order to find the critical sequence and hence, the type I error probability that would give us the same results as the unrestricted model.

For copper we found that a right-sided critical sequence of 73% would be required to replicate the unrestricted estimation results. Thus, replicating the unrestricted Markov-estimation results for Copper will have required an unacceptable level of Type I error probability. Our analysis on the 3-state results (Section 3.7) requires similarly unrealistic Type I probabilities to obtain the unrestricted results. Therefore, we contend that conventional critical sequences (1%, 5%, and 10%) suffice for the restricted estimation and the results are not very sensitive at these levels. We did carry out 2-state

analysis for 5 other non-ferrous metals; however, for brevity we only visit these results briefly but relevant figures and analysis are available upon request.

Table 3.2 reports the switching parameter estimates from the restricted and unrestricted regressions while Table 3.3 reports the transition matrices and corresponding steady state probabilities under the unrestricted and restricted transition probabilities for the metals in our sample. The general pattern appears to be reasonably accurate; we observe both mean reverting ($\phi_t < 0$)and explosive ($\phi_t > 0$) states in the unrestricted as well as the restricted regressions. The smoothed probabilities on the other hand differ greatly and it is here that our methodology proves most useful. We next move on to an interesting application of our methodology and discuss results for all metals in greater detail.

3.7 Testing the efficient market hypothesis

The framework presented in the previous sections can be extended to test the efficient market hypothesis. To test market efficiency we consider 3 states instead of 2; the market can be in the efficient state (i.e. prices behave like a random walk), the explosive or bubble state as considered before or in a stationary state where the market corrects itself or in the case of commodities there is excess inventory or oversupply. We can continue using equation 3.6 for our analysis and use the switching regression estimation methodology. For the 3-state scenario we maximize the following likelihood function:

$$\ln L = \sum_{t=1}^{T} \ln\{ \sum_{j=0}^{2} \sum_{j=0}^{2} f(P_t | s_{j,t}, s_{j,t-1}, I_{t-1}) Pr(s_{j,t}, s_{j,t-1} | I_{t-1})\} \tag{3.14}$$

where $f(P_t | s_{j,t}, s_{j,t-1}, I_{t-1})$ is the same as given before and $\Pr(s_{j,t}, s_{j,t-1} | I_{t-1})$ can be found using GSADF tests. The model continues to be the Markov Switching Auto Regressive model of order 1 as in equation 3.8 above although this time we have 3 possible values for the auto-regressive parameter instead of 2. Thus, the estimated model is:

$$\Delta P_{t+1} = [\psi_0 s_{0,t} + \psi_1 s_{1,t} + \psi_2 s_{2,t}] + [\phi_0 s_{0,t} + \phi_1 s_{1,t} + \phi_2 s_{2,t}] P_t + \sigma_e e_t \tag{3.15}$$

where $s_{j,t} = 1$ when the process is in state j at time period t and is 0 otherwise.

In order to detect the periods during which the commodity series is stationary or mean reverting i.e. $\phi_j < 1$, we again rely on the GSADF test; however, instead of just considering a right-sided Dickey Fuller test we also use the original unit root test which is a left-sided test. The main difference between this procedure and the normal augmented Dickey Fuller test is that instead of relying on a singular value we work with a sequence of left-sided Dickey Fuller statistics similar in line to the

right-sided Dickey Fuller sequence which was used to date explosive regimes. Thus, we date the explosive regime by considering the 95% critical sequence and the mean reverting regime by considering the 5% critical sequence.

Simulations are used as they were for the 2-state case, but instead of the 95% confidence interval we also use the 5% confidence interval. Thus, any time the GSADF sequence for the commodity under question is below the 5% confidence interval the series is mean reverting. The efficient market hypothesis would require the price process to act like a random walk; therefore, any statistically significant break from the random walk i.e. stationarity or explosiveness will indicate a deviation from the efficient markets hypothesis. We consider the case of copper in more detail but also present results for other commodities. Figures are only provided for copper although these are available for other metals upon request. Table 3.4 reports transition probabilities and steady state probabilities for all metals.

Figure 3.4 shows regions where the market for copper is efficient, mean reverting and explosive (the green line represents the mean reverting sequence). Using this procedure we can divide the copper series into 3 different states. In the context of our efficient markets test we primarily rely on the transition matrix and the corresponding steady state probabilities but also report the smoothed state probabilities. The steady state probabilities vector will indicate how much time the series spends in each state thereby helping us understand how often the market deviates from the random walk or efficient state.

Stationary periods tend to occur during periods of oversupply or when inventories are high. These are periods during which the market is correcting itself and thus becomes somewhat predictable. It should be noted that just like the bubble literature, there is a probability that the series will go from being stationary to being a random walk and this may influence investor behaviour. For Copper, we find that there are substantially more switches to the stationary state than there are to the bubble state. Another thing we note in our results is that there are no switches from the bubble state to the stationary state or vice versa.

The longest periods during which copper prices stayed mean reverting occur during 1976-1979, 1984-1987 and 2009-2011. These periods are marked by falling copper prices and are usually separated from an explosive episode by at least a year. For copper specifically, these were periods when copper production increased and there was some evidence of oversupply. While the GSADF procedure in itself provides a good test for the efficient market hypothesis, we go a step further and estimate transition probabilities and steady state vectors for the 3 states. The steady state probabilities can also serve as a metric for market efficiency and can indicate the probability of a particular asset market being in an efficient state.

We will compare the transition matrix and the steady state probabilities vector across the restricted model (which we estimate from the GSADF test) and the unrestricted model (where we obtain the transition probabilities through directly estimating a 3-state Markov-switching auto-regression). Note that the unrestricted model is not confined in any way; thus, the random walk state may not be the first state. In fact we could have two stationary or two explosive states. Thus, we need an assumption in order to facilitate the comparison. For copper, we use the sign and magnitude of the parameter estimates as well as the magnitude of the steady state probability to identify the random walk state.

The random walk state should have a parameter estimate close to zero and under the null (the market is efficient) it should also have the highest steady state probability (if the market is mostly efficienct). The explosive state on the other hand should have a parameter estimate greater than zero and the stationary state should have a parameter estimate below 0 and significant. This identification strategy will allow us to compare the transition matrix and steady state probabilities estimated through the Markov-State regression to the transition matrix and steady state probabilities estimated from the GSADF test.

The 3-state Markov switching parameter results for Copper obtained through direct estimation of the model are contained in the table below:

Three State Markov Switching Regression. Parameter estimates for Copper

Parameter	Parameter Estimate	p Value
ϕ_0	-0.0005	0.00
ϕ_1	0.0056	0.00
ϕ_2	-0.0103	0.00

Using the identification strategy outlined above we note that state 1 is the random walk state; state 2, the explosive state and state 3 is the stationary state as it has a negative coefficient and a higher magnitude than state 1. The smoothed-state probabilities are shown in Figure 3.5. The series oscillates between stationary and explosive states with some interspersed episodes of stationarity. The random walk state dominates for most of the period although there are a number of switches to the explosive state particularly during the 60's, 70's and the 2000's. It is relatively easier to distinguish these three states despite the fact that we did not impose any restrictions on the transition probabilities. Note, however, that this result is an exception rather than the norm as results for other metals are different. In general, we are unable to distinguish the explosive and stationary states. The unrestricted transition matrix for copper is given below:

Unrestricted Transition matrix for Copper =

$$\begin{bmatrix} 0.95 & 0.04 & 0.01 \\ 0.04 & 0.84 & 0.12 \\ 0.14 & 0.26 & 0.60 \end{bmatrix}$$

The transition matrix suggests an expected duration of 21.34 months for the random walk state, 6.39 months for the explosive state and 2.49 months for the stationary state. Thus, the random walk state appears to be much shorter than the previous case (i.e. in the 2 state case). Another thing to note is that the results suggest a switch from the explosive state to the stationary state (transition probability of 0.12). We do not observe such switches in the data as seen from Figure 3.4. In fact we only observe a switch from the explosive to the stationary state in 1 out of the 6 metals we test (the exception being Zinc). The transition matrix also suggests switches from the stationary to the explosive state which is not present in the transition matrices estimated from the GSADF test. Overall these results suggest that copper prices deviate from the efficient market hypothesis for a substantial duration.

The steady state probabilities for Copper implied by the unrestricted regression are:

$$\begin{bmatrix} 0.57 \\ 0.32 \\ 0.11 \end{bmatrix}$$

Thus, the market is expected to be efficient 57% of the time, explosive 32% of the time and stationary 11% of the time according to the unrestricted model. The duration for the explosive regime has been significantly overestimated. It appears that a proportion of the random walk state as per the GSADF test has been identified as an explosive state by the algorithm.

Next, we estimate the transition matrix and the steady state probabilities for Copper from the GSADF test using Figure 3.4. When the GSADF sequence for Copper exceeds the 95% GSADF sequence the series is said to be in the explosive regime and when the GSADF Sequence for Copper is below the 5% GSADF sequence the series is considered stationary. Transition probabilities can then be calculated using estimated expected durations as specified in equation (3.13). This leads to the following transition matrix for Copper:

$$\begin{bmatrix} 0.96 & 0.01 & 0.03 \\ 0.16 & 0.84 & 0 \\ 0.07 & 0 & 0.93 \end{bmatrix}$$

Here again, the first state is the random walk state, the second state is the explosive state and the third state is the mean reverting state. Transition probabilities for state 1 to state 2 and state 1 to state 3 were based on the ratio of number of switches to each state from state 1. We noted a total of 19 switches from the random walk or efficient state. Of these 19, 5 were to the explosive state and 14 to

the stationary state. The expected duration is 24.2 periods for the random walk state which leads to a probability of 0.96. The remaining probability is apportioned between the explosive and stationary states respectively based on the number of switches from the random walk state. While the procedure we have outlined does not require re-estimation of the 3-state Markov autoregressive model, we perform this step for completeness. Figure 3.6 below shows the smoothed state probabilities based on a restricted 3-state Markov-Switching model where the transition matrix was restricted to the one estimated from the GSADF test.

The only common feature between the restricted and unrestricted models is that the random walk state is the dominant state. In stark contrast to the previous result, the stationary state is the second most frequent state with most switches being from the random walk to the stationary state rather than the explosive state. There are interspersed periods of explosiveness but these are short and tend to correspond with dates that were obtained from the GSADF test. In particular we note high probabilities for the explosive state in 1978-79 and in the run up to the great recession. The steady state probabilities for Copper are shown below:

$$\begin{bmatrix} 0.67 \\ 0.04 \\ 0.29 \end{bmatrix}$$

The steady state probability vector is very different from the unrestricted vector as we note that the stationary state occurs much more often than suggested by the unrestricted model. We also note shorter explosive states and a higher probability of the efficient state. Thus, our comparison provides a more convincing case for the efficient markets hypothesis for Copper than the unrestricted 3-state regression. It is also more in line with the results in chapter 4. Price tends to be mean reverting when inventories are high or when the global economy is recovering from a recession as outlined in Section 3.5 above (Edelstein, 1999). The random walk or efficient state tends to prevail for a majority of the length of the series.

The analysis is extended to include other metals. Table 3.4 in the appendix shows the transition matrices and the corresponding steady state probability vectors for both estimation techniques (3-state Markov regressions versus direct estimation from the GSADF test). The results provide further evidence in favour of the GSADF steady state probabilities. An analysis of the unrestricted transition matrices and steady state probabilities reveals that the probability of the random walk or efficient state is greatly underestimated and would suggest significant deviation from the efficient market hypothesis. In fact, the highest efficient markets duration under the unrestricted regressions is for copper where the efficient market operates for 57% of the time period covered by the copper price series. Thus, naïve estimation often makes identification of states difficult.

In contrast the Tin market is efficient only 39% of the time under the unrestricted regression. There are also frequent switches from the stationary state to the explosive state which is implausible (though not impossible). It is difficult to differentiate across the remaining two states. The transition or steady state probabilities are not suggestive; the parameter estimates and the smoothed state probability graphs (not provided but available upon request) do not help either so it becomes difficult to distinguish between the explosive and the random walk states for other metals.

This is in contrast to what we would have expected in the Tin market. Up until 1985, the International Tin Council, a consortium of 22 tin producing countries, played a stabilizing role in the market (Mallory, 1990). When the International Tin Agreement collapsed in 1985, there was some uncertainty in the commodity markets but this did not last long enough to warrant a high incidence of inefficiency. For Nickel and Aluminium the estimated model suggested two explosive states and a random walk state; there is no mean reverting state. Thus, the 3-state Markov state auto-regression with no restrictions does not allow us to distinguish different states and comment on market efficiency.

The results from our restricted regression (where we impose a transition probability matrix) on the other hand provide a more consistent and intuitive picture. Steady state probabilities for the efficient state are higher relative to their unrestricted counterpart, bubble state durations are low while the stationary/mean-reverting state durations tend to vary across different metals. Some metals are more efficient than others with the efficient state operating over 75% of the time. These include Lead (89%), Tin (90%) and Zinc (79%). For these relatively more efficient markets, the stationary state occurs less often.

The least efficient commodity is Nickel according to this analysis where the efficient state prevails for only 59% of the time (although this is still higher than the unrestricted estimate – 52%). Incidences of explosiveness are high in only 2 of the 6 metals under consideration: Aluminium (14%) and Nickel (15%). These probabilities are much lower than the unrestricted estimates. At the same time, these two metals also show a high incidence of mean reversion, 18% and 26% respectively. This suggests that explosive states are counteracted by long stationary periods. Looking at the descriptive statistics of the series, the high kurtosis for the two series (8.57 for Aluminium and 13.86 for Nickel) are indicative of the extremes the series reach and are reflected in our results.

Overall we obtain intuitive results using the restricted regression compared to the unrestricted regression. Thus, our methodology offers better intuition and a deeper insight into the efficiency of these markets as compared to estimates obtained from unrestricted three state Markov regressions.

3.8 Conclusion

Using a simple partial equilibrium framework, we have provided a motivation for using Markov switching regression to estimate commodity prices. The model emphasizes the role of inventory which behaves differently in different states. The two states we are particularly interested are a mean reverting state and an explosive or bubble state. In chapter 5 we provide a more structural model that also arrives at similar conclusions.

While the framework suggests the use of a Markov-state regression, we find that using naïve Markov-state regressions without using prior information on switches does not lead to particularly accurate results. For the case of Copper and other non-ferrous metals we find that the unrestricted Markov-state regressions often suggest high smoothed probabilities for the explosive state when no explosiveness is found in the actual series. In contrast, when the transition matrix is estimated directly using recursive Augmented Dickey Fuller tests, the smoothed state probabilities tend to agree with what we observe in the price series i.e. the probability for the explosive state is high when the GSADF test detects a bubble. This also provides us with a metric for market efficiency.

The analysis was further extended to assess the efficiency of the non-ferrous metals market. This requires estimation of a three state regression with the additional state being a mean reverting or stationary state. In order to date the mean reverting state we again relied on the GSADF test and used the left sided Dickey Fuller critical values. The results from our 3 state analysis indicated that with the exception of Nickel and Aluminium, non-ferrous metal markets behaved largely efficiently with some periods of explosiveness and mean reversion. While Nickel and Aluminium also stay in the random walk state more than 50% of the time, there are large periods of deviation from the efficient state which is due to high volatility in these markets. We also found that unrestricted estimation of transition probabilities can often lead to erroneous conclusion and lead one to believe that metal markets are inefficient.

Thus, the framework we have provided can be used to estimate commodity prices or can be employed to understand market efficiency. Furthermore, the smoothed Markov-state probabilities indicate when these states may have occurred. Further research may be carried out on the efficient markets scenario. The framework we have provided can help us understand why econometric methods are unable to detect instances of explosiveness or mean reversion in a market which displays all three states highlights above. However, the technique does have its limitations.

While we can introduce additional states in the model and measure transition probabilities, it will be difficult to identify more than 3 states using econometric tests such as the ones we have employed here. While using a 95% sequence to identify explosiveness and a 5% sequence to identify mean reversion is intuitively appealing, additional states may be more difficult to explain. Secondly, we

have also limited the form of the auto-regressive parameters and have employed the use of a switching drift term. In the following chapter, we introduce a different metric for market efficiency which also allows us to specify models differently.

APPENDIX 3A: Conditions for a steady-state distribution for a switching AR-1 Process

In this section we outline the conditions that need to be satisfied by the commodity price in order to have a steady-state distribution. We do not calculate the moments. The interested reader is referred to Chapter 4 for an analysis of moments. The following analysis follows Knight et al's (2011) terminology closely. The reduced form of the commodity price process is:

$$P_t = \phi_{t-1}P_{t-1} + \eta_t \text{ where } \eta_t \sim N(0, \sigma_{\eta_t}^2) \tag{3.A1}$$

Through repeated substition, the solution of the above process becomes:

$$P_t = \sum_{n=1}^{\infty} S_n(t)\eta_{t-n} + \eta_t, \text{ where } S_n(t) = \prod_{m=1}^{n} \phi_{t-m}$$

It follows that as $n \to \infty$:

$$\lim_{n\to\infty} \frac{1}{n}\ln|S_n(t)| = E[\ln|\phi_{t-m}|] \tag{3.A2}$$

From Knight et al, we know that (3.A1) has a stationary solution if

$$E(\ln|\phi_{t-m}|) < 0 \text{ or } E\left(\ln\left|1+\frac{\beta+\gamma}{\alpha_t}\right|\right) < 0 \tag{3.A3}$$

We first consider the two-state scenario. For the above condition to be satisfied, if we have one explosive state i.e. $\phi_t > 1$ which implies that $\alpha_t > 0$, this necessarily requires that in the second state $\phi_t < 1$ so that we can have a situation where $E(\ln|\phi_{t-m}|) < 0$. It follows from 3.A3 that this will require: $-1 < \frac{\beta+\gamma}{\alpha_t} < 0$. If we label the switching parameter in state 1, α_1, and the switching parameter in state 2, α_2, we have the following bounds for the two paramaters:

$$\alpha_1 > 0$$

$$\alpha_2 + \beta + \gamma < 0$$

If the steady state probability of state 1 is π_1 and the steady state probability for state 2 is π_2 and $\pi_1 + \pi_2 = 1$, the following condition needs to be satisfied for P_t to have a steady state distribution:

$$\pi_1 \ln\left|1+\frac{\beta+\gamma}{\alpha_1}\right| + \pi_2 \ln\left|1+\frac{\beta+\gamma}{\alpha_2}\right| < 0 \tag{3.A4}$$

If we now consider a 3rd state, i.e. a random walk or efficient state, the results can be extended further. For the commodity price process to behave like a random walk we require that α_3 is large and

negative so that $\left|1 + \frac{\beta + \gamma}{\alpha_3}\right| \approx 1$ or $\ln\left|1 + \frac{\beta + \gamma}{\alpha_3}\right| \approx 0$. If π_3' represents the probability of state 3 and $\pi_1' + \pi_2' + \pi_3' = 1$ ($\pi_1' \neq \pi_1$ and $\pi_2' \neq \pi_2$) condition A4 becomes:

$$\pi_1' \ln\left|1 + \frac{\beta + \gamma}{\alpha_1}\right| + \pi_2' \ln\left|1 + \frac{\beta + \gamma}{\alpha_2}\right| + \pi_3' \ln\left|1 + \frac{\beta + \gamma}{\alpha_3}\right| < 0 \qquad (3.A5)$$

However, since $\ln\left|1 + \frac{\beta + \gamma}{\alpha_3}\right| \approx 0$, the above condition simplifies to:

$$\pi_1' \ln\left|1 + \frac{\beta + \gamma}{\alpha_1}\right| + \pi_2' \ln\left|1 + \frac{\beta + \gamma}{\alpha_2}\right| < 0 \qquad (3.A6)$$

Thus, in order to have a steady state distribution we required the commodity price to exhibit stationary behaviour for a certain amount of time. If we just consider an explosive and a random walk state, we will not obtain a steady-state distribution.

APPENDIX 3B – Tables and Figures

Table 3.1 – AR(1) with a Dummy for the bubble state - Copper

Dependent Variable: COPPERDIFF
Method: Least Squares
Date: 05/19/15   Time: 11:59
Sample (adjusted): 1959M12 2014M03
Included observations: 652 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
COPPER(-1)	-0.002376	0.002872	-0.827271	0.4084
BUBBLECOPP	209.8690	45.15446	4.647803	0.0000

R-squared	0.030805	Mean dependent var	9.141089
Adjusted R-squared	0.029314	S.D. dependent var	241.4073
S.E. of regression	237.8427	Akaike info criterion	13.78416
Sum squared resid	36769959	Schwarz criterion	13.79790
Log likelihood	-4491.636	Hannan-Quinn criter.	13.78949
Durbin-Watson stat	1.328374		

Table 3.2 – Markov-Switching Parameter Estimates for ϕ_t

Commodity	State 1 – Unrestricted	State 2- Unrestricted	State 1 – Restricted	State 2 – Restricted
Copper	0.0001 (0.00)	0.0011 (0.00)	0.0002 (0.00)	0.0011 (0.00)
Aluminium	0.0000 (0.00)	0.002913 (0.00)	-0.0002 (0.00)	0.0012 (0.00)
Lead	-0.0007 (0.00)	0.0015 (0.00)	-0.0007 (0.00)	0.0017 (0.00)
Nickel	-0.0010 (0.00)	0.0057 (0.00)	0.0000 (0.00)	0.0005 (0.00)
Tin	-0.0001 (0.00)	0.0007 (0.00)	0.0000 (0.00)	0.0009 (0.00)
Zinc	-0.0001 (0.00)	0.0007 (0.00)	0.0002 (0.00)	0.0011 (0.00)

The empirical model we estimate is: $\Delta P_t = \alpha_t + \phi_t P_{t-1} + \eta_t$, where P_t is in logs; thus, we estimate a return regression.

Table 3.3 – Restricted and unrestricted transition probability matrices

Commodity	Unrestricted transition matrix	Unrestricted steady state probabilities	Restricted transition matrix	Restricted steady state probabilities
Copper	$\begin{pmatrix} 0.9637 & 0.0363 \\ 0.0754 & 0.9246 \end{pmatrix}$	$\begin{pmatrix} 0.6849 \\ 0.3151 \end{pmatrix}$	$\begin{pmatrix} 0.9910 & 0.0090 \\ 0.1613 & 0.8387 \end{pmatrix}$	$\begin{pmatrix} 0.9472 \\ 0.0528 \end{pmatrix}$
Aluminium	$\begin{pmatrix} 0.8679 & 0.1321 \\ 0.0426 & 0.9574 \end{pmatrix}$	$\begin{pmatrix} 0.2438 \\ 0.7562 \end{pmatrix}$	$\begin{pmatrix} 0.9906 & 0.0094 \\ 0.2174 & 0.7826 \end{pmatrix}$	$\begin{pmatrix} 0.9586 \\ 0.0414 \end{pmatrix}$
Lead	$\begin{pmatrix} 0.9438 & 0.0562 \\ 0.0637 & 0.9363 \end{pmatrix}$	$\begin{pmatrix} 0.5314 \\ 0.4686 \end{pmatrix}$	$\begin{pmatrix} 0.9844 & 0.0156 \\ 0.1159 & 0.8841 \end{pmatrix}$	$\begin{pmatrix} 0.8814 \\ 0.1186 \end{pmatrix}$
Nickel	$\begin{pmatrix} 0.8941 & 0.1059 \\ 0.0566 & 0.9434 \end{pmatrix}$	$\begin{pmatrix} 0.3483 \\ 0.6517 \end{pmatrix}$	$\begin{pmatrix} 0.9869 & 0.0131 \\ 0.0492 & 0.9508 \end{pmatrix}$	$\begin{pmatrix} 0.7897 \\ 0.2103 \end{pmatrix}$
Tin	$\begin{pmatrix} 0.9417 & 0.0583 \\ 0.0502 & 0.9498 \end{pmatrix}$	$\begin{pmatrix} 0.4626 \\ 0.5374 \end{pmatrix}$	$\begin{pmatrix} 0.9920 & 0.0080 \\ 0.2000 & 0.8000 \end{pmatrix}$	$\begin{pmatrix} 0.9615 \\ 0.0385 \end{pmatrix}$
Zinc	$\begin{pmatrix} 0.9728 & 0.0272 \\ 0.0466 & 0.9534 \end{pmatrix}$	$\begin{pmatrix} 0.6315 \\ 0.3685 \end{pmatrix}$	$\begin{pmatrix} 0.9903 & 0.0097 \\ 0.2000 & 0.8000 \end{pmatrix}$	$\begin{pmatrix} 0.9537 \\ 0.0463 \end{pmatrix}$

State 1 represents the case where $\phi_t \leq 1$ whereas state 2 represents the case where $\phi_t > 1$. The restricted transition matrix was estimated through the GSADF test while the unrestricted transition matrix was estimated directly from the MSAR algorithm.

Table 3.4 – Efficient markets test using a 3 state estimation

Commodity	Unrestricted transition probabilities – 3-state estimation			Unrestricted - Steady state probabilities	Restricted transition probabilities using GSADF sequences			Restricted – Steady state probabilities
Copper	$\begin{bmatrix}0.9531 & 0.0391 & 0.0078\\0.0345 & 0.8436 & 0.1220\\0.1436 & 0.2575 & 0.5989\end{bmatrix}$			$\begin{bmatrix}0.5697\\0.3215\\0.1088\end{bmatrix}$	$\begin{bmatrix}0.9586 & 0.0109 & 0.0305\\0.1613 & 0.8387 & 0\\0.0741 & 0 & 0.9259\end{bmatrix}$			$\begin{bmatrix}0.6760\\0.0457\\0.2783\end{bmatrix}$
Aluminium	$\begin{bmatrix}0.8873 & 0.1127 & 0\\0.0474 & 0.9214 & 0.0312\\0 & 0.0869 & 0.9131\end{bmatrix}$			$\begin{bmatrix}0.2365\\0.5619\\0.2016\end{bmatrix}$	$\begin{bmatrix}0.9742 & 0.0172 & 0.0086\\0.0808 & 0.9192 & 0\\0.0325 & 0 & 0.9675\end{bmatrix}$			$\begin{bmatrix}0.6768\\0.1441\\0.1791\end{bmatrix}$
Lead	$\begin{bmatrix}0.9338 & 0.0453 & 0.0209\\0.0315 & 0.8413 & 0.1272\\0.2554 & 0.2907 & 0.4540\end{bmatrix}$			$\begin{bmatrix}0.5547\\0.3440\\0.1013\end{bmatrix}$	$\begin{bmatrix}0.9855 & 0.0097 & 0.0048\\0.1316 & 0.8684 & 0\\0.1034 & 0 & 0.8966\end{bmatrix}$			$\begin{bmatrix}0.8927\\0.0656\\0.0417\end{bmatrix}$
Nickel	$\begin{bmatrix}0.9016 & 0 & 0.0984\\0.00 & 0.9890 & 0.0110\\0.1678 & 0.0320 & 0.8002\end{bmatrix}$			$\begin{bmatrix}0.3036\\0.5183\\0.1781\end{bmatrix}$	$\begin{bmatrix}0.9688 & 0.0173 & 0.0134\\0.0690 & 0.9310 & 0\\0.0306 & 0 & 0.9694\end{bmatrix}$			$\begin{bmatrix}0.5897\\0.1521\\0.2582\end{bmatrix}$
Tin	$\begin{bmatrix}0.8782 & 0.1154 & 0.0064\\0.0829 & 0.8550 & 0.0621\\0 & 0.0731 & 0.9269\end{bmatrix}$			$\begin{bmatrix}0.2627\\0.3860\\0.3513\end{bmatrix}$	$\begin{bmatrix}0.9856 & 0.0126 & 0.0018\\0.1750 & 0.8250 & 0\\0.0500 & 0 & 0.9500\end{bmatrix}$			$\begin{bmatrix}0.9025\\0.0650\\0.0325\end{bmatrix}$
Zinc	$\begin{bmatrix}0.9521 & 0 & 00479\\0.1297 & 0.7307 & 0.1396\\0 & 0.1942 & 0.8050\end{bmatrix}$			$\begin{bmatrix}0.5317\\0.1962\\0.2727\end{bmatrix}$	$\begin{bmatrix}0.9796 & 0.0111 & 0.0093\\0.1220 & 0.8537 & 0.0243\\0.0583 & 0 & 0.9417\end{bmatrix}$			$\begin{bmatrix}0.7893\\0.0599\\0.1509\end{bmatrix}$

State 1 represents the case where $\phi_t = 1$, state 2 represents the case where $\phi_t > 1$, whereas state 2 represents the case where $\phi_t < 1$. The restricted transition matrix was estimated through the GSADF test while the unrestricted transition matrix was estimate directly from the MSAR algorithm.

Figure 3.1 Copper GSADF Test
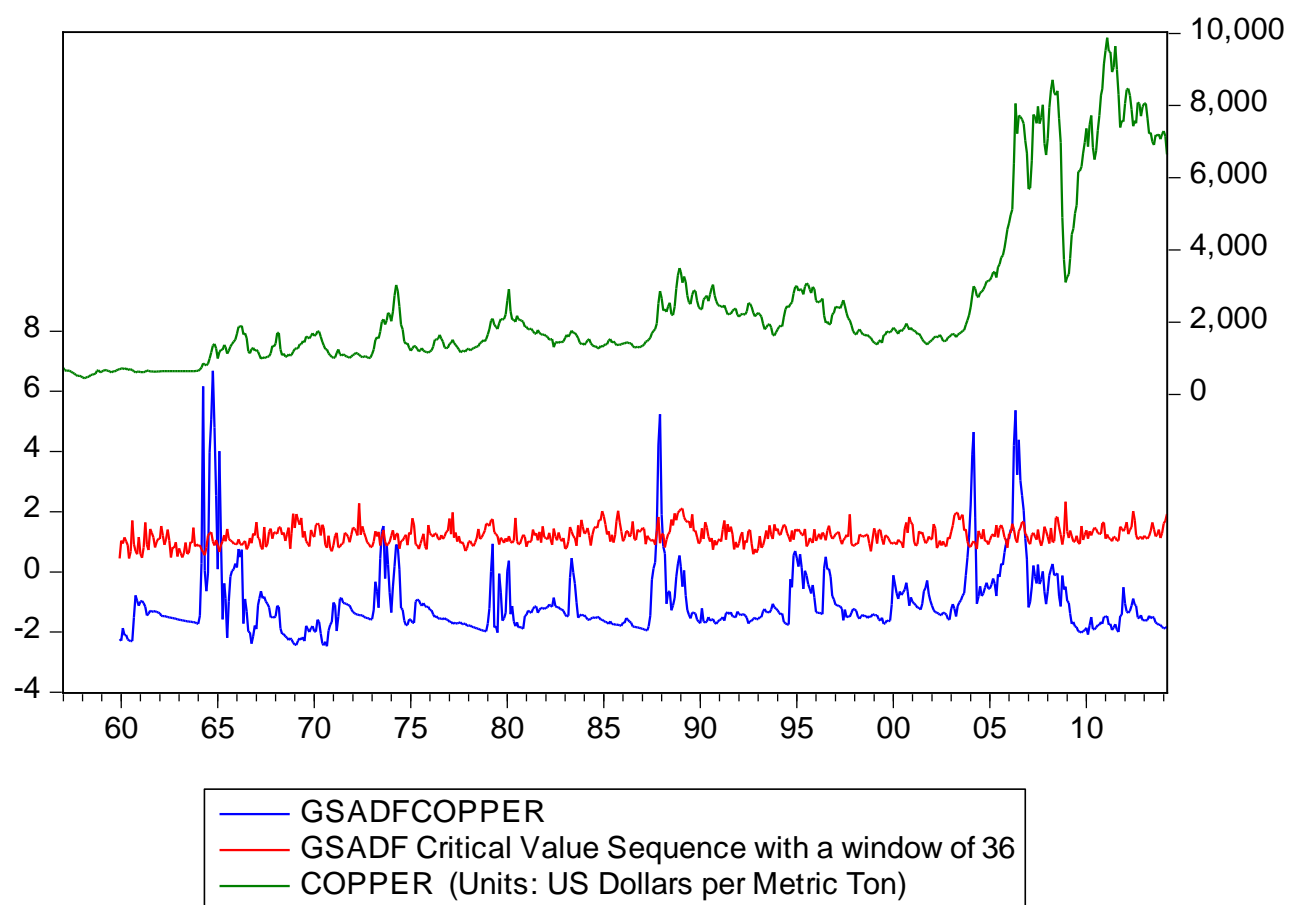
Figure 3.2 Copper MSAR (Unrestricted Markov State Probabilities) - 2 states

Smoothed Markov-State Probabilities - Unrestricted (Copper)
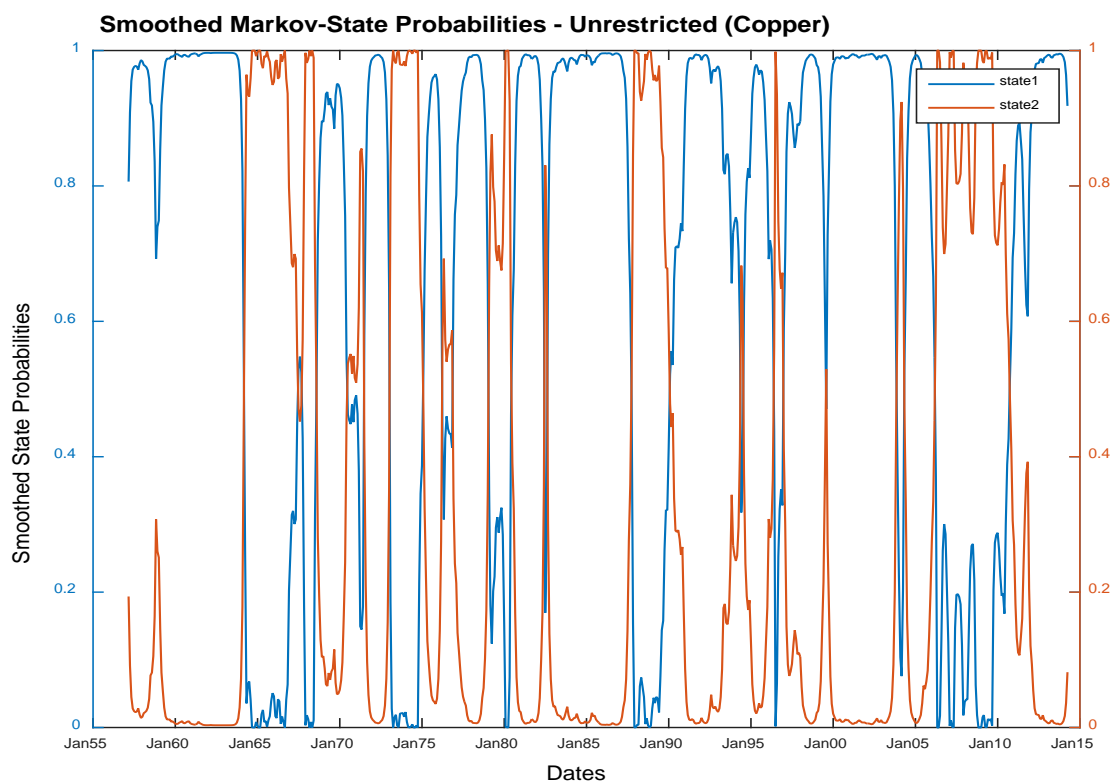
Figure 3.3 – Copper MSAR (Restricted Markov-State Probabilities) – 2 states
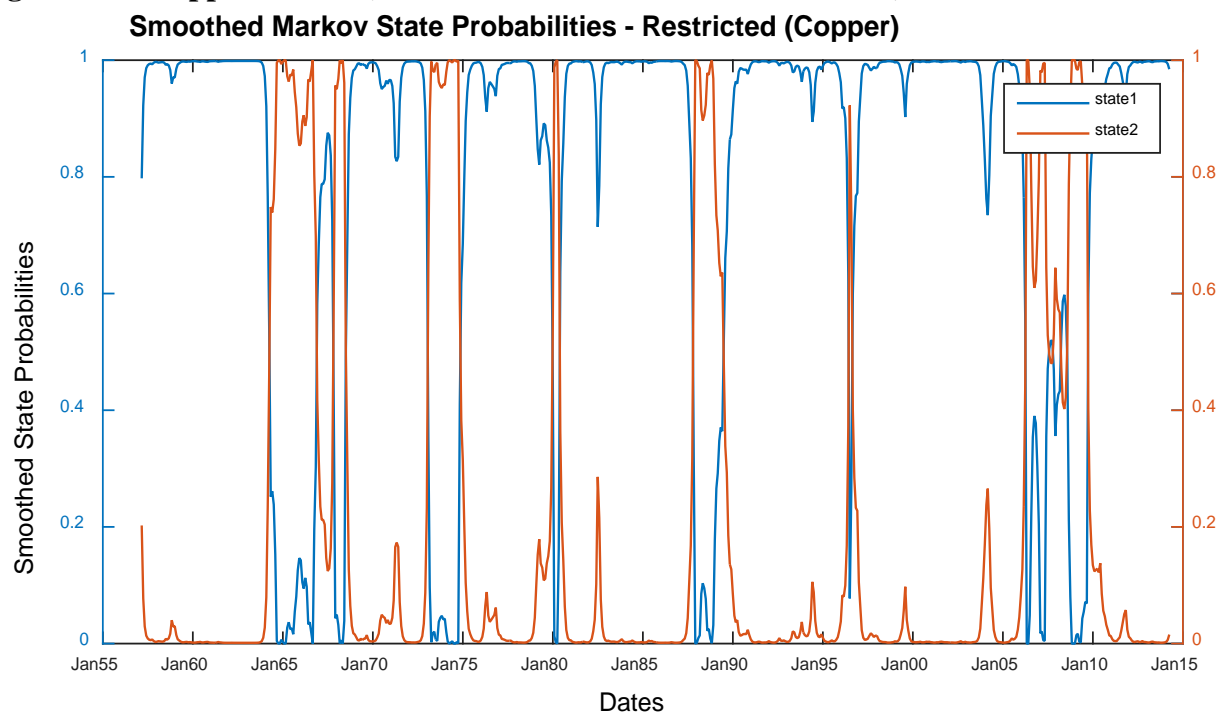
Smoothed Markov State Probabilities - Restricted (Copper)

Figure 3.4

GSADF sequence for Copper
GSADF Critical Sequence at 95%
GSADF Critical Sequence at 5%
COPPER (Units: US Dollars per Metric Ton)

Figure 3.5

Smoothed Markov-State Probabilities - Unrestricted (Copper)

Figure 3.6

Smoothed Markov-State Probabilities - Restricted (Copper)

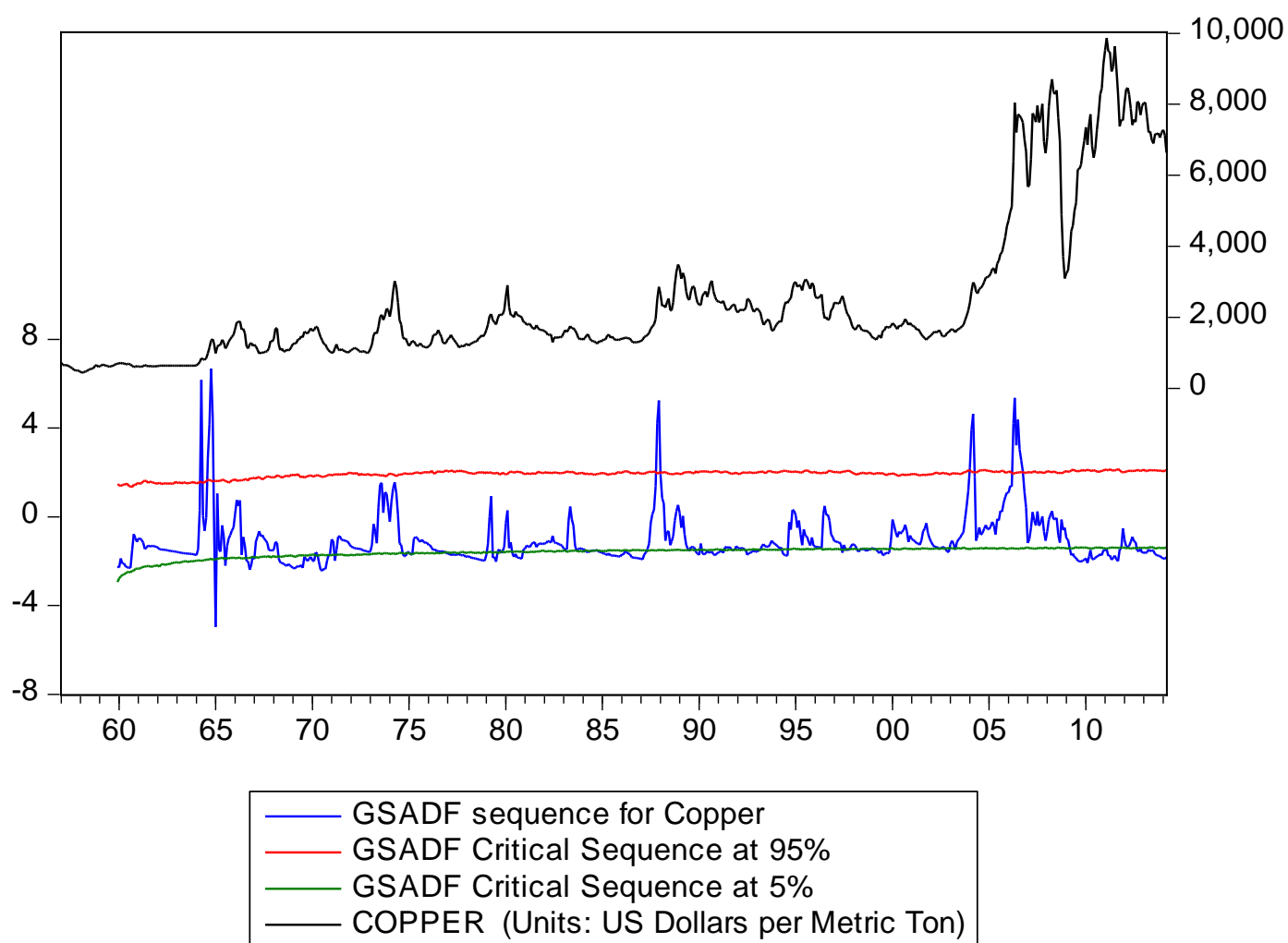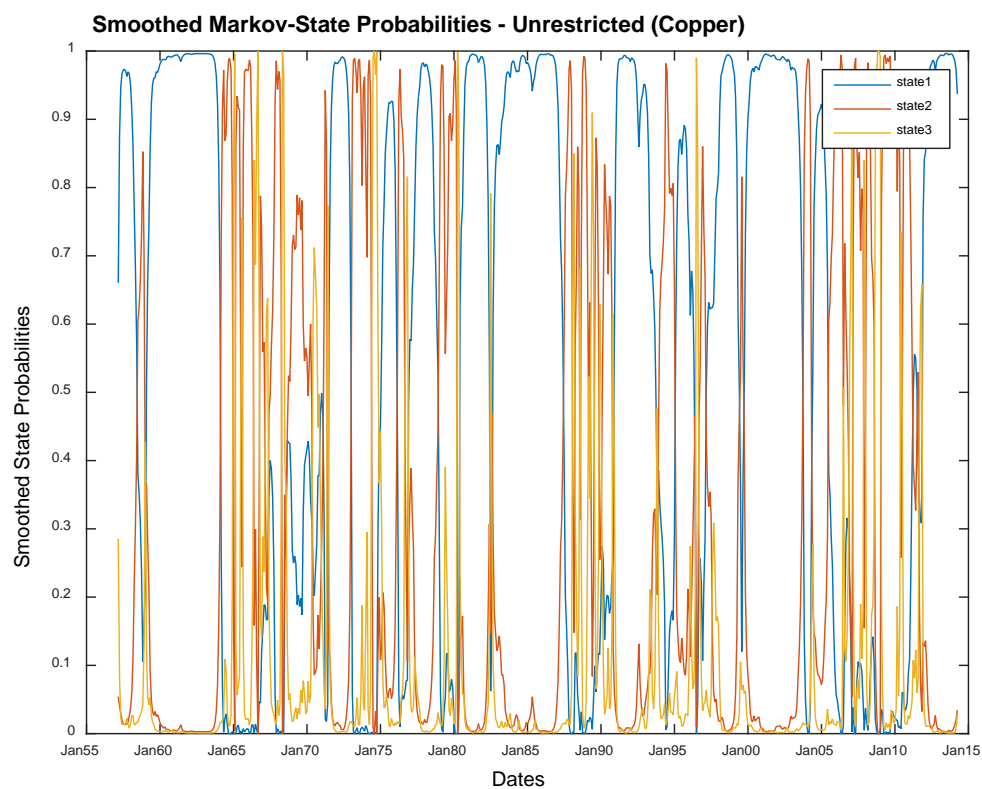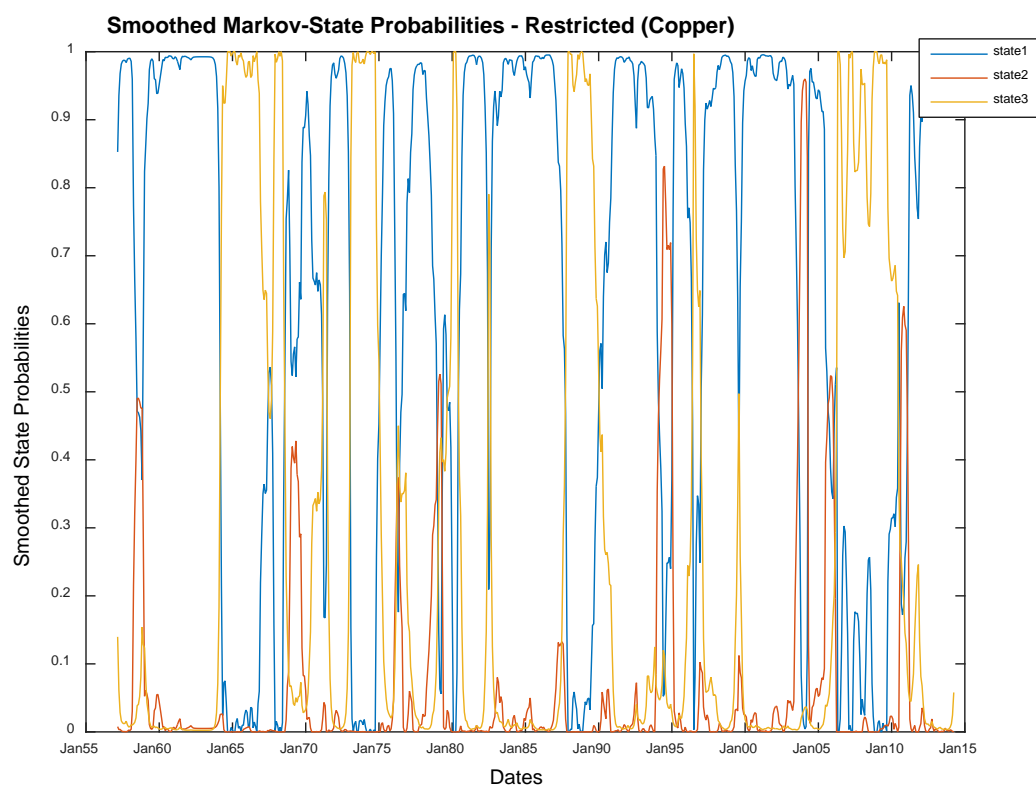# CHAPTER 4: *What Proportion of Time is a particular Market inefficient?...A Method for analysing the frequency of market efficiency when equity prices follow Threshold Autoregressions.*

In this chapter we generalize existing results for threshold autoregressive models, first presented in Knight and Satchell (2011) for the existence of a stationary process and the conditions necessary for the existence of a mean and a variance; we also present formulae for these moments. Using a simulation study we explore what these results entail with respect to the impact they can have on tests for detecting bubbles or market efficiency. We find that bubbles are easier to detect in processes where a steady state stationary distribution does not exist. Furthermore, by assuming that asset returns follow threshold autoregressive forms, we explore how these models may enable us to identify how often asset markets are inefficient. We find, unsurprisingly, that the fraction of time spent in an efficient state depends upon the full specification of the model; the notion of how efficient a market is, in this context at least, a model-dependent concept. However, our methodology allows us to compare efficiency across different asset markets. In doing so we provide a metric for market efficiency that allows for a more general model specification than the metric in chapter 3.

4.1 Introduction

The tendency of asset prices to go through locally explosive and mean reverting states is well documented and has intrigued both theoretical and empirical economists. Regime switching models, such as the ones introduced by Goldfeld and Quandt (1973), Tong (1978) and Hamilton (1989) have often been employed to empirically estimate asset prices with regime changes. These models include hidden Markov-state models as well as Threshold autoregressive models. Hansen (2011) provides an important review of applications of threshold autoregressive models in economics.

Knight and Satchell (2011) study the steady state properties of asset prices that are estimated using threshold auto-regressive models. Their article formalises necessary and sufficient conditions for the existence of a stationary distribution for regime-switching threshold models with 2 states. Analytical expressions for the mean, variance, co-variance and the distribution are also derived for the 2 states case. While Knight et al carry out most of their analysis under the assumption of an independent and

identically distributed trigger variable they also consider the case of a threshold auto-regressive model (TAR henceforth) with a Markov trigger.

The current chapter builds on their results and generalizes the conditions required for a TAR (1) model. We only consider the case where the TAR model is driven by an independent and identically distributed (i.i.d) exogenous variable that triggers regime-switches. We derive analytical expressions for the mean and variance of these models, noting the conditions that need to be satisfied for their existence. The two moments are derived with a switching drift, with a constant drift and with no drift.

Instead of calling a market universally efficient or inefficient, using TAR models we can treat market efficiency as a state. By considering efficiency as one of several states, we make several contributions to the literature on market efficiency. We contribute to the bubble testing literature by carrying out a simulation study which compares the power of bubble detection tests in situations where the stationary distribution conditions are satisfied against situations where they are not satisfied. Here, the explosive or bubble state is one state of a multi-state price process.

Evans (1991) pointed out in his seminal study that bubble detection tests are less useful when an observed series contains multiple instances of collapsing bubbles. His study showed that such tests lose power when the number of bubbles and collapses in a series increases. A number of studies have attempted to address this criticism through alternative methodologies. The most notable ones among them are Hall et al (1999) and Phillips et al (2013). Hall et al used a Markov-state regime switching model to estimate the probability of an asset being in an explosive state. Phillips et al on the other hand have devised a recursive procedure using the Augmented Dickey Fuller test which allows them to not only test for explosiveness but also date these bubbles. The GSADF test as it is now called has proven to be popular with macroeconomists and financial economists. Our simulations show that while this test is statistically powerful, in empirical application it has its limitations.

Using both i.i.d and Markov-switching triggers, we show that when a time series resembling an asset price fails to satisfy the conditions for a stationary distribution, the GSADF test has high power. On the other hand, the power of the GSADF test falls considerably when the process has a stationary distribution even though locally explosive regimes continue to be present. Thus, our simulation study builds on our theoretical results and further elaborates on observations made in Knight et al (2014), who outline reasons for the failure of bubble detection tests when a series has a stationary distribution.

Our analysis provides a limiting feature of the GSADF test as the test is premised on a process not having a mean reverting state. Phillips et al show that the GSADF test has higher power than other alternative bubble detection test; thus, we contend that these results should have external validity for other bubble tests. Furthermore, we note that the power of the GSADF test increases the farther the

explosive regime parameter is from unity. This observation is also supported by the formulae we derive in our workings and is discussed in the relevant section.

Finally and perhaps most importantly, we make a further contribution to the market efficiency literature by providing a methodology that may be used to estimate how often an asset market is efficient and also allows us to compare efficiency across different markets (and thereby gives the chapter its name). As we discussed in Chapter 1, the efficient market hypothesis is perhaps the most well-known as well as the most divisive hypothesis in economics. While economists were aware of market efficiency for a very long time before him, Fama (1965) was the first to conceptualize market efficiency in his seminal article on stock prices where he concluded that stock market prices followed a random walk.

Since then a large number of economists have contributed to this literature with both proponents and opponents of the hypothesis contributing. Seminal contributions have been made to this literature by Samuelson (1965), De Bondt and Thaler (1985), Marsh and Merton (1986), Shiller (2000) among others (see Chapter 1 for further details). The literature around the efficient market hypothesis is vast enough to cover several volumes; thus, we refer the interested reader to Sewell's (2011) article which provides a brief but useful chronological history of the hypothesis and its evolution. The hypothesis itself states that for a given information set I_t, systematic gains cannot be made by trading on the information set alone. In fact, it may be argued that the threshold auto-regressive model literature and the bubble testing literature is a subset of the efficient markets literature as the techniques developed to estimate asset prices have often been discussed in the context of efficient markets.

Note that in this chapter whenever we mention market efficiency we are referring to the weak form of market efficiency which states that returns cannot be predicted based on prior information i.e. the impact of prior prices is already reflected in the current price. In econometric terms this suggests that asset price follow a random walk with (or without) drift i.e. $P_t = \alpha + P_{t-1} + \epsilon_t$. In the context of this research, thus, a deviation from efficiency should be understood as a deviation from the random walk process or any deviation from the specification above. Instead of arguing for or against the efficient market hypothesis we recognize that although markets may be mostly weak form efficient, they can deviate from efficiency for significant periods of time. Semi-strong or strong form efficiency may not be easy to analyse with our methodology. Thus, we introduce a new notion of market efficiency; an asset market may be efficient sometimes and not efficient at other times. Classifying asset markets as either efficient or inefficient would be tantamount to oversimplifying a rather complex set of markets.

Given the setting above, we provide an estimation methodology for asset markets where market efficiency is one of several states. Our estimation methodology aims at identifying how long periods of efficiency and inefficiency last. To the best of our knowledge this research would be the first to

provide a metric for market efficiency using threshold auto-regressions. Historically, research has taken a binary view towards market efficiency and has been based around the presence or absence of market efficiency taken on average over a sample period; our metric provides a more detailed view. We believe that markets can undergo efficient as well as inefficient states and our methodology helps us determine and estimate how long such states last.

We provide an illustrative empirical application of our methodology through estimating a TAR(1) model for the S&P500 and FTSE 100 stock market indices where the parameter switches due to an exogenous trigger variable. We estimate the TAR(1) using a constant drift, a regime-switching drift and no drift. Our results indicate that the inclusion of a drift term, particularly a regime-switching drift term, reduces the impact of the regime-switching slope parameter. A switching drift term is able to explain changes in the return process and thus, the high variance observed during explosive periods. With a switching drift term, markets appear to be more efficient than with no drift as we are unable to reject the random walk hypothesis for a number of coefficients. This observation indicates that in the context of regime-switching models, our metric for determining market efficiency depends on model specification. However, our methodology does allow us to compare efficiency across different markets for the same model specification.

Other measures of efficiency based on trading volumes or number of informed traders may be used to gain estimates of market efficiency. This would require estimating a trading model for the former and a heterogeneous agent model for the latter. We argue that our estimation methodology is much simpler and the data required for estimation (i.e. prices or returns) are much easily available compared to trading volumes and private information of traders. The methodology may be used for forecasting purposes; however, no data testing on forecasting is carried out as it is goes beyond the scope of our research which primarily aims at specifying a metric for market efficiency.

In summary, our main contributions to the literature are a generalization of conditions required for a threshold auto-regressive model to have a mean and variance when a steady state distribution exists, a simulation study of how bubble tests behave when processes do or do not have steady state distributions and an empirical methodology for analysing market efficiency through threshold autoregressive models driven by an exogenous trigger variable. Section 4.2 generalizes the results in Knight et al (2011) to a finite number of regimes and presents results on the mean and variance of TAR(1) models. Section 4.3 explains the results of the simulation study using both i.i.d and Markov chain exogenous triggers. Section 4.4 illustrates how the model in section 4.2 may be estimated in practice and how it can be used to construct efficiency measures. Section 4.5 concludes.

4.2 Conditions for the existence of a mean and variance:

Let p_t be the price (or log price) of an asset. We assume that:

$$p_t = \psi_{t-1} + \phi_{t-1}p_{t-1} + \eta_t \text{ where } \eta_t \sim N(0, \sigma_\eta^2) \qquad \textbf{(4.1)}$$

For illustrative purposes we consider the 3 state case although the results will be applicable to a finite number of 'k' states. For k = 3 we specify values of the trigger or driving variable Z_t for which the parameters switch between values. Knight and Satchell (2011), only consider two states and a constant drift term.

Thus, for a 3-state case we have:

$$\psi_{t-1} = \alpha_1, \phi_{t-1} = \beta_1 \text{ if } -\infty < Z_{t-1} < c_1$$

$$\psi_{t-1} = \alpha_2, \phi_{t-1} = \beta_2 \text{ if } c_1 \le Z_{t-1} < c_2$$

$$\psi_{t-1} = \alpha_3, \phi_{t-1} = \beta_3 \text{ if } c_2 \le Z_{t-1} < \infty$$

c_1 and c_2 are threshold levels which trigger the switch between states. The above framework can be generalized to k intervals and k+1 constants where $c_0 = -\infty$ and $c_k = \infty$.

We need Z_{t-1} to be i.i.d to derive our results. The probability that Z_{t-1} will take a value between any two constants is assumed to be π_{j+1} i.e. $P(c_j \le Z_{t-1} < c_{j+1}) = \pi_{j+1}$ and

$$\sum_{j=1}^{k} \pi_j = 1$$

As a result there will be '2k' different parameters. We denote the regime specific parameter by α_j, β_j.

We note that conditions for the existence of a stationary distribution will be similar if Z_t followed a Markov process but the moments will be different. (Knight et al 2011 Theorem 2).

If $E(\ln|\phi_{t-m}|) = \sum_{j=1}^{k} \pi_j \ln|\beta_j| < 0$ and $\psi_{t-1} = \alpha$ (Knight et all 2011) then the TAR model given by (4.1) has the solution

$$p_t = \alpha(\sum_{n=0}^{\infty} S_n(t)) + \sum_{n=0}^{\infty} S_n(t)\eta_{t-n}, \text{ where } S_n(t) = \prod_{m=1}^{n} \phi_{t-m} \qquad \textbf{(4.2)}$$

and $S_0(t) = 1$, (Quinn,1982).

When $\psi_{t-1} = \alpha_i$ for i = 1, 2….k

$$p_t = \psi_{t-1} + \phi_{t-1}P_{t-1} + \eta_t$$

$$= \psi_{t-1} + \phi_{t-1}(\psi_{t-2} + \phi_{t-2}P_{t-2} + \eta_{t-1}) + \eta_t$$

$$= (\psi_{t-1} + \phi_{t-1}\psi_{t-2} + \phi_{t-1}\phi_{t-2}\psi_{t-3} + \cdots.) + \sum_{n=0}^{\infty} S_n(t)\eta_{t-n}$$

$$p_t = \sum_{n=0}^{\infty} \psi_{t-1-n}S_n(t) + \sum_{n=0}^{\infty} S_n(t)\eta_{t-n} \qquad (4.3)$$

It follows that as $n \to \infty$:

$$\lim_{n\to\infty} \frac{1}{n}\ln|S_n(t)| = E[\ln|\phi_{t-m}|] \qquad (4.4)$$

We note that the finiteness of the first term in (4.3) is governed by the behaviour of $S_n(t)$ but that equation (4.4) being satisfied is enough to ensure the existence of (4.3). We further note that existence of the process does not imply existence of the mean so the first term may not converge to a finite limit in expectation.

In the following sub-sections we derive the mean and variance for the general case $\psi_{t-1} = \alpha_i$ and discuss special cases i.e. $\psi_{t-1} = \alpha$ and $\psi_{t-1} = 0$.

Mean:

$$E(p_t) = E(\psi_{t-1}) + E(\phi_{t-1}p_{t-1}) \qquad (4.5)$$

$$= E(\psi_{t-1}) + E(\phi_{t-1})E(p_{t-1})$$

As the $\psi's\ and\ \phi's$ switch independently (due to Z_{t-1} being i.i.d), we have:

$$E(\phi_{t-1}) = \sum_{j=1}^{k} \pi_j\beta_j \qquad (4.6)$$

$$E(\psi_{t-1}) = \sum_{j=1}^{k} \pi_j\alpha_j \qquad (4.7)$$

It follows from (4.6) and (4.7)

$$E(p_t) = \frac{\sum_{j=1}^{k} \pi_j\alpha_j}{1 - \sum_{j=1}^{k} \pi_j\beta_j} \qquad (4.8)$$

Thus, the mean will exist if $\sum_{j=1}^{k} \pi_j \beta_j < 1$

If $\psi_{t-1} = \alpha$,

$$E(p_t) = \frac{\alpha}{1 - \sum_{j=1}^{k} \pi_j \beta_j} \qquad (\mathbf{4.9})$$

The mean is zero if there is no drift term.

<u>Variance:</u>

<u>For $\psi_{t-1} = \alpha_i$</u>

Under independence of η_t and Z_{t-1} it follows from (4.1) that:

$$Var(p_t) = Var(\psi_{t-1}) + Var(\eta_t) + Var(\phi_{t-1}p_{t-1}) + Cov(\psi_{t-1}, \phi_{t-1}p_{t-1}) \quad (\mathbf{4.10})$$

We evaluate each term in (4.10) separately

$$var(\psi_{t-1}) = E(\psi_{t-1}^2) - [E(\psi_{t-1})]^2$$

$$= \sum_{j=1}^{k} \pi_j \alpha_j^2 - \left(\sum_{j=1}^{k} \pi_j \alpha_j \right)^2 \qquad (\mathbf{4.11})$$

$$Var(\eta_t) = \sigma_\eta^2 \qquad (\mathbf{4.12})$$

$$Var(\phi_{t-1}p_{t-1}) = E(\phi_{t-1}^2)E(p_{t-1}^2) - [E(\phi_{t-1})E(p_{t-1})]^2$$

$$= E(\phi_{t-1}^2)[Var(p_{t-1}) + [E(p_{t-1})]^2] - [E(\phi_{t-1})E(p_{t-1})]^2$$

$$= var(\phi_{t-1})[E(p_{t-1})]^2 + E(\phi_{t-1}^2)Var(p_{t-1}) \qquad (\mathbf{4.13})$$

(4.13) again makes use of the fact that Z_{t-1} is an i.i.d process which implies independence between ϕ_{t-1} and p_{t-1}.

Using the definition of variance we know that:

$$var(\phi_{t-1}) = E(\phi_{t-1}^2) - [E(\phi_{t-1})]^2$$

$$= \sum_{j=1}^{k} \pi_j \beta_j^2 - \left(\sum_{j=1}^{k} \pi_j \beta_j \right)^2 \qquad (\mathbf{4.14})$$

Finally, we evaluate $Cov(\psi_{t-1}, \phi_{t-1}p_{t-1})$

$$Cov(\psi_{t-1}, \phi_{t-1}p_{t-1}) = E(\psi_{t-1}\phi_{t-1}p_{t-1}) - E(\psi_{t-1})E(\phi_{t-1})E(p_{t-1})$$

$$= E(\psi_{t-1}\phi_{t-1})E(p_{t-1}) - E(\psi_{t-1})E(\phi_{t-1})E(p_{t-1})$$

$$= Cov(\psi_{t-1}, \phi_{t-1})E(p_{t-1}) \qquad (\mathbf{4.15})$$

(4.15) relies on the independence of ψ_{t-1} and ϕ_{t-1} from p_{t-1}.

Thus, (4.11)-(4.15) allow us to calculate $Var(p_t)$

$$Var(p_t) = [Var(\psi_{t-1}) + \sigma_\eta^2 + Var(\phi_{t-1})[E(p_{t-1})]^2 + E(\phi_{t-1}^2)Var(p_{t-1})$$
$$+ Cov(\psi_{t-1}, \phi_{t-1})E(p_t) \qquad (\mathbf{4.16})$$

Rearranging (4.16) and recognizing that if p_t has a stationary distribution $Var(p_{t-1}) = Var(p_t)$ and $E(p_{t-1}) = E(p_t)$

$$Var(p_t)\left(1 - E(\phi_{t-1}^2)\right) = Var(\psi_{t-1}) + \sigma_\eta^2 + Var(\phi_{t-1})[E(p_t)]^2 + Cov(\psi_{t-1}, \phi_{t-1})E(p_t)$$

Thus,

$$Var(p_t) = \frac{\left[Var(\psi_{t-1}) + \sigma_\eta^2 + Var(\phi_{t-1})[E(p_t)]^2 + Cov(\psi_{t-1}, \phi_{t-1})E(p_t)\right]}{1 - E(\phi_{t-1}^2)}{}_4 \qquad (\mathbf{4.17})$$

The condition for the existence of a finite variance with switching regimes is thus,

$$E(\phi_{t-1}^2) = \sum_{j=1}^{k} \pi_j \beta_j^2 < 1 \qquad (\mathbf{4.18})$$

The expression in (4.17) suggests that the variance of the price series is increasing in the variance of the drift parameter ψ_{t-1}, the variance of the error term σ_η^2, the variance of the coefficients ϕ_{t-1}, the expected value of p_t (which includes the absolute value of the drift) and the Covariance between the drift and coefficient parameters. The latter will be positive since we assume that the same exogenous parameter causes a regime switch in both the drift and the coefficient parameter. If we have regimes with slope coefficients deviating far from unity (a case we will be interested in when considering the efficient market hypothesis), we will get a much higher variance for the price series. Similarly, if there is a large drift term, it can lead to the variance of price being high. This also gives us an early indication that model specification may be an important determinant in analysing asset price series.

---

[4] $Var(\psi_{t-1}), Var(\phi_{t-1})$ and $E(\phi_{t-1}^2)$ are as before.
$Cov(\psi_{t-1}, \phi_{t-1}) = \sum_{j=1}^{k} \pi_j \alpha_j \beta_j - (\sum_{j=1}^{k} \pi_j \alpha_j)(\sum_{j=1}^{k} \pi_j \beta_j)$

Given the amount of variation in an observed set of series, different specifications will lead to the variation being captured by different parameters. Indeed, this is what we observe in Section 4.4.

Now we consider the case where the drift is constant i.e. $\psi_{t-1} = \alpha$.

A constant drift implies $Var(\psi_{t-1}) = 0$ and $Cov(\psi_{t-1}, \phi_{t-1}) = 0$. Therefore, the expression for the variance of the asset price with a constant drift term reduces to:

$$Var(p_t) = \frac{[\sigma^2 + Var(\phi_{t-1})[E(p_{t-1})]^2]}{1 - E(\phi_{t-1}^2)} \qquad (\mathbf{4.19})$$

With a constant drift term the variance of the process relies on the variance of the error term, the variance of the coefficient parameter and the absolute value of the expected price which itself is a function of the drift term; thus, the variance of price is dependent on the absolute value of the drift term.

Finally, we consider the case where $\psi_{t-1} = 0$. With no drift term, (4.9) implies that $E(p_{t-1}) = 0$. Thus,

$$Var(p_t) = \frac{\sigma^2}{1 - E(\phi_{t-1}^2)} \qquad (\mathbf{4.20})$$

In the vicinity of a unit root, $E(\phi_{t-1}^2)$ is likely to be close to 1, which will lead to a very large variance for the process. Nevertheless, the variance will be finite and will exist as long as the condition in (4.18) is satisfied. Sections 4.3 and 4.4 analyse the implications of the formulae derived above. Section 4.3 considers a series of simulations to show how violations of the criterion for a stationary steady state distribution specified above and the variance of the parameter coefficient impact the ability of statistical tests to detect explosive roots or bubbles. Section 4.4 uses an illustrative empirical study to show how specifications such as (4.1) may be used to analyse the efficiency of asset markets.

4.3 Simulation Results

Our theoretical analysis has multiple applications, of which we discuss two in this Chapter. First, we carry out a simulation study to understand how tests for detecting explosiveness perform when a series is steady state stationary and when a series is not stationary. As we mentioned in the introduction to this chapter, Evans (1991) in his seminal article argued that tests for detecting explosiveness in a series had less power if there were multiple episodes of collapsing bubbles. Philips et al (2013) devised the GSADF statistic to address that criticism and showed that they could detect explosiveness even in the presence of multiple collapses. We show that while their test continues to have high power when a series does not have a steady state distribution, the power decreases

considerably when the series does satisfy the conditions for a steady state distribution. In particular, the presence of a mean reverting state in addition to explosive and efficient or random walk states, makes bubble detection difficult using the GSADF statistic.

We simulate series with a switching autoregressive parameter and a standard normal error term. We do not consider a switching drift term for our simulations as a constant drift adequately addresses the issue we wish to highlight. The simulated series takes the following form:

$$y_t = \alpha + \phi_{t-1}y_{t-1} + \epsilon_t \text{ where } \epsilon_t \sim N(0,1) \qquad (\mathbf{4.21})$$

In this context y_t can be thought of as the logarithm of a price variable so that if $\phi_{t-1} = 1$, the return is $y_t - y_{t-1} = \alpha + \epsilon_t$. In the simulations below we consider the case when $\alpha = 0, \alpha = 0.01$ and $\alpha = 0.025$.

Our simulation study considers the 3-state case for computational ease, although the results will also hold for a finite number of k-states. The switching parameter depends on the pseudo sentiment variable, Z_{t-1} which in our simulations is either a multinomial vector or a Markov chain variable. For the multinomial vector case, we select the probability with which each state occurs. Thus, Z_{t-1} is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ when in state 1, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ when in state 2 and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ when in state 3 with probabilities π_1, π_2 and π_3 respectively. The value taken by ϕ_{t-1} in a particular period depends on the parameter vector specified.

On the other hand if Z_{t-1} is a Markov chain variable, it takes on the values 1, 2 or 3 depending on which state the series is in. For the Markov chain simulations we need to specify a transition matrix instead of a probability vector. A Markov chain is more intuitive for the type of series we are concerned with as states tend to be more persistent in this case. It is also more comparable to the kind of simulations used in the literature related to tests for explosiveness or bubbles. Evans and Philips et al consider similar series in their respective articles. In addition to the process followed by the switch inducing variable Z_{t-1}, we also need to specify values for the switching parameters. Together, the switching parameter and the probability of Z_{t-1} enable us to verify if the criterion for a steady state distribution, set forth in the previous section, is satisfied.

In order to illustrate what happens when the criterion is satisfied and when it is not we carry out the GSADF test, developed by Philips, Shi and Yu (PSY henceforth), on each simulated series to check if the test is able to detect explosiveness in the series. This takes the form of a power test. Each simulated series contains periods of explosiveness, taking the series away from stationarity. Each series also has a mean reverting state which tends to make the whole series more stationary.

The GSADF statistic is being used in different areas of economics and PSY have shown it to have high power in detecting explosiveness or bubbles. The test involves estimating recursive regressions of the following form:

$$\Delta y_t = \alpha + \phi y_{t-1} + \sum_{i=1}^{k} \psi^i \Delta y_{t-i} + \epsilon_t$$

In the above regression the parameter of interest is ϕ which is estimated through expanding, rolling windows with a minimum window size specified by the researcher. We have considered the test in chapter 2 and 3, so we comment on it briefly. For each regression a right-sided unit root statistic is calculated. The supremum(sup) of all right-sided unit root statistics thus calculated is the GSADF statistic. The sup value can be compared to simulated critical values, allowing the user to comment on whether a bubble may be present in the series under consideration. We refer the interested reader to Phillips et al, 2013 for further details on the test procedure and asymptotic properties of the statistic.

4.3.1 Multinomial trigger variable:

For clarity, we indicate the specific form taken by our simulated series. When Z_{t-1} is a multinomial vector taking values $Z_{t-1,j}$, it takes the following form: $y_t = \alpha + (\sum_{j=1}^{3} \beta_j Z_{t-1,j}) y_{t-1} + \epsilon_t$

$$y_t = \alpha + \beta_1 y_{t-1} + \epsilon_t \; when \; Z_{t-1,1} = \begin{matrix} 1 \\ 0 \\ 0 \end{matrix} \; with \; probability \; \pi_1$$

$$y_t = \alpha + \beta_2 y_{t-1} + \epsilon_t \, when \; Z_{t-1,2} = \begin{matrix} 0 \\ 1 \\ 0 \end{matrix} \; with \; probability \; \pi_2$$

$$y_t = \alpha + \beta_3 y_{t-1} + \epsilon_t \, when \; Z_{t-1,3} = \begin{matrix} 0 \\ 0 \\ 1 \end{matrix} \; with \; probability \; (\pi_3 = 1 - \pi_1 - \pi_2)$$

Section 4.2 considers a similar exogenous trigger. To aid understanding, the first element of the Z_{t-1} vector indicates the mean reverting state, the second element indicates the random walk or efficient state and the third element indicates the explosive or bubble state. For each set of parameter and probability values we generated 500 simulated series and the GSADF test was conducted on each series. We simulate the state variable Z_{t-1} by randomly drawing from the vector of probabilities; this then enables us to simulate the process y_t. Each series is 1000 observations long and the minimum

window size for the GSADF test was stipulated to be 10% of the series or 100 observations. Critical values were generated separately using the MATLAB code provided by PSY. The GSADF test was conducted at the 5% level (critical values for the GSADF test at the 5% level for series of length 1000 with initial window size of 100 is 2.16 for series without drift and 2.233 with drift). Since we assume that each simulated series is explosive 10% of the time, each series exhibits the type of explosive behaviour that the GSADF test seeks to detect. The power is simply calculated by dividing the number of bubbles detected by 500 for each set of 500 simulations.

Table 4.1 (see APPENDIX) shows the results of our simulations along with the parameter values and the probability vector. Column 3 shows the value of the criterion for a steady state distribution. The criterion is said to be satisfied whenever $\sum_{j=1}^{3} \pi_j \ln |\beta_j| < 0$. We ensured that we chose a range of values so that for some values the criterion was satisfied and for other values it was not. For the multinomial vector case, we note that the power of the GSADF test is much higher when the criterion is not satisfied. We illustrate our results by considering some sets of parameter and probability values. For parameter values [0.96 1 1.05] and a probability vector [0.10 0.80 0.10] we obtain a criterion value of 0.00080 and as per our theoretical results the series should not have a stationary distribution. We see that when the criterion threshold is breached, we get a power of 17.6% from the GSADF test. Note that the existence of a stationary distribution does not guarantee the existence of moments. With two exceptions (parameter vector = [0.98 1.02 1.05] and [0.96 1 1.03]), all other sets of values do not have a mean or variance even though the distribution may exist (when the criterion is satisfied).

Contrast this with cases when the criterion is satisfied e.g. when the parameter vector is [0.95 1 1.05] and the probability vector is [0.10 0.80 0.10]. The intensity of the bubble or explosive behaviour stays the same; i.e. the bubble increases the value of the series by 10% each period and the explosive state occurs 10% of the time in the long run. We note that even with no change in the bubble state parameter value the power of the GSADF test reduces markedly down to 9.0%. This supports our theoretical results and shows that if bubbles occur in assets which may have a long run steady state distribution, they may be harder to detect.

In their article PSY do not use a mean reverting state. Their analysis is based on a random walk and a mildly explosive regime which will not satisfy the criterion and is thus, likely to result in a higher power for their test based on what we observe in our simulations. While the test has undoubtedly been useful in many applications, it is important to keep its limitations in mind particularly when it is unable to detect bubbles in an asset which may otherwise be thought to have gone through periods of explosiveness. The types of series considered by PSY are closer to the Markov-chain simulations in the following sub-section so some of the low power detected in this sub-section may be attributed to the choice of our i.i.d exogenous trigger.

We also note that the power of the test increases the farther apart from unity the explosive state is i.e. $var(\phi_{t-1})$. For instance, when we reduce $var(\phi_{t-1})$ and consider the parameter vector [0.98 1 1.02] with the same probability vector as before, the power reduces to 5.6% even though the criterion is smaller than before. We also consider the case with 1 mean reverting and two explosive states with the mean reverting state occurring 80% of the time. With the same probability vector, this set of values attained a power of only 1.2%. This is the only set of values for which both a mean and a variance exists.

Tables 4.2a and 4.2b on the other hand report results for simulations which include a constant drift term. Table 4.2a contains results for a drift of 0.01. Table 2b contains results for a drift of 0.025. The higher drift value is chosen in order to illustrate how the power of the test depends on the drift term. Note that as per the results in section 4.2, the size of the drift plays a role in determining the variance and is likely to impact the results. Recall that the variance of a series with switching-regimes and a constant drift is $\frac{[\sigma^2 + Var(\phi_{t-1})[E(y_{t-1})]^2]}{1 - E(\phi_{t-1}^2)}$ where $[E(y_{t-1})]^2 = \left[\frac{\alpha}{1 - \sum_{j=1}^{k} \pi_j \beta_j}\right]^2$ and is thus, dependant on α.

When a constant drift term is included the power of the GSADF test goes up for all sets of values compared to the case with no drift and our main results continue to hold i.e. the power is lower if the criterion for a steady state distribution is satisfied and the variance of the parameter vector is low despite the fact that the mean and variance for the set of values chosen do not exist. As per the expressions in Section 4.2 for the mean and variance of the simulated series, a higher alpha implies not just a higher mean but also a higher variance. While it is not clear from our results whether a higher α necessarily leads to a higher power for the GSADF test, the powers attained are higher than the no drift case. We do note, however, that for the case where both a mean and a variance exist a higher α leads to a higher power. The relationship between the power of the GSADF test and α stems from the impact the drift term has on the variance. A higher α results in higher variance and thus more extreme values which may make bubble detection easier.

4.3.2 Markov Chain trigger variable:

In this sub-section we consider a Markov-chain trigger variable instead of an exogenous trigger variable. Instead of a 3-dimensional multinomial vector, Z_{t-1} is now a Markov-chain variable dependant on a transition matrix; thus, we specify a transition matrix instead of a probability vector, which determines the value of Z_{t-1}, the state variable. Thus, the process becomes:

$$y_t = \alpha + \beta_1 y_{t-1} + \epsilon_t \ when \ Z_{t-1} = 1$$

$$y_t = \alpha + \beta_2 y_{t-1} + \epsilon_t when \ Z_{t-1} = 2$$

$$y_t = \alpha + \beta_3 y_{t-1} + \epsilon_t when \ Z_{t-1} = 3$$

with the transition matrix

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}$$

The probability that Z_{t-1} takes a value of 1 given that it was 1 in the previous period is p_{11}; the probability that Z_{t-1} takes a value of 2 given that it took a value of 1 in the previous period is p_{12} and so on. Thus Z_{t-1} is a one period Markov-chain variable. Simulating the series in this way has the desirable property that states tend to be more persistent compared to the multinomial case. While over the long term the duration of each state is similar to the multinomial case, as evident from the steady state probability vector, each state tends to last longer. Thus, it can be argued that series generated in this way share more properties of actual asset price series.

As in the previous section, we carry out 500 simulations for each set of values and each series consists of 1000 observations. The first step in the simulations as before is to generate the state variable Z_{t-1}, but instead of using a probability vector we generate the state variable using the transition matrix. In order to aid comparison we ensure that the transition matrix was such that the steady state probabilities of states were similar to those used in the multinomial series. One simplification made in selecting values for the transition matrix is that there are no switches from the explosive state to the mean reverting state and vice versa. Thus, whenever there is a switch from either the explosive or the mean reverting state, it is to the random walk state in the first instance. This is in line with our findings in chapter 3, where we estimated the transition matrices by estimating real commodity series using Markov-switching regresisons.

Simulating and testing Markov-chain series further strengthens the results obtained from the previous sub-section. Using Markov-chains instead of multinomial vectors, the simulated series exhibit more asset price like properties and due to state persistence we obtain higher powers for each set of values compared to the multinomial vector counterpart. Table 4.3 reports results for Markov-chain simulations below. When the criterion is set to -0.00004 the multinomial vector series have a power of 5.6% compared to 26.8% for the corresponding Markov-chain simulations. A similar pattern is observed for the remaining values. This observation may be attributed to state persistence introduced by the Markov-chain which enables detection via the GSADF test.

As noted previously, $var(\phi_{t-1})$ influences the results. When the criterion is not satisfied and the non-efficient states significantly deviate from 1, we get very high power. For example, when the parameter vector is $\begin{pmatrix} 0.98 \\ 1 \\ 1.02 \end{pmatrix}$ with a transition matrix so chosen to give a steady state probability vector $\begin{pmatrix} 0.10 \\ 0.80 \\ 0.10 \end{pmatrix}$ we find a power of 26.8% (criterion value -0.00004). Keeping the explosive state at 1.02, if the mean reverting state is made more persistent (0.99 from 0.98), the criterion is violated (0.00098) and we get a higher power for the GSADF test at 38.2%. If we increase the deviations from the random-walk state while maintaining the same steady state probabilities, the power increases further even though the value of the criterion itself does not change significantly. For example, if the parameter vector is $\begin{pmatrix} 0.96 \\ 1 \\ 1.05 \end{pmatrix}$ with the same steady state probabilities and a criterion value of 0.0008, the power increases to 73.8%.

We also consider the case when we have multiple explosive states and a mean reverting state (in this case the mean and variance of the process exist according to the criteria set in Section 4.2. Among the 2 explosive states one is more explosive than the other but both explosive states have the same steady state probability. The parameter vector is $\begin{pmatrix} 0.98 \\ 1.02 \\ 1.05 \end{pmatrix}$ with steady state probabilities $\begin{pmatrix} 0.80 \\ 0.10 \\ 0.10 \end{pmatrix}$ yielding a criterion value of -0.0093. Corresponding to these values we get a power of 25.8% for the GSADF statistic. With the multinomial regime-switching variable we observed a power of only 1.2% for the same set of values. Thus, even with two explosive states we note that if the steady state distribution criterion is satisfied, explosive behaviour may not be discernible using conventional right-sided unit root tests such as the GSADF test.

We also report results for simulations which include drift terms, $\alpha = 0.01\ and\ 0.025$). Results are reported in Table 4.4a and Table 4.4b. In the Markov-switching case we note that the size of the drift term matters significantly. With a small drift term we do not note a significant change in power compared to the case with no drift (note that the critical values for the two tests are different). For the Markov chain simulations we also note that the power of the test increases as the drift term is increased from 0.01 to 0.025 for all sets of values except one. PSY carried out the test for small deviations from random walk while our simulations include much larger deviations which explains why we observe much higher power despite including a mean reverting term. Nevertheless the results are consistent with the previous set of simulations and the same set of factors namely the value of the criterion, the existence of a drift term and the variance of the parameter vector tend to determine the power of the GSADF test.

Thus, the Power of bubble detection tests will be higher if the estimated parameters and state probabilities are such that the criterion for a stationary distribution is not satisfied. We also note that the power of such tests is higher when $var(\phi_{t-1})$ is high.

4.4. A metric for market efficiency:

We use the results in section 4.2 and 4.3 to outline a methodology that may be employed as a measure for market efficiency. Estimating different states in a process will enable us to understand the stationarity properties of an asset price series and whether the process has a steady state stationary distribution. It will also allow us to evaluate how much of the time the process spends in each of the states. While we use time as a metric for market efficiency, other potential metrics may also be employed (these could include but not be limited to metrics using trading volumes or market liquidity); however, we argue that our metric is the simplest to estimate and understand. We outline our strategy for estimating market efficiency below.

4.4.1 Methodology:

Consider equation (4.22):

$$\Delta p_t = \psi_{t-1} + (\phi_{t-1} - 1)p_{t-1} + \eta_t \text{ where } \eta_t \sim N\left(0, \sigma_\eta^2\right) \qquad (\mathbf{4.22})$$

where,

$$\psi_{t-1} = \alpha_1, \phi_{t-1} = \beta_1 \text{ } if -\infty < Z_{t-1} < c_1$$

$$\psi_{t-1} = \alpha_2, \phi_{t-1} = 0 \text{ } if \text{ } c_1 \leq Z_{t-1} < c_2$$

$$\psi_{t-1} = \alpha_3, \phi_{t-1} = \beta_3 \text{ } if \text{ } c_2 \leq Z_{t-1} < \infty$$

We define an n-point grid over the extreme values taken by the trigger variable, Z_{t-1}. Thus, our grid takes values $\{m_1, m_2 \dots m_n\}$. Our recursive procedure starts by considering an initial pair of values for c_1 and c_2 (m_1 and m_2 respectively). Using the values of c_1 and c_2, the sample is divided into 3 sub-samples i.e. the first sub-sample contains all values of the asset price that occur when the trigger variable is less than c_1, the second sub-sample contains all values of the asset price when the trigger variable is between c_1 and c_2 and the third contains all values of the asset price when the trigger variable is greater than or equal to c_2. Under most circumstances we would start with $c_1 = c_2$, with no efficient state considered.

The number of grid points is chosen so that we always have more than 2 observations between two consecutive grid points. Following similar terminology to sections 4.2 and 4.3, β_2 is restricted to a value of 0 in order to ensure that the second sub-sample is consistent with an efficient market (if we run the regression in levels/logs instead of returns, β_2 would equal 1). Thus, the market is efficient between the thresholds c_1 and c_2. If $c_1 = c_2$ then the market is not efficient for any amount of time. We estimate parameters β_1 and β_3 for the other two sub-samples using least squares. β_1 is estimated using all values of p_t that correspond to $Z_{t-1} < c_1$ and β_3 is estimated using all values of p_t that correspond to $Z_{t-1} > c_2$.

We calculate the total sum of squared residuals, $\sum_{t=2}^{T} \epsilon_t^2 = \sum_{t=2}^{T} (\Delta p_t - \psi_{t-1} - (\phi_{t-1} - 1)p_{t-1})^2$ for the full sample. In the next iteration of the algorithm we keep c_1 fixed at m_1 and change the value of c_2 to m_3. The above process is repeated and a new set of parameter estimates for β_1 and β_3 are obtained. The procedure is repeated until $c_2 = m_n$, the last point on the grid. Following this first recursion, the recursive procedure is re-started by altering the value taken by c_1 to m_2 and c_2 to m_3. The above double recursion is repeated until $c_1 = m_{n-1}$ and $c_2 = m_n$, which give us our final parameter estimates. The values of c_1 and c_2 and the corresponding β_1 and β_3 that minimize the total sum of squared residuals for the full sample are selected and parameters estimated along with their asymptotic standard errors.

Once we have found the switching parameters and found the thresholds for the trigger variable we can use a time metric to measure market efficiency. Using the thresholds, we can divide the data into efficient periods (corresponding to periods when $\beta_j = 0$) and inefficient periods (corresponding to periods when statistically $\beta_j \neq 0$). If we have a total sample size n with n_1 efficient and n_2 inefficient periods, we can argue that the proportion of time the asset market is efficient is $\frac{n_1}{n}$ and the proportion of time it is inefficient is $\frac{n_2}{n}$.

Note, that if the c_1 and c_2 that minimize the sum of squared residuals are close, it implies that markets are rarely fully efficient (provided that the auto-regressive parameters for the sub-samples are significantly different from 0). Following convention from Section 4.2, α_1, β_1 indicate the first state, α_2, β_2 the second state and so on. If the series contains one mean reverting, one efficient and one explosive state we should find that $\beta_1 < 0$, $\beta_2 = 0$ and $\beta_3 > 0$ or that $\beta_1 > 0$ and $\beta_3 < 0$ given how the recursive procedure and the thresholds operate (it also depends on the relationship between Z_t and p_t).

Note, that unlike chapter 3, we neither need to restrict the number of states nor the values that the parameter β_j will take (except the β_j corresponding to the efficient state). Thus, we may observe multiple mean reverting or explosive states. We restrict our empirical application to the 3 state case

for tractability. Increasing the number of states beyond 3 will increase computational time substantially. Next, we provide an illustrative example using stock-market indices.

We would like to note that the example is illustrative in nature and is primarily aimed at showing how the methodology should be used in empirical work. To be more rigorous any empirical application aiming to derive firm and robust conclusions on market efficiency will need to justify the use of a trigger and also be specific with model selection (i.e. whether to use a specification with or without a switching or constant drift term). Since our example is illustrative, we provide results for all model specifications and comment on how model specification changes the efficiency metric.

4.4.2 Empirical application:

For our empirical study we use monthly returns for the S&P 500 and FTSE 100 indices. Data for the two indices were obtained from Google Finance. The setting in section 4.2 suggests that the states are triggered by an exogenous i.i.d variable. We use the University of Michigan's Index of Consumer Sentiment as our exogenous trigger (MCSI henceforth). The lower bound on the MCSI is 51 and the upper bound is set at 112 so that the regression in (4.22) can be estimated. We use a 200 point grid on the MCSI.

The MCSI is a monthly survey collected by the University of Michigan. It asks questions on personal finance and economic trends of individuals and households through telephonic interviews. Survey respondents are representative of the American population and each month more than 500 interviews are conducted. More information on the Survey and sample design is available on the MCSI website. Other candidates such as the VIX were considered and testing was carried out; it is available upon request. We reported results for the MCSI instead of the VIX because the data for MCSI includes periods from the oil price crisis; the VIX on the other hand starts from the 1990s and misses some important periods. Data for the S&P500 index and the MCSI were obtained from January 1978 to June 2015. The underlying assumption required for the MCSI to be a valid trigger variable in this setting is that contemporaneous and one-period lagged values of the two indices do not impact the MCSI.

While data for the MCSI are available from 1964, the survey was initially collected twice a year and the index only becomes a monthly index in 1978. The FTSE 100 index on the other hand is used from its inception in 1984. All data are monthly. The use of MCSI as a trigger variable for FTSE 100 is justified by the strong correlation between the S&P 500 and FTSE 100 indices (when considering the monthly series in levels). For the data in question the series have an auto-correlation of 0.967 (from January 1984 to June 2015). The MCSI does not appear to be independent (first order autocorrelation > 0.9); thus, it is closer to being a Markovian trigger variable.

Although we have not explicitly calculated moments for the case when the trigger variable is Markovian, we refer the interested reader to Knight and Satchell (2011), which discusses these results for the two states case with a constant drift. Since our example is illustrative in nature, the MCSI will suffice. Selecting a Markovian variable does not impact our estimators or our estimation strategy; however, we will not be able to use the formulae derived in section 4.2 to calculate the mean and variance of our assets.

High consumer sentiment i.e. a positive outlook towards personal finance and general business environment in the country is reflected through a high value of the index. On the other hand low consumer sentiment is reflected as a lower value. We posit that high consumer sentiment that persists for long periods is indicative of explosiveness or bubbles i.e. if consumers have a very positive outlook they are likely to invest in assets and if a large number of consumers enter asset markets or in this case the stock market the increase in demand could lead to a switch from a random walk state to an explosive state.

Similarly when lower values persist we posit that the market is correcting itself and we get mean reverting behaviour. Figure 4.1 shows the MCSI and the log of the S&P500 index indicating how the MCSI varies with the log of S&P500 index. We see a spike in consumer sentiment in the run up to the dot com bubble. A similar increase is seen near the 2008-09 financial crisis. Mean reverting behaviour is observed after the 1979 oil crisis as well as in the aftermath of the financial crisis. While the MCSI may not always respond contemporaneously to movements in the S&P500 index, it nevertheless acts as a valuable trigger variable for our illustrative example.

We estimate the autoregressive and drift parameters in (4.22) using 3 states for the log of S&P500 and FTSE 100 indices respectively. The dependant variable in the regression is asset returns instead of log prices in order to ensure consistency of standard errors. We could have used log levels instead of returns but returns are more intuitive; secondly, using levels we would find some non-stationary states. Using the return formulation also ensures that the criterion for the existence of a stationary distribution, specified in Section 4.2, is always satisfied.

We estimate the model with and without the switching drift term ψ_{t-1}. When we do use a drift term, we report results with both a switching drift term i.e. the drift changes in each state and a constant drift term i.e. the drift does not change across states. The most commonly employed specification in related literature is that with a constant drift. Our aim is to find thresholds c_1 and c_2 for the MCSI that minimize the residual sum of squares for the threshold auto-regression which in turn also yield the parameter estimates for the 2 inefficient states.

Note, that we do not impose any restrictions on the parameters of the other two states; both states may be mean reverting or explosive. The only restriction imposed is $c_1 > c_2$. We use the procedure

outlined above to estimate the thresholds, c_1 and c_2. For our assets these values are reported in Table 4.5 in the appendix and include results for both stock market indices with and without a (switching/constant) drift term. The table also notes the time the stock-market index is in each of the 3-states; this will allow us to comment on the proportion of time each of the stock-market indices is efficient or inefficient.

We postulate that $\beta_1 < 0$ if MCSI is low and $\beta_3 > 0$ when MCSI is high; the postulated relationship will vary based on our choice of trigger variables). Alternatively, the explosive behaviour could result due to a large drift term. Columns (8) and (9) in table 4.5 report the thresholds corresponding to the minimum sum of squared residuals. Since we use different model specifications for this example it is no surprise that the results in table 4.5 present a mixed picture. When we consider the case of a switching drift term in addition to a switching slope coefficient, a lot of the variance in the series is captured by the drift terms. Inclusion of a switching drift substantially reduces the impact of the switching slope terms and we find no more than 32 observations in non-efficient regimes (20 for S&P500 and 32 for FTSE100). We also find little evidence of explosive behaviour due to the slopes.

Thus, explosive and mean reverting episodes under a model with a moving drift are primarily caused by the change in drift. Note that the drift terms are larger in magnitude and appear farther apart which implies that they have a higher variance. As per our formulae in Section 4.2, a higher variance of the drift parameter leads to a higher variance of the series. We find that the drift term corresponding to periods of very high price increases is statistically significant and greater in magnitude than the drift term in other regimes. Econometricians and financial experts modelling asset prices with this specification will thus argue that asset markets are mostly efficient. Any deviations from market efficiency are in fact caused by investors' expectation of a higher return for some period of time and not by a change in slope.

For models with a switching drift, the criterion for a steady state stationary distribution is trivially satisfied as for both FTSE100 and S&P500 we do not find an explosive slope coefficient. Figure 4.2 shows the areas that fall under the different states under this specification for log prices based on the thresholds estimated by the procedure outlined above.

We note that the first state corresponds to periods of relative slow down i.e. in the aftermath of the 2nd oil price crisis, in the immediate aftermath of the financial crisis and in late 2011 when fears of a double dip recession abounded. The other non-efficient state occurs in the run up to the East Asian financial crisis and the dot-com bubble when consumer sentiment was at an all-time high. Our grid-search results do not indicate a deviation from a random walk during the financial crisis. The area under the non-random walk states has been shaded (blue for mean-reverting and red for explosive). Figure 4.3 shows similar results for the FTSE100 index. Note that apart from the dot-com bubble

period in early 2000, 1998 is identified as a period of explosiveness for both indices. Both indices attained historical highs in the 1998 which is reflected in consumer sentiment.

One way to systematically beat the market in such a situation will be through predicting when the switches will occur provided that investors are aware of what state they are in as soon as the switch has occurred (and thereby becomes a part of the information set). Thus, we are referring to efficiency in a broader sense. In the conventional auto-regressive sense, a market is said to be efficient if the auto-regressive parameter is 1 i.e. the process is a random walk or more precisely, a martingale, so that the only change in asset returns is due to unpredictable factors and no gains can be made based on the existing information set. In the threshold auto-regressive case, in addition to the restriction on the auto-regressive parameter we would also require the state switches to be unpredictable; although once the switch occurs everyone becomes aware of it. Thus, the information set will also include information about the exogenous trigger or the state of the asset market. If markets are weak form efficient all rational investors will find out about the switch at the same time although they may not know when the switch may occur.

For the specification without a drift the results are closer to the behaviour observed in the simulations i.e. we observe 3 states although the deviation from efficiency is very small. When $\beta_3 > 0$ i.e. we are in the explosive or bubble regime, we observe additional annualized gains of only 1.5% in the S&P500 index and 1.1% in the FTSE100 index. It may be argued that the additional annualized gains being captured by the parameter are in fact accounting for the missing drift term.

By regressing log prices on their lags instead of returns (i.e. add 1 to each coefficient estimated in table 4.5), we can calculate the value of the criterion function specified in Section 4.2 which allows us to comment on whether the series has a stationary distribution. For the case without drift the value of the criterion for the S&P500 and FTSE100 is 0.0010 and 0.0005 respectively (the criterion in this case is calculated as $\sum_{i=1}^{3} \hat{\pi}_j(\hat{\beta}_j+1)$. Neither of the two indices satisfies the criterion for a steady state stationary distribution under specifications without a drift. This indicates that any test for explosiveness that either assumes a constant drift term without shifting slope coefficients (not reported) or that drop the drift term are more likely to find the criterion violated for the S&P500 and FTSE100 indices.

As mentioned before, if MCSI was an independent and identically distributed variable we would be able to use our formulae from Section 4.2 and be able to calculate the mean and variance for both the S&P500 and the FTSE100 series when they are estimated using a threshold auto-regression. This will have allowed us to compute metrics such as Sharpe ratios enabling us to comment further on market efficiency and investor behaviour. Since MCSI is closer to a Markovian variable we are unable to use the formulae derived earlier. However, our results do allow us to compare efficiency across the two

markets. In the following discussion when we talk about inefficient states we are referring to the number of periods spent by each index in a state that is statistically significantly different from the random-walk.

When specifications with a drift are considered, the FTSE100 index appears more inefficient than the S&P500 index. The S&P500 index is inefficient for 2% of the time with the switching drift specification and 86% of the time with a constant drift. In contrast the FTSE 100 index is inefficient for 6.7% of the time under the switching drift specification and 93% of the time under the constant drift specification. On the other when no drift is included, the S&P500 appears mostly inefficient (95.5%) compared to the FTSE 100 (68.7%). If we compare similar periods i.e. from 1981 onwards, the results remain robust. The mixed results do not offer a clear answer as to which market appears more inefficient; nevertheless the methodology is applicable to other assets. If an asset appears to spend more time in inefficient states under all specifications compared to another we will be able to conclude that the market for that particular asset is inefficient more often. We consider further specifications and other assets in the following sub-section.

Our empirical results supplement our findings in Sections 4.2 and 4.3. Explosiveness is more likely to be detected in asset price series where the criterion function is violated and the variance of the switching slope parameters is large (i.e. there are many regimes or the regimes are much farther apart). Inclusion of a switching drift term may explain most of the explosiveness and may make the price series appear efficient. Another way of analysing results could be through comparing the different series.

The set of results reported above also depends on the selection of the trigger variable. Finding an appropriate trigger variable that may indicate switches in regimes is non-trivial in practice and will require a rigorous theoretical, empirical or experimental basis so that regime identification criteria can be appropriately set. Secondly, while we use contemporaneous values of the MCSI to identify state switches it may be argued that the MCSI is a leading indicator of switches. This again requires judgement on the part of the researcher, the specific asset price being considered and the relationship between the asset price and the trigger variable. Additionally, the researcher also needs to consider the number of states to be used. A price series could exhibit multiple explosive or mean reverting states.

Our methodology can thus work in practice for different asset markets. We have shown with our example that the methodology may be used to identify the incidence of market efficiency for different assets while also highlighting the importance of model-specification when testing for market efficiency. Model specification is not important just in terms of estimation but completely changes the theoretical meaning of the results. Once a researcher has identified an appropriate specification for an asset price or return based on either technical analysis or through solving a structural model, our methodology will allow her to comment on market efficiency for that asset, given an information set.

However, irrespective of the specification, the results may still be used to compare different markets and identify which markets are more efficient for the given information set.

4.4.3 Additional Examples:

In this sub-section we present some additional examples. Table 4.6 reports results for the S&P 500 and FTSE 100 indices under the assumption that consumer sentiment is a leading indicator of investment behaviour. Specifically, a 6 period lag is used for the MCSI i.e. if consumer sentiment is very high; the actual switch in investor behaviour and in the parameter estimates takes 6 months. In addition, we also report results for two commodities, West Texas Intermediate Petroleum and Copper. 2 and 3 period lags yield results very similar to results without lags.

With a switching drift, we note that when a 6 month lag is included the S&P500 results do not change significantly as the index is mostly efficient and earns an average monthly return of 0.7%. Instead of being in the non-efficient state, the S&P500 is in the efficient state most of the time when a constant drift specification is used. When no drift is included, the S&P500 appears to be slightly explosive as before. On the other hand we do note one significant change for the FTSE 100 index. When a switching drift is included for the FTSE 100, instead of being primarily efficient, the series is mostly in an inefficient state. The other two states see a negligible impact.

This shows how changing the lag structure may influence the results we derive, further driving home the point we made earlier. Since the exogenous trigger variable is a US based consumer sentiment index it may be argued that the MCSI influences FTSE 100 with a lag through its impact on the US financial market which takes some months to permeate through the global financial system. If we compare the two markets in terms of efficiency we will reach the same conclusion as before i.e. with a drift (switching or constant) the S&P500 appears to be more efficient whereas without a drift the FTSE100 appears more efficient. Therefore, the efficiency results appear to be robust to the inclusion of lags. Inclusion of lags further highlights the sensitivity of the results to model specification. We tried different lag lengths for the MCSI but they were closer to one of the two results that are reported in this chapter.

We also consider results for WTI oil and copper. In contrast to the indices, the two commodities do display significant deviations from efficiency under all specifications. Regardless of the specification and lag structure employed, both WTI oil and copper are in inefficient states for at least 9% of the duration of the series although the departures from the random walk are not always statistically significant. When we use the switching drift specification, we note evidence of a statistically significant mean switch even though the coefficient on lagged prices does not appear to be significantly different from zero. The switching drift specification is unable to distinguish the Copper

return series from a random walk; for the WTI series on the other hand there is evidence of a switching drift term without a switch in the coefficient.

Specifying a constant drift leads to the most statistically significant results for both Copper and WTI respectively. Irrespective of lag structure we note significant deviation from efficiency based on the time spent by both series in non-random walk states. While the efficiency results for Copper stay robust to lag structure, we do note a difference for oil. WTI oil appears less inefficient when contemporaneous values of MCSI are used to estimate the thresholds. Thus, our results for WTI oil are not robust to lag structure. When we compare the two assets, we note that in all specifications used, the Copper price series is inefficient more often than the WTI series. One caveat to note about the results for Copper and WTI is that the MCSI may not be the most relevant trigger variable for commodities and alternatives such as a measure of global industrial production should be used.

In fact, one may argue that there may be an element of endogeneity present in the trigger variable so in this case it could be argued that consumer sentiment reacts to market conditions rather than the other way round. In such a situation, the correct first step would be to obtain the exogenous variation in the trigger variable by using an instrumental variable regression. The exogenous variation in the trigger variable may then be employed for our metric.

4.5 Conclusion:

In this chapter we have extended existing general conditions that need to be satisfied by threshold autoregressive models in order to have a steady state distribution and for a mean and variance to exist and have provided formulae for them. The results have been extended to include the case where a switching drift is included in addition to a switching coefficient parameter. We have also considered the case of models with and without drift, specifically considering an i.i.d variable as an exogenous regime switching trigger.

We believe that the results can be extended to other types of trigger variables, such as Markovian trigger variables, although we do not evaluate analytical expressions for such cases. We have shown that when a steady state distribution does exist for a TAR(1) model with a switching drift term, the variance depends on the variance of the error term, the variance of the drift parameters, the variance of the coefficient parameters as well as the covariance between the drift and coefficient parameters.

A simulation study is carried out to evaluate the power of the GSADF bubble detection test under conditions where a steady state distribution may not exist. Our simulation study has shown that if a series has a steady state distribution, bubbles may be more difficult to detect. We further note that the power of such tests increases with the variance of the regime parameters i.e. the farther apart the parameters are from unity, the higher is the power. These results enable us to understand why bubble tests may fail to detect explosiveness even though it may be locally present in a series.

Our most substantial contribution in this chapter is a methodology for estimating the proportion of time an asset market may be efficient which use threshold autoregressive models with both exogenous and Markov-chain trigger variables. The methodology is an improvement upon the one introduced in chapter 3 as it is less restrictive and allows the user to estimate any number of finite states. In order to show how the methodology may be used in practice we also provide an illustrative example using stock market indices. It also further enhances upon the over-arching theme of this thesis i.e. that efficiency is one state of nature and deviations from this state are a norm rather than an exception. The methodology outlined in this chapter provides an objective way of identifying efficient and non-efficient states.

Our empirical results indicate that model specification is critical when analysing weak form market efficiency using price series. Series that may appear to exhibit inefficiency when a financial analyst assumes no drift will appear efficient when a regime-switching drift term is used which highlights the need for carefully considering model specification prior to estimation. Thus, before this methodology is employed to carry out empirical or policy driver work, it is important to provide a theoretical backing for the specified model. In the previous chapter we noted how a reduced form for 4.2 may be

obtained. In the following chapter we use a commodity storage model which highlights the conditions under which different states may occur.

Our empirical results vindicate our theoretical findings i.e. the variance of a price process depends not only on the regime-switching coefficients but also on the regime-switching drift term. Additionally we also extend the notion of efficiency to include the predictability of state switching i.e. a market is more efficient if state switching is unpredictable. We believe this methodology is applicable to a variety of different markets including commodity and foreign exchange markets. The methodology also allows us to compare different asset markets and comment on their efficiency relative to one another for a given model specification.

Multiple avenues of further research open up as a result the contributions made in this chapter. The theoretical results may be further expanded to include a Markovian process as the regime switching variable or consider results for TAR (p) models. Indeed, this is something that I have currently been working on in collaboration with my supervisor.

Our simulation results highlight one limitation of the GSADF test and also the need for having bubble tests that may be applied locally or on sub-samples as considering the full price process may make detection difficult. This may justify the use of tests such as the one devised in Chapter 2.

The methodology we have introduced opens up a vast array of possibilities for financial analysts and econometricians alike who may be interested in understanding market efficiency in different markets. In particular, trend-following commodity trading analysts could use such procedures to determine which markets are efficient most of the time and avoid trading dynamically in them unless their mean-variance properties make them intrinsically appealing. This chapter also raises the question of what exogenous trigger variables may be most appropriate for a particular asset market. Finally, identification of a suitable trigger may have policy implications i.e. the government may try to influence expectations about these variables in order to move asset markets towards efficiency although we do caution against using tenuous relationships to draw policy conclusions.

APPENDIX 4 – Tables and Figures

TABLE 4.1: Power test for the GSADF statistic using a multinomial vector ($\alpha = 0$)

Parameter Vector	Multinomial probability vector	Criterion $\sum_{j=1}^{k} \pi_j \ln |B_j| < 0$	Power
[0.98 1 1.02]	[0.10 0.80 0.10]	-0.00004	5.6%
[0.99 1 1.02]	[0.10 0.80 0.10]	0.00098	12.6%
[0.96 1 1.03]	[0.10 0.80 0.10]	-0.0011	3.6%
[0.97 1 1.03]	[0.10 0.80 0.10]	-0.00009	5.8%
[0.98 1 1.03]	[0.10 0.80 0.10]	0.00094	11.0%
[0.97 1 1.04]	[0.10 0.80 0.10]	0.0009	15.8%
[0.95 1 1.05]	[0.10 0.80 0.10]	-0.00025	9.0%
[0.96 1 1.05]	[0.10 0.80 0.10]	0.00080	17.6%
[0.98 1.02 1.05]	[0.80 0.10 0.10]	-0.0093	1.2%
[0.90 1 1.10]	[0.10 0.80 0.10]	-0.0010	23.4%

TABLE 4.2a: Power test for the GSADF statistic using a multinomial vector with drift ($\alpha = 0.01$)

Parameter Vector	Multinomial probability vector	Criterion $\sum_{j=1}^{k} \pi_j \ln |\beta_j| < 0$	Power
[0.98 1 1.02]	[0.10 0.80 0.10]	-0.00004	8.4%
[0.99 1 1.02]	[0.10 0.80 0.10]	0.00098	13%
[0.96 1 1.03]	[0.10 0.80 0.10]	-0.0011	4.2%
[0.97 1 1.03]	[0.10 0.80 0.10]	-0.00009	8%
[0.98 1 1.03]	[0.10 0.80 0.10]	0.00094	16.6%
[0.97 1 1.04]	[0.10 0.80 0.10]	0.0009	20.2%
[0.95 1 1.05]	[0.10 0.80 0.10]	-0.00025	11.2%
[0.96 1 1.05]	[0.10 0.80 0.10]	0.00080	20.4%
[0.98 1.02 1.05]	[0.80 0.10 0.10]	-0.0093	1.3%
[0.90 1 1.10]	[0.10 0.80 0.10]	-0.0010	29.6%

TABLE 4.2b: Power test for the GSADF statistic using a multinomial vector with drift ($\alpha = 0.025$)

Parameter Vector	Multinomial probability vector	Criterion $\sum_{j=1}^{k} \pi_j \ln |\beta_j| < 0$	Power
[0.98 1 1.02]	[0.10 0.80 0.10]	-0.00004	6.6%
[0.99 1 1.02]	[0.10 0.80 0.10]	0.00098	12.2%
[0.96 1 1.03]	[0.10 0.80 0.10]	-0.0011	4.6%
[0.97 1 1.03]	[0.10 0.80 0.10]	-0.00009	8%
[0.98 1 1.03]	[0.10 0.80 0.10]	0.00094	18.8%
[0.97 1 1.04]	[0.10 0.80 0.10]	0.0009	18.6%
[0.95 1 1.05]	[0.10 0.80 0.10]	-0.00025	15%
[0.96 1 1.05]	[0.10 0.80 0.10]	0.00080	20.8%
[0.98 1.02 1.05]	[0.80 0.10 0.10]	-0.0093	1.4%
[0.90 1 1.10]	[0.10 0.80 0.10]	-0.0010	31.6%

TABLE 4.3: Power Test for the GSADF statistic using a Markov-Chain trigger ($\alpha = 0$)

Parameters	Transition matrix			Criteria	Steady State probabilities	Power
[0.98 1 1.02]	.80	.20	0	-0.00004	[0.10 0.80 0.10]	26.8%
	.025	.95	.025			
	0	0.20	.80			
[0.99 1 1.02]	.80	.20	0	0.00098	[0.10 0.80 0.10]	38.2%
	.025	.95	.025			
	0	0.20	.80			
[0.96 1 1.03]	.80	.20	0	-0.0011	[0.10 0.80 0.10]	41%
	.025	.95	.025			
	0	0.20	.80			
[0.97 1 1.03]	.80	.20	0	-0.00009	[0.10 0.80 0.10]	45.4%
	.025	.95	.025			
	0	0.20	.80			
[0.98 1 1.03]	.80	.20	0	0.00094	[0.10 0.80 0.10]	52.6%
	.025	.95	.025			
	0	0.20	.80			
[0.97 1 1.04]	.80	.20	0	0.0009	[0.10 0.80 0.10]	68.4%
	.025	.95	.025			
	0	0.20	.80			
[0.95 1 1.05]	.80	.20	0	-0.00025	[0.10 0.80 0.10]	68.2%
	.025	.95	.025			
	0	0.20	.80			
[0.96 1 1.05]	.80	.20	0	0.0008	[0.10 0.80 0.10]	73.8%
	.025	.95	.025			
	0	0.20	.80			
[0.98 1.02 1.05]	.025	.95	0.025	-0.0093	[0.80 0.10 0.10]	25.8%
	0.20	.80	0			
	0.20	0	.80			
[0.90 1.00 1.10]	.80	.20	0	-0.0010	[0.80 0.10 0.10]	91.4%
	.025	.95	.025			
	0	0.20	.80			

TABLE 4.4a: Power Test for the GSADF statistic using a Markov-Chain trigger ($\alpha = 0.01$)

Parameters	Transition matrix	Criteria	Steady State probabilities	Power
[0.98 1 1.02]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	-0.00004	[0.10 0.80 0.10]	29%
[0.99 1 1.02]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	0.00098	[0.10 0.80 0.10]	38.8%
[0.96 1 1.03]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	-0.0011	[0.10 0.80 0.10]	42.2%
[0.97 1 1.03]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	-0.00009	[0.10 0.80 0.10]	47.8%
[0.98 1 1.03]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	0.00094	[0.10 0.80 0.10]	54.8%
[0.97 1 1.04]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	0.0009	[0.10 0.80 0.10]	70.8%
[0.95 1 1.05]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	-0.00025	[0.10 0.80 0.10]	71.6%
[0.96 1 1.05]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	0.0008	[0.10 0.80 0.10]	75.6%
[0.98 1.02 1.05]	$\begin{pmatrix} .025 & .95 & 0.025 \\ 0.20 & .80 & 0 \\ 0.20 & 0 & .80 \end{pmatrix}$	-0.0093	[0.80 0.10 0.10]	26%
[0.90 1.00 1.10]	$\begin{pmatrix} .80 & .20 & 0 \\ .025 & .95 & .025 \\ 0 & 0.20 & .80 \end{pmatrix}$	-0.0010	[0.80 0.10 0.10]	92.4%

TABLE 4.4b: Power Test for the GSADF statistic using a Markov-Chain trigger

($\alpha = 0.025$)

Parameters	Transition matrix			Criteria	Steady State probabilities	Power
[0.98 1 1.02]	.80	.20	0	-0.00004	[0.10 0.80 0.10]	32%
	.025	.95	.025			
	0	0.20	.80			
[0.99 1 1.02]	.80	.20	0	0.00098	[0.10 0.80 0.10]	41.2%
	.025	.95	.025			
	0	0.20	.80			
[0.96 1 1.03]	.80	.20	0	-0.0011	[0.10 0.80 0.10]	44.2%
	.025	.95	.025			
	0	0.20	.80			
[0.97 1 1.03]	.80	.20	0	-0.00009	[0.10 0.80 0.10]	50.4%
	.025	.95	.025			
	0	0.20	.80			
[0.98 1 1.03]	.80	.20	0	0.00094	[0.10 0.80 0.10]	61.6%
	.025	.95	.025			
	0	0.20	.80			
[0.97 1 1.04]	.80	.20	0	0.0009	[0.10 0.80 0.10]	69.6%
	.025	.95	.025			
	0	0.20	.80			
[0.95 1 1.05]	.80	.20	0	-0.00025	[0.10 0.80 0.10]	74.8%
	.025	.95	.025			
	0	0.20	.80			
[0.96 1 1.05]	.80	.20	0	0.0008	[0.10 0.80 0.10]	75.8%
	.025	.95	.025			
	0	0.20	.80			
[0.98 1.02 1.05]	.025	.95	0.025	-0.0093	[0.80 0.10 0.10]	27.4%
	0.20	.80	0			
	0.20	0	.80			
[0.90 1.00 1.10]	.80	.20	0	-0.0010	[0.80 0.10 0.10]	94.4%
	.025	.95	.025			
	0	0.20	.80			

Table 4.5: Non-linear least squares regression results (with standard errors)

Index	α_1 (s.e)	β_1 (n_1) (s.e)	α_2 (s.e)	β_2 (n_2) (s.e)	α_3 (s.e)	β_3(n_3) (s.e)	c_1	c_2
S&P500 w/mving drift	0.1809 (0.127)	-0.037(11) (0.0203)	**0.0089** (0.002)	0(429) N/A	**1.6368** (0.4359)	**-0.2304**(9) (0.0606)	59.7	107.5
S&P500 w/const. drift	**0.0212** (0.0055)	**-0.012**(11) (0.0022)	0 N/A	0(51) N/A	0 N/A	**-0.002**(387) (0.0009)	59.6	68.7
S&P500 w'out drift	0 N/A	**-0.009**(11) (0.0038)	0 N/A	0(9) N/A	0 N/A	0.0013(429) (0.0003)	59.7	62.2
FTSE100 w/ mving drift	1.157 (0.820)	-0.143(7) (0.0966)	**0.0061** (0.0024)	0(345) N/A	**2.406** (0.575)	**-0.276**(25) (0.0661)	59.6	105.7
FTSE100 w/ const. drift	**0.0302** (0.0088)	**-0.0101**(7) (0.0022)	0 N/A	0(24) N/A	0 N/A	**-0.0031**(346) (0.0011)	59.6	68.7
FTSE100 w'out drift	0 N/A	**-0.0066**(7) (0.0020)	0 N/A	0(111) N/A	0 N/A	**0.0009**(259) (0.0003)	59.6	82.6

n_1, n_2 and n_3 are time periods for which the respective index is in that state. SE represents standard errors

Table 4.6: Non-linear least squares with lags and for additional assets

Series	α_1 (s.e)	B1 (n1) (s.e)	α_2 (s.e)	B2 (n2) (s.e)	α_3 (s.e)	B3 (n3) (s.e)	C1	C2
S&P500 6 lags w/swtch drift	0.0589 (0.0709)	-0.0024(9) (0.01132)	**0.0066** (0.0021)	0(429) N/A	-4.2601 (1.8906)	0.5843(5) (0.2618)	57.80	108.94
S&P500 6 lags w/const. drift	**0.0066** (0.0021)	**0.0058(9)** (0.0023)	0 N/A	0(430) N/A	0 N/A	**-0.0077**(4) (0.0030)	57.80	109.20
S&P500 6 lags w'out drift	0 N/A	**0.0069**(9) (0.0017)	0 N/A	0(9) N/A	0 N/A	**0.0011**(425) (0.0003)	57.80	67.65
FTSE 6 lags w/ swtch drift	0.6845 (0.4872)	-0.0754(6) (0.0578)	-0.0113 (0.0204)	0(6) N/A	**0.1101** (0.0395)	**-0.0128**(353) (0.0047)	58.07	62.24
FTSE 6 lags w/ const. drift	**0.0970** (0.0294)	-0.0064(5) (0.0042)	0 N/A	0(1) N/A	0 N/A	**-0.0112**(365) (0.0036)	57.53	58.60
FTSE100 6 lags w'out drift	0 N/A	**0.0058**(6) (0.0009)	0 N/A	0(29) N/A	0 (N/A)	**0.0008**(330) (0.0003)	57.80	69.52
WTI no lags w/swtch drift	**0.0373** (0.0226)	-0.009(392) (0.0063)	**-0.0588** (0.022)	0(18) N/A	0.0598 0.119	-0.0123(39) (0.0392)	97.69	103.04
WTI no lags w/const. drift	0.0067 (0.0042)	**-0.0132(14)** (0.0053)	0 N/A	0(358) N/A	0 N/A	-0.0036(77) (0.0032)	60.75	96.08
WTI no lags w'out drift	0 N/A	-0.0142(9) (0.0132)	0 N/A	0(401) N/A	0 N/A	**0.0073**(39) (0.0039)	58.07	103.04
WTI 6 lags w/swtch drift	**0.1987** (0.0678)	**-0.0522**(83) (0.0176)	**0.0308** (0.0067)	0(38) N/A	0.03557 (0.0266)	-0.0103(322) (0.0078)	72.18	76.18
WTI 6 lags w/cnst. drift	**0.0366** (0.0110)	**-0.0109**(80) (0.0036)	0 N/A	0(41) N/A	0 N/A	**-0.0106**(322) (0.0035)	71.65	76.18
WTI 6 lags w'out drift	0 N/A	-0.0071(24) (0.0063)	0 N/A	0(340) N/A	0 N/A	0.0041(79) (0.0028)	63.13	95.88
Copp no lags w/swtch drift	0.2495 (0.462)	-0.0408(15) (0.0550)	0.0043 (0.0150)	0(14) N/A	0.0057 (0.1523)	0.0001(420) (0.0047)	60.99	64.45
Copp no lags w/const. drift	**0.0185** (0.0070)	**-0.0133(15)** (0.0020)	0 N/A	0(74) N/A	0 N/A	**-0.00181**(360) (0.0010)	60.99	72.98
Copp no lags w'out drift	0 N/A	**-0.0111**(14) 0.0038	0 N/A	0(15) N/A	0 N/A	**0.0009**(420) (0.0003)	60.99	64.46
Copp 6 lags w/swtch drift	-0.5697 (0.5341)	0.0713(10) (0.0652)	-0.134 (0.138)	0(2) N/A	0.0485 (0.0387)	-0.0056(431) (0.0049)	58.86	59.93
Copp 6 lags w/const. drift	**0.0461** (0.0162)	**-0.0065**(96) (0.0021)	0 N/A	0(13) N/A	0 N/A	**-0.0053**(334) (0.0021)	73.51	74.58
Copp 6 lags w'out drift	0 N/A	-0.0014(69) (0.0013)	0 N/A	0(24) N/A	0 N/A	**0.0008**(350) (0.0003)	70.05	73.25

Figure 4.1: Log S&P500 Index and the Michigan Consumer Sentiment Index
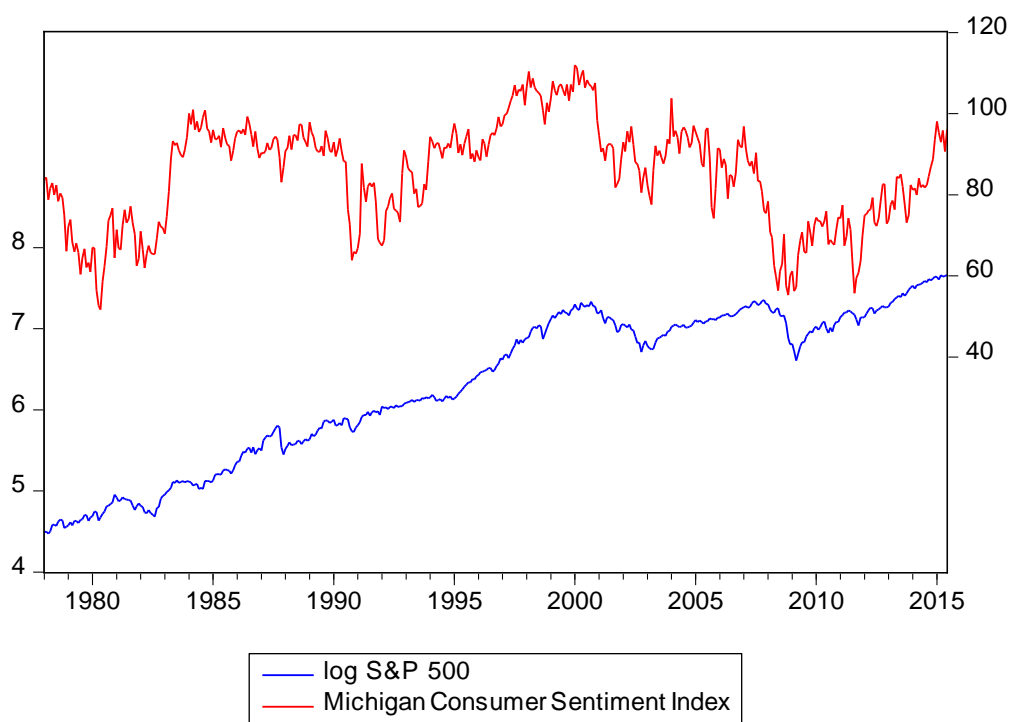

Figure 4.2 S&P500 index: Results for non-linear least squares with a switching drift
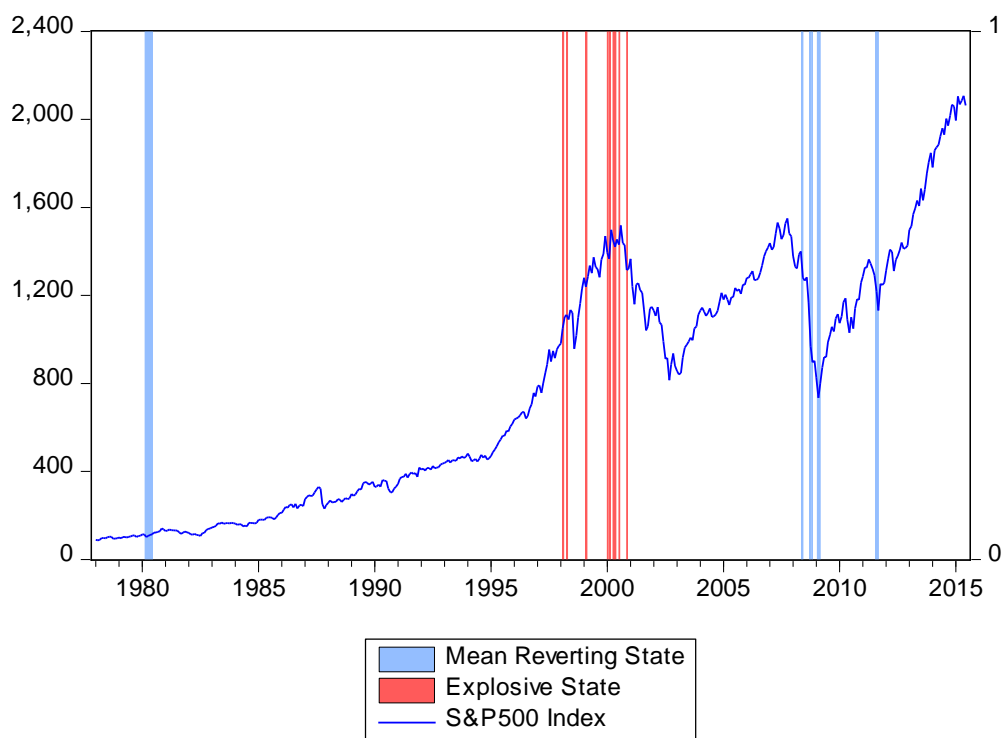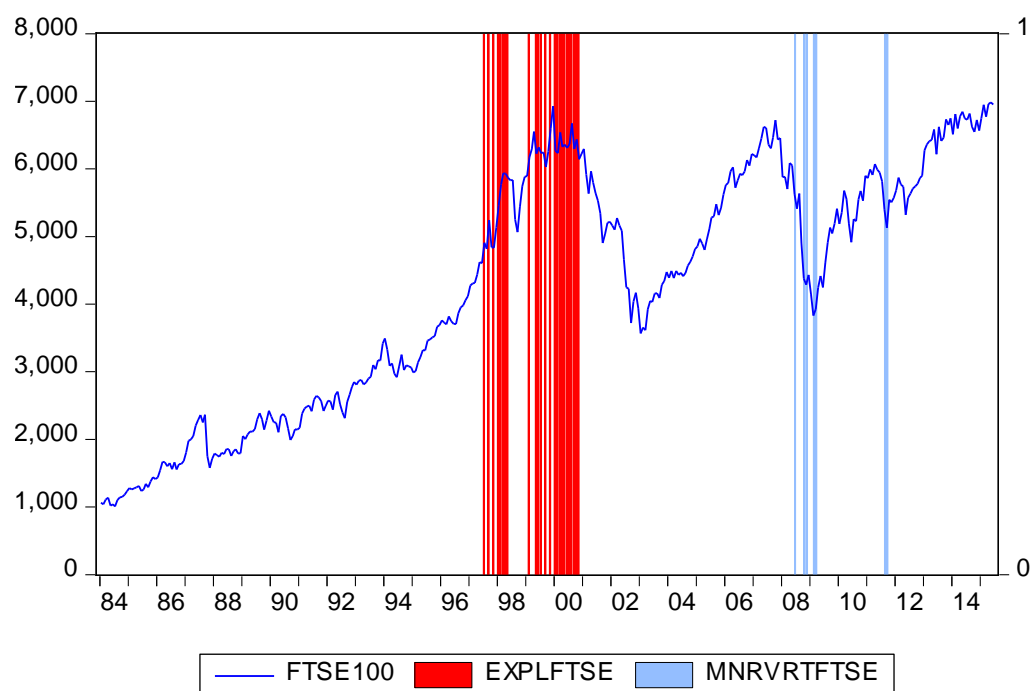
Figure 4.3 FTSE 100 Index: Results for non-linear least squares with a switching drift

CHAPTER 5: *Making the Commodity Storage Model Empirically relevant*

William and Wright's (1991) seminal commodity storage model aimed at capturing properties of real commodity data but despite going through various iterations it could not capture first order auto-correlations accurately. In this chapter we make three important changes to the model; We use an iso-elastic demand curve, add a convenience yield and introduce implicit saving behaviour. The resulting model retains the simplicity of the original model but is able to capture not only the first-order autocorrelation but also skewness and kurtosis similar to that found in real commodity data. The saving behaviour introduced also offers an explanation for occasionally explosive behaviour that is observed in commodity prices; this provides an insight into what may cause the state-switching behaviour discussed in the previous chapters. An empirical application of the augmented model is provided to show that, when correctly calibrated, the model comes close to replicating distributional properties of selected commodities

5.1 Introduction

This chapter modifies William and Wright's (1991) partial equilibrium commodity storage model so that it is able to better replicate the distributional properties typically exhibited by commodity prices. Storable commodities such as grains like wheat and sugar are characterized by two main features: a strong first order auto-correlation and occasional spikes or increases in volatility. Any model attempting to describe commodity prices needs to capture these two features of commodity prices. While there have been numerous iterations of William and Wright's (WW henceforth) seminal work, they have not been successful at capturing both features simultaneously; at least not without complicating the model substantially. The model does not explicitly consider behavioural aspects which Chapter 3 alludes to and instead focusses on one primary shock mechanism, supply.

Such modifications usually fail to capture the strong auto-correlation that is a feature of most commodities. Storage fundamentally sets apart commodities from other assets and forms the basis of a strong first-order auto-correlation; thus, WW's commodity storage model provides a seminal study for

understanding commodity markets. At the outset, it is worth mentioning that the commodity storage model is primarily meant for agricultural commodities which have a 6 month to 1 year growing cycle as opposed to industrial metals or precious commodities which are mined on a more continuous basis.

Through using an iso-elastic demand curve which also reacts to the amount of storage at hand, a multiplicative convenience yield similar to Ng and Murcia (2000) which reduces the frequency of stockouts and using a parameterized expectations algorithm to solve for the partial equilibrium model ala Gouel (2013) we show that simulations based on the solution of the partial equilibrium model can replicate the features commonly found in commodity prices. By calibrating the model as per convention, we observe auto-correlations of around 0.80 while still achieving occasional spikes that are typically found in commodity prices. Unlike other studies which focus solely on the auto-correlation aspect we also test for explosiveness in our simulated price; this helps us compare our simulated results to real data at an additional level. Since the model attempts to replicate real and not nominal prices we do not always observe statistically significant evidence of explosiveness although that is a feature shared with real data. Another way of analysing locally non-stationary behaviour in commodity prices is through the fourth moment or Kurtosis. A positive excess Kurtosis is a feature of pricing processes that exhibit extreme peaks and troughs frequently.

Before we introduce the model and the solution methodology, it is important to understand how the original model has evolved and what features various modifications to the model have introduced. While WW were the original creators of the partial equilibrium model, the model shot to prominence through a series of articles written by Deaton and Laroque (1992, 1995, 1996). Deaton and Laroque (DL henceforth) showed that the standard storage model of WW has a Stationary Rational Expectations Equilibrium (SREE) and in their article they solved an analytical version of the model. However, their model is unable to capture essential features of commodity prices and predicts a much higher frequency of stock outs than generally observed. The use of an autoregressive harvest processes is among the modifications they have applied to the standard storage model although the modified model was still unable to capture the high auto-correlation. Their major contribution in addition to proving that the model will have an SREE solution is a strategy for empirically estimating the model for different commodities using a maximum likelihood approach.

Miranda (1993) estimates a nonlinear REE commodity storage model which includes government stockholding in addition to private stockholding. He finds that the existence of government stocks crowds out private stockholding. Miranda and Rui (1999) add storage costs to the model in addition to depreciating stock and use Chebychev orthogonal collocation to find a numerical price function which provides a better fit to the auto-correlation function. They consider both a constant storage cost and a semi-log cost of storage function similar to a convenience yield (Kaldor, 1939). Inclusion of a semi-log cost of storage function eliminates stock outs and since the main mechanism through which auto-

correlation is introduced in the model is storage, the lack of stock outs thus increases auto-correlation. However, this higher auto-correlation comes at the cost of less volatility in the price function since price in this model typically increases in response to a stock out. Thus, one feature of commodity prices is substituted for another.

Ng and Murcia (2000) add a number of extensions to the model to explain the auto-correlation without compromising on volatility. They consider the cases of staggered production, long contracts and convenience yields. The convenience yield that we use in our augmented model is based on the specification in their article. They find that the inclusion of convenience yields can explain the auto-correlation better than other alternatives although this comes at the cost of reduced volatility in the price process. However, their auto-correlation measures are around 0.6 on average as opposed to the 0.8 observed in real commodity data. They also consider an MA(1) harvest process which increases auto-correlation. While an MA(1) harvest process can increase auto-correlation, we abstain from using this in our model as this is more likely to be a feature of a model aiming to replicate quarterly or monthly as opposed to annual data. In addition, they use a linear demand curve.

Other articles that have contributed to this literature more recently include Wright (2010), Miao et al (2011), Guitterez et al (2014), Gouel (2013) and Guerra et al (2015) although the latter papers focus primarily on empirical estimation of the model through extending DL's methodology. Gouel (2013) provides a useful summary of different numerical approaches that may be employed to numerically estimate the commodity storage model including Value Function Iterations, Projection Methods and Parameterized expectations algorithm (PEA). He shows that the PEA algorithm converges faster and that the solution has better properties than other methodologies. This algorithm is also fairly close to the original solution algorithm adopted by WW. Thus, we proceed with this numerical methodology in this chapter. Miao et al (2011) have attempted to augment the model in similar vein to the current chapter; however, in doing so they increase computational burden considerably and our results come closer to real data than theirs. They also employ the use of a fluctuating real interest rate which improves their results. While adding some complexity to the commodity storage model we ensure that the computational burden stays manageable. The main similarity between our augmented model and Mioa et al's is the use of an iso-elastic variation of the demand curve.

A recent article worth mentioning separately is Arsenau and Leduc's (2013) study which extends the partial equilibrium model to a general equilibrium setting. They extend the canonical model by introducing a production sector and incorporating the consumption savings decisions. Thus, consumption, the real interest rate, storage and production are all simultaneously determined in the model. While their methodology for solving the model is similar to WW and Gouel (2013), they do not provide a simulation analysis since their main motivation is policy analysis. Since we are able to explain commodity prices within the confines of the canonical model, we do not extend our analysis

to the general equilibrium setting as it substantially increases the dimensionality of the problem. We do, however, try to incorporate saving behaviour in our partial equilibrium framework through augmenting the demand curve.

None of the articles discussed above do any formal analysis on the volatility of simulated prices. We use the test developed by Philips, Shi and Yu (2013) to check if there is substantial explosiveness observed in simulated prices. We note that simulated prices tend to replicate real prices as opposed to nominal prices and thereby significant mean reversion is noted; as a result, the GSADF test is often unable to detect explosiveness as we pointed out in Chapter 4. The next section describes the original model and introduces our modifications to it. Section 5.3 elaborates on the numerical solution methodology while Section 5.4 compares the properties of our model to other versions of the model. Section 5.5 tests the sensitivity of the model solution to various parameters of interest while Section 5.6 attempts to empirically replicate the properties for five commodities. Section 5.7 concludes.

5.2 The Commodity Storage Model

While we use modifications introduced in Ng and Murcia (2000) and Gouel (2013), it is instructive to look at the original model by Williams and Wright. The methodology we use to solve the model is fairly close to the algorithm adopted in the original model as well.

Supply and Demand:

The standard WW model is characterized by a linear demand curve and no supply response. Since the model is set in a partial equilibrium setting, the supply and demand functions are given.

$$P_t = a - bQ_t \qquad\qquad (5.1)$$

$$h_{t+1} = \bar{h}(1 + v_{t+1}) \qquad\qquad (5.2)$$

P_t is the price of the commodity, Q_t is consumption demand, $\bar{h}$ is average yield/production and v_t is a normally distributed weather shock. Each period the realization of the weather shock v_t determines the yield in the following period, h_{t+1}. The commodity is planted one period ahead and the realization of the weather shock determines the actual output obtained. If planned production responds to price, the above planned production function becomes a function of the incentive price P_t^r i.e. the expected price anticipated given the distribution of the shock process:

$$h_{t+1} = f(P_t^r) \qquad\qquad (5.2)$$

In the standard model the weather shock is normally distributed with a mean of 0 and a standard deviation of 0.10. Specification of the weather shock can influence the solution of the model and the resulting decision rules. DL have investigated an AR(1) weather process while Ng and Murcia have

employed an MA(1) weather process. While building persistence in shocks this way inflates correlation in the resulting simulations, the results still fail to match real data. We find that reducing the standard deviation of shocks can also lead to higher correlations; we comment on these findings in Section 5.5.

Investors:

The model solves the investors' problem. Investors or warehouse owners are risk neutral agents looking to maximize profits by optimally choosing an inventory level. This leads to an arbitrage condition that determines inventory demand. Investors purchase and carry forward an amount S_t today at price P_t. Inventory is stored at constant cost k_t per unit and undergoes depreciation δ_t. Inventory is sold in the following period at P_{t+1}; however, since the storage decision has to be undertaken one period prior to the actual sale of inventory, investors choose an inventory based on their expected price $E_t(P_{t+1})$. The investors' problem is thus:

$$E\big(\Pi_{t+1}(S_t)\big) = \left[\frac{1-\delta}{1+r}E_t(P_{t+1}) - P_t - k_t\right]S_t \tag{5.3}$$

Investors select storage to maximize their profits. Since storage cannot be negative, maximizing 5.3 with respect to storage implies the following arbitrage condition (Scheinkman and Schechtman, 1983):

$$S_t \geq 0 \; if \; \frac{1-\delta}{1+r}E_t(P_{t+1}) - P_t - k_t \geq 0 \tag{5.4}$$

Otherwise, $S_t = 0$

Thus, investors only enter the commodity market if they expect the price to rise by an amount that is sufficient to cover their cost of capital, the storage cost, the purchase cost and depreciation. In the standard storage model, WW do not employ the use of a convenience yield ala Kaldor (1939). The investors' problem is thus, non-linear. Given supply (5.2), demand (5.1) and equilibrium conditions (5.5 and 5.6), we can solve for the optimal amount of storage and hence, derive availability using numerical methods.

To close the model and find equilibrium storage, we define availability or the amount available each period to satisfy consumer demand and storage demand as

$$A_t = S_t + Q_t \tag{5.5}$$

This demand is satisfied through storage carried over from the previous period and the realized production in the current period (which depends on the weather shock in the current period, v_t).

$$A_t = (1-\delta)S_{t-1} + h_t \tag{5.6}$$

Although availability is the state variable of the system, most solution algorithms solve for storage rules and through the storage rule, derive a decision rule for availability. This is due to the non-linearity of storage. Once a decision rule has been found for the part for which a storage rule exists we are able to derive the availability decision rule for the availability space.

5.2.1 Augmented Storage model

Our modifications to the standard storage model are three fold. We use an iso-elastic demand function instead of a linear demand function ala Gouel (2013), we employ the use of a multiplicative convenience yield similar to the one used by Ng et al (2000) and introduce a response to storage in the demand function. We also use an iso-elastic supply function instead of opting for no supply response. While this does increase computational time, a supply response is more intuitively appealing and theoretically justifiable in a more general setting. In a number of applications of the WW model, we noted that although authors acknowledged a supply response in their narrative, they often solved the model assuming no supply response.

Planned production is given by:

$$h_{t+1} = E_t[(P_{t+1}v_{t+1})^{\epsilon_s}] \qquad \text{where } \epsilon_s > 0 \qquad\qquad \textbf{(5.2a)}$$

$P_{t+1}v_{t+1}$ represents the incentive price to which commodity producers respond. ϵ_s is the supply elasticity of production. We propose a more general form for the demand function. In a general equilibrium setting, as in Arseneau and Leduc, consumers try to maximize utility by selecting not only current but also future expected consumption. This implies that current period consumption or demand is a function of future expected consumption in addition to being a function of current price. Since we are primarily interested in studying the impact of storage on prices, we represent this, more general demand function as $Q_t(P_t, E[Q_{t+1}])$. However, expected future consumption is itself a function of storage; thus, the general form of the function becomes $Q_t(P_t, E[Q_{t+1}(S_t)]) = Q_t(P_t, S_t)$ where $Q_{P_t} < 0$ and $Q_{S_t} > 0$.

When storage is low or 0, consumers anticipating lower consumption in the future (due to lower availability) start consuming a lower amount today; lower consumption in the current period may lead to a lower equilibrium price which may make more storage affordable for some investors. As a result, the number of stockouts is lower. The opposite is true when storage is positive. This can also be thought of as consumers transferring some of their consumption to the future in response to an anticipated price increase in future periods (i.e. consumption smoothing). Although a variety of

functional forms may be used to model this behaviour we rely on the simplest functional form. Our Inverse demand function is given by

$$P_t = \left(\frac{1}{S_t^2 + C}\right) Q_t^{1/\epsilon_d} \qquad \text{where } \epsilon_d < 0 \qquad\qquad \textbf{(5.7)}$$

C is a constant and is chosen to determine how much consumers react to storage and how much consumption smoothing they prefer; ϵ_d is the constant price elasticity of demand. Price responds by a constant percentage to change in consumption but its response to storage is more volatile. We believe that this behaviour will help us explain the skewness and excess kurtosis in observed data. This also aligns well with the asset price behaviour we have worked with in previous chapters. Note, that due to the nonlinearity of storage and (5.5), both (5.7) and (5.2a) also become non-linear functions. While saving behaviour may be more accurately captured in a general equilibrium setting, since our main objective is to understand commodity price movements we are able to reduce the dimensionality of our problem and save significant computational time through incorporating saving behaviour within the consumption demand function.

The notion of a convenience yield was first introduced by Kaldor in his seminal 1939 article where he discussed a commodity model with storage. Put simply convenience yield is the convenience of having goods to hand. Having goods to hand enables storage owners to meet unexpected increases in demand. Thus, precautionary or speculative demand may give rise to a convenience yield. As a result, including a convenience yield increases the demand for storage. Kaldor(ibid) provides additional justifications for a convenience yield.

While there are different methods of modelling a convenience yield, we use the multiplicative convenience yield employed by Ng et al (2000). Prior to Ng et al, Miranda (1997) had used a logarithmic version of the convenience yield. Using a logarithmic convenience yield removes the non-linearity in the model giving rise to continuous storage and price functions. However, this also substantially reduces the volatility of simulated price as there are no stock outs.

The multiplicative convenience yield is expressed as a function of storage, $\phi(S_t)$. Holding on to more inventory will increase the convenience yield; however, this is likely to be at a decreasing rate. Thus, $\phi(S_t)$ can be expressed as a concave function. We use the parameterization used by Ng et al:

$$\phi'(S_t) = \theta + (1 - \theta)g(S_t)$$

where

$$\theta = \frac{(1 + r)(1 - \varepsilon)}{1 - \delta}$$

ε is arbitrarily small; for a normalizing constant 'L', $g(S_t)$ is expressed as

$$g(S_t) = \frac{S_t}{(S_t + L)}$$

Setting L= 0 eliminates the convenience yield. Ng et al discuss more properties and implications of this convenience yield. Introducing the multiplicative convenience yield along with the demand curve (5.2a) in the arbitrage condition (5.4) implies:

$$S_t \geq 0 \; if \; \frac{\phi'(S_t)(1-\delta)}{1+r} E_t(P_{t+1}) - P_t(A_t - S_t) - k \geq 0 \qquad\qquad \textbf{(5.8)}$$

In the resulting solution, stock outs are less common and more storage is held per period due to the convenience of having storage to hand. Storage builds more persistence in the model and the lower frequency of stock outs implies fewer episodes of explosive prices.

In summary, our system solves for optimal storage in equation 5.8 given the following relationships:

Augmented demand:

$$P_t = \left(\frac{1}{S_t^2 + C}\right) Q_t^{1/\epsilon_d}$$

future planned production, (5.2a)

$$h_{t+1} = E_t[(P_{t+1} v_{t+1})^{\epsilon_s}]$$

and the identities implied by (5.5) and (5.6)

$$A_t = S_t + Q_t$$

$$A_t = (1 - \delta)S_{t-1} + h_t$$

DL have proved that the above system has a Stationary rational expectations equilibrium. (5.2a) and (5.8) define the equilibrium of the commodity storage model. Our solution algorithm is outlined in the following section.

5.3 Solution Algorithm

In order to aid comparison with other research in the area we adopt the solution algorithm used by WW (1991). Gouel (2013), Ng et al (2000) and Arseneau et al (2013) adopt the same algorithm with minor amendments. The algorithm is often employed in the Macroeconomics literature and is referred to as the Parameterized expectations algorithm. In a nutshell, the algorithm aims to parameterize $E_t(P_{t+1})$ as a function of storage over a pre-determined grid. Although the actual state variable in the

model is availability, A_t, due to the non-linearity of the model we use storage, S_t, as the state variable and derive the decision rule for availability.

The state space for the harvest/weather shock is approximated as a discrete process over an M point grid. Although there are different ways of discretizing a shock process, we use the same values and probabilities that WW employ (i.e. M=9). Each weather state $1 + v_{t,j}$ occurs with a probability $\pi_{t,j}$ .Note that $v_{t,j}$ has a standard deviation of 0.10. A Gauss-Hermite quadrature would have discretized the weather process in the same way. Discretizing the weather shock enables us to calculate the expectation functions[5]. We also define an N point grid over storage $S_{t,n} = [S_{t,1}, S_{t,2} \ldots \ldots S_{t,N}]$. $\psi_p(S_t)$ denotes the polynomial for expected price in terms of storage. We initialize the algorithm by selecting values for the polynomial coefficients i.e. $E_t[P_{t+1}] = \psi_p(S_t)$ over all values of the storage grid. We formulate a guess for initial planned production $\widetilde{h}_t$.

Actual production is found for each realization of the shock i.e. $h_{t,m} = \widetilde{h}_t(1 + v_{t,m})$. For every realization of the shock and production, h_t, we find availability, $A_t = (1 - \delta)S_{t,i} + h_{t,m}$. Using this level of availability and our current guess for the polynomial $\psi_p(S_t)$ we use MATLAB's non-linear solver on equation (5.8) to find the equilibrium level of storage for each realization of the shock. If the equilibrium level of storage for any realization of the shock is negative we re-set the equilibrium storage level to zero. Equilibrium storage and availability allow us to find consumption and therefore price for each realization of the shock using (5.7). Expected price can thus be found as $E_t(P_{t+1}) = \sum_{j=1}^{m} \pi_{t,j} P_{t+1,j}$. $E_t(P_{t+1})$ in turn allows us to find the producer's incentive price and thereby actual planned production conditional on the distribution of the weather shock, $\widehat{h}_t$. For every level of storage on the grid we find $\widehat{h}_t - \widetilde{h}_t$. If the calculated difference is below our tolerance threshold (we set the tolerance threshold to 10^{-8}) we continue the algorithm; otherwise, we change $\widetilde{h}_t$ to $\widehat{h}_t$ and repeat.

Once planned production has converged, we regress $E_t(P_{t+1})$ on all storage levels across the grid, S_t and find new coefficient estimates for the polynomial which we label $\widehat{\psi_p}(S_t)$. If $|\widehat{\psi_p} - \psi_p|$ is below our threshold tolerance we will have found the self-replicating polynomial for $E_t(P_{t+1})$ as a function of storage and hence solved for the rational expectations equilibrium. If the values do not converge, we set $\psi_p = \widehat{\psi_p}$ and repeat the procedure from the initial step. Note that in addition to the polynomial for expected price, we can also find a polynomial expressing planned production as a function of storage from the above steps. Armed with these two polynomials we can find the equilibrium functions for P_t, S_t, Q_t, h_t for given availability, A_t.

---

[5] $v_{t,j} = [-0.20 - 0.15 - 0.10 - 0.05 \ 0 \ 0.05 \ 0.10 \ 0.15 \ 0.20]$;
$\pi_{t,j} = [0.0401 \ 0.0659 \ 0.1212 \ 0.1745 \ 0.1966 \ 0.1745 \ 0.1212 \ 0.0659 \ 0.0401]$;

For our analysis we used a third-order polynomial such that $\frac{\partial \psi(S_t)}{\partial S_t} < 0$ and $\frac{\partial^2 \psi(S_t)}{\partial S_t^2} \geq 0$. As mentioned before, we discretize the error process over 9 grid points and choose a 50 point grid for storage.

Equilibrium functions allow us to find values of the control variables in the commodity storage model for given values of the state variable A_t in the current period. The support for these equilibrium functions is defined over the availability state space [0.8 – 1.2]. We use spline interpolation to plot the equilibrium functions as some of the functions are non-linear. Figure 5.1 shows the equilibrium functions for the initial calibration.

The resulting splines allow us to run simulations and thus, observe the dynamic behaviour of commodity prices over time. The features of these simulated prices are then compared with actual prices. In particular we are interested in finding the first order autocorrelation and the occurrence of explosiveness or bubbles. Table 5.1 in the appendix states our calibrated values for initial specifications. Initial calibrations are derived from Ng et al (2000) for the convenience yield and Gouel (2013) for the elasticities. Our calibrations are chosen so that steady state availability and planned production are around 1. In our analysis we also provide a comparison with the original model and various other modifications to it.

Note the prominent features of the equilibrium functions. Storage tends to stay zero until threshold availability is reached; once the threshold level has been reached it grows linearly with availability. In similar vein, planned production stays constant at 1.03 until availability increases beyond the threshold. Once producers observe storage behaviour they are forced to reduce planned production as their incentive price decreases. These features are shared across other specifications of the commodity storage model.

Price and consumption on the other hand show slightly different behaviour due to some of the modifications we have introduced. Price is more volatile and reacts strongly when storage is low or zero. We believe that the strong reaction better captures real pricing behaviour. Consumption on the other hand tends to flatten out once availability increases beyond the threshold level. There is some evidence of consumption smoothing behaviour beyond the threshold level of availability. Consumers start saving a greater proportion of their income as prices are lower and there is evidence of an income effect. While the equilibrium functions are for one period only, simulations allow us to use these equilibrium functions to analyse dynamic properties of our model which is the subject matter of the next section. We also test the sensitivity of our model to various parameters in a separate section.

5.4 Simulation Results:

The equilibrium functions derived in Section 5.3 above allow us to simulate the commodity storage model. Thus, we are able to simulate a time series for commodity prices which incorporates the

impact of storage and weather shocks on the price of commodities. As stated previously, the objective of such an exercise is to replicate real commodity behaviour through matching the properties of its distribution. One time period for all the analysis carried out in this chapter is to be understood as one full commodity cycle, normally thought to be one year. Secondly, the model aims at replicating real as opposed to nominal prices. Before presenting simulation results we briefly discuss the properties of real commodity prices that we seek to replicate. The commodities we consider are coffee, cotton, maize, sugar and wheat.

Our price data for annual commodity prices is the Grilli and Yang (1988) data series as augmented by Pfaffenzeller et al (2007). We use the latest (2011) update for our results. The data reports results on 24 commodities along with an index of all 24 commodities from 1900-2011. Table 5.2 reports the properties of the 5 commodity series mentioned above. We report moments for both nominal and real values of each commodity. The real price series were obtained by dividing the nominal price series by the US CPI obtained from the Bureau of Labour statistics (base year: 1977-79 average). In addition to the moments and correlation we also report results for the GSADF test of explosiveness. Figure 5.2 shows the nominal and real sugar price series.

The graph highlights important features of commodity markets. The real price of sugar has been falling over the years primarily due to developments in technology (although this is not as high as the fall observed in other commodities such as cotton or wheat). However, both price series remain volatile and we continue to observe occasional spikes and troughs in both the real and nominal price series. Note that using annual instead of monthly data will make it more difficult to observe long periods of explosiveness. As the analysis in chapter 2 indicates, periods of explosiveness tend to last from anywhere between 2 months to 1 year. Thus, conducting the GSADF test on the real series instead of the nominal series and using annual instead of monthly data should reduce the probability of detecting explosiveness as noted in chapter 4. A 'Yes' in the GSADF test result row indicates that we were able to detect explosiveness in the series.

Some features of the distributions of our commodities are common. All commodities have a high first order auto-correlation (around 0.8 on average), are positively skewed and have excess kurtosis although the extent of excess kurtosis tends to vary. These features are shared across nominal as well as real prices although the degree of positive skewness and excess kurtosis is lower in real as opposed to nominal prices on average. As outlined above, except Cotton we are able to detect explosive behaviour in all Nominal price series. On the other hand we are unable to detect statistically significant evidence of explosiveness in the real series even though their distributional characteristics appear similar to their nominal counterparts. Thus, for our simulation results to hold validity we would expect a high auto-correlation, positive skewness and excess kurtosis. Since our simulations are closer to real instead of nominal prices we may not be able to detect explosiveness very often.

For our simulation analysis we solve and simulate four different versions of the commodity storage model which highlight the importance of each feature introduced in the augmented version of the model. The first set of simulations is carried out for the fully augmented model described in section 5.2.1 above. Next, we simulate the model without any savings behaviour and use a simple iso-elastic demand curve. The third set of simulations uses a linear demand function (equation (5.1)) instead of an iso-elastic demand function. Our final variation is the augmented model without a convenience yield. To get statistical properties we solve each model and simulate the model 2000 times; each simulated series is 700 periods in length. The results reported in Table 5.3 are averages.

5.4.1 Augmented Model

In addition to reporting auto-correlation, skewness, kurtosis and the percentage of series exhibiting bubbles, we also report the average frequency of stockouts in the simulated price series. Table 5.3 reports the results for our 4 model variations.

As noted in Table 5.3, our augmented model achieves a correlation of 0.7960 which is fairly close to the average correlation attained by the 5 commodities as noted in Table 5.2. Note that the simulations continue to be calibrated as per the values in Table 5.1. Our sensitivity analysis in the following sections suggests that the auto-correlation is sensitive to some parameters. We also observe positive skewness and excess kurtosis. While the skewness measure appears to be within the expected range, we do obtain a higher than expected value for kurtosis. None of the commodities in Table 5.2 exhibit Kurtosis higher than 8 whereas we obtain a kurtosis nearly double that number. This is partially due to the way we have modelled consumption behaviour but also reflects the impact of modelling the complex dynamics that result from storage.

One of the major criticisms of the original WW model was that it predicted a much higher number of stockouts than generally observed in commodity markets. We observe stockouts on average 1.6% of the time or 11 periods out of 700. Statistically significant evidence of explosiveness was found in 14.12% of the series which indicates that explosiveness may not be readily detectable in simulations of the commodity storage model although the simulated series continues to exhibit peaks and troughs that are very characteristic of actual series. Figure 5.3 shows a simulated price series using the augmented model.

The simulated price series in Figure 5.3 shares properties with the Real price of Sugar; we observe a long term decline in price with occasional spikes or volatility. Due to the overall mean reverting nature of this process we were unable to detect a bubble in this particular simulation despite the occasional explosiveness in the series (particularly around period 60).

5.4.2 Comparison with other models

Columns 3, 4 and 5 in table 5.3 present results from different variations of the augmented model. In the first variation we remove any feedback from storage to consumption; thus, the demand curve simplifies to:

$$P_t = Q_t^{1/\epsilon_d}$$

As we have mentioned before, storage influences two aspects of commodity prices: correlation and volatility. These two aspects often conflict with each other. Storage ensures that quantities, and thereby consumption, across two consecutive periods are closer together which tends to increase auto-correlation; when we observe a stockout, volatility increases and we typically see a spike in the price of the commodity which tends to reduce correlation. With the simple demand curve above, we note that the correlation reduces considerably although the frequency of bubbles increases. Measures of skewness and kurtosis on the other hand are out of sync with the skewness and kurtosis observed in real prices. The excessive kurtosis (57.3) is primarily responsible for the large number of series exhibiting statistically significant explosive behaviour (62.8%). Thus, we obtain a reasonably high correlation and explosive frequency but other properties of real data are not captured.

For the linear demand model we used equation (5.1) which was calibrated based on the parameter values chosen by Ng et al that gave them one of the higher correlation values. Since Ng et al's study is empirical in nature we chose the parameters for wheat and re-estimated the model. Due to the convenience yield in our model we were able to obtain an average correlation that is higher than what Ng et al obtained for the same specification however it is still much lower than the correlation observed in real data. The frequency of stock-outs is also higher but a similar amount of explosiveness is detected. Thus, the model with linear demand does much better than the standard model but it is unable to match the results of the fully augmented model.

In the final variation we exclude the convenience yield while keeping saving behaviour and the iso-elastic demand curve; thus we use equation (5.4) instead of equation (5.8) as the arbitrage relationship. This results in significant differences in the result. Having a convenience yield increases the demand for storage by virtue of assigning a value to having stock at hand; this reduces the number of stockouts observed. We note that excluding the convenience yield greatly increases the number of stockouts in the model. No stocks are held for more than a third of the period in our simulations. Unsurprisingly this results in a much lower correlation. Skewness and kurtosis are much smaller but are of the right sign. None of our 2000 simulations for this particular specification exhibited explosive behaviour which is an extreme result. In fact a number of simulations exhibited stationary or mean reverting behaviour. It is also worth mentioning that this variation comes closest to the original WW model with the only difference being that WW used a linear demand curve.

Despite its limitations, the augmented model comes closest to exhibiting characteristics ascribed to real commodity prices. Not only do we obtain a high enough correlation, we also observe fewer stock-outs and the moments are of comparable magnitude to real data. The frequency of bubbles is low but as we have highlighted in Table 5.2b, it is often difficult to detect explosiveness, even in real data, in the presence of mean reverting behaviour which is observed in periods of exceptionally high harvests or high availability. The results also show the importance of each aspect of the augmented model as compared to the standard model. While the convenience yield appears to be the most important addition, the iso-elastic demand and saving behaviour also play a significant role. In the remaining sections of this chapter we proceed with the augmented model.

5.4.3 Impulse Responses

We can also use simulations to derive impulse responses. Figure 5.4 to 5.6 show the impact of positive and negative weather shocks in the first period on availability, storage, equilibrium price and planned production. Due to the non-linearity present in the model, the impact of a positive shock is different from that of a negative shock if the negative shock leads to a stock out. Since the steady state stock is 0.2745 we consider 3 different shocks; a positive 1 standard deviation shock to weather which increases realized production in period t; a negative 1 standard deviation weather shock which reduces storage but does not drive it to 0; a negative 3 standard deviation shock which leads to a stock-out. Note that for a different parameter specification, the steady state stock would be different and it is plausible that a much more frequently occurring weather shock would lead to a stock-out. Additionally, we also consider impacts of persistent shocks i.e. a weather shock that lasts 2 periods (Figure 5.7).

Figure 5.4 shows the response (in terms of percentage deviation from steady state values) to a 1 standard deviation positive shock to the weather for availability, storage, price and planned production. The positive shock leads to a bumper crop; a large proportion of the bumper crop is absorbed by storage which jumps around 40% in the period the shock is realized. The impact of the bumper crop is very persistent and storage stays above its steady state value for well over 30 periods. The impact on Availability and price on the other hand is much smaller. Availability increases by 8% while price falls by 4.5%. Recall that the price elasticity of demand in the baseline case is inelastic which explains why the price does not go into free fall. The inelastic price elasticity of supply is also prominent as there is less than a 1% fall in planned production. Since planned production does not fall a great deal, availability and storage continue to be high in subsequent periods. If supply were more responsive, we would see the response to the shock dissipate much quicker.

The response sizes are similar but of the opposite sign with a negative 1 standard deviation shock. As mentioned before, a negative weather shock which causes a poor crop yield but is not sufficient to

cause a stock-out decreases stocks in the period of the shock and continues to impact stocks for a substantial period of time. The contemporaneous response is through a 40% fall in stock; as before, planned production increases only marginally by 1%.

The impact is much more significant when there is a negative 3 standard deviation shock, as seen in Figure 5.6. We observe a stock-out in the period of the shock; availability falls by 25% and the price jumps by almost 20%. Thus, the response is asymmetric and increases due to the stock-out. Planned production increases by around 3.4%. As before, the impact is persistence and lasts for well over 20 periods although the impact is halved in less than 5 periods.

Contrast these shocks with a 2 standard deviation shocks that occurs in two consecutive periods (see Appendix 5B for an explanation of what we mean by a two period shock). Weather shocks such as these are not uncommon as evident from long droughts observed in a number of countries around the world. Persistent shocks can have a much greater impact on real prices than single period shocks as evident from Figure 5.7. In the second period of the shock, availability falls by 30% and most of the shock is absorbed by storage (70%). Prices jump by around 12% and planned production increases slightly. However, the second period shock compounds the impact greatly. Storage is driven to 0 as availability is now 30% less than steady state values. This results in a nearly 60% jump in real prices. Thus, multi-period shocks tend to explain the large increases in prices much better. Persistent droughts or period of bad weather are more likely to cause large jumps in prices than large shocks that occur for only one period.

From the impulse response analysis we observe that extreme weather events can have a significant impact on the real price of a commodity. The impact on nominal price is likely to be greater still which likely results in the explosive path often observed in asset prices. Our analysis also hints at what could happen if extreme weather becomes more common as a result of climate change. If global temperatures continue to rise, leading to a change in the distribution of weather socks, persistent negative shocks to output may become more common. We also see that in cases where it is not easy to increase production, supply may be slow to respond.

5.5 Sensitivity Analysis

In this section we analyse the sensitivity of the model's solution to its parameters. Parameter values for the baseline case are reported in table 5.1. The sensitivity of the model is considered by changing values of one parameter at a time. For each part of our analysis we changed the value of the parameter of interest and ran 2000 simulations for 700 periods each; we compare the first order auto-correlation, skewness, kurtosis, percentage of stockout periods and percentage of simulated series exhibiting statistically significant explosiveness. These results are reported in Table 5.4. In each case the second column represents the parameter value in the baseline scenario.

The sensitivity is evaluated with respect to depreciation δ, Price elasticity of Demand ϵ_d, Price elasticity of Supply ϵ_s, the standard deviation of the weather shock σ, the real rate of interest r and the parameter C which determines the demand function's reaction to stocks. The most pronounced effects are observed for the weather shock. Reducing the standard deviation of the weather shock considerably increases autocorrelation. This is due to the low frequency of stockouts that results. When weather is more stable and predictable across periods, there are fewer episodes of bad weather which could result in stockouts. As a result we observe that only 0.2% or around 5 to 6 periods out of every 700 see a stockout. Since stocks are carried in almost every period, the price across periods becomes highly auto-correlated. We also observe that the distribution of prices follows the normal distribution more closely as the skewness and kurtosis measures are not statistically significantly different from the skewness and kurtosis of a normal distribution. Due to climate changes over the past few decades, however, this particular case does not appear particularly relevant; a more interesting scenario is that of higher weather volatility.

When the volatility of the weather process is increased to 0.15 from 0.10 the change is significant. The correlation reduces to 0.52 and skewness and kurtosis measures increase considerably, indicating the number of extreme weather events. The number of stockout periods rises to 4.64% and as a result the incidence of extreme prices also tends to increase as evidenced by the number of bubbles detected. While global warming may not lead to such drastic changes in reality (as other parameters such as supply and demand elasticity are also likely to change), it is nevertheless a good indication of the direction of change if weather becomes more volatile.

Other important determinants of the price process are the price elasticities of demand and supply. Increasing the magnitude of the price elasticity of demand, ϵ_d, increases the correlation while reducing skewness and kurtosis. It also tends to increase the degree of stockouts; however, this does not always lead to an increase in the number of bubbles detected. The price elasticity of demand acts through the demand function (equation 5.7) and the storage function. With a higher price elasticity of demand, quantity demanded reacts more to a given change in price. If price increases the fall in consumption will be greater in the current period; simultaneously an increase in price encourages investors to hold more inventories. Thus, higher price elasticity translates into larger changes in quantity demanded and thus, storage. With higher storage we are likely to observe a higher first-order auto-correlation.

Supply elasticity ϵ_s on the other hand has the opposite impact. If we have perfectly inelastic supply i.e. $\epsilon_s = 0$, correlation increases, stockouts reduce but the incidence of explosive prices increases. If supply is inelastic, the impact of the weather shock is magnified as producers are unable to respond to changes in prices and the amount of supply solely depends on the weather shock. If supply is perfectly inelastic, the amount supplied every period, h_t is more predictable leading to more predictable

outcomes for consumption and inventory which we observe through the high auto-correlation. If extreme weather events are observed, we see a much higher increase in equilibrium price, which explains the higher incidence of bubbles when $\epsilon_s = 0$. When supply elasticity is higher, the auto-correlation tends to be lower and the equilibrium price is more skewed. The impact on the frequency of bubbles is less clear when price elasticity is high.

The other parameters do not significantly impact our results. Due to the inclusion of the convenience yield the impact of the real interest rate becomes negligible. The depreciation rate on the other hand tends to have a small impact on the auto-correlation of equilibrium price. This impact comes through availability in equation (5.6) rather than the arbitrage equation (5.8). If the convenience yield were not included, both depreciation and the real rate of interest would likely have had a higher impact.

5.6 Empirical Results

As mentioned at the outset, a substantial amount of literature exists on the commodity model with storage but its empirical success has been limited. In this section, we show that our augmented model is able to capture features of actual data. Getting reliable estimates for price elasticities of demand and supply has been a difficult task, worthy of its own chapter. We rely on recent methods introduced by Roberts and Schlenker (2013) and some of their data to estimate the price elasticity of demand, ϵ_d and use the estimated price elasticity of demand in our model to analyse if the model is able to capture features of real data. Roberts et al also lay out a methodology for measuring the price elasticity of supply by converting production data into calories. We do not use estimates for price elasticity of supply in our empirical analysis as that is beyond the scope of this article.

Identification of the price elasticity of demand requires the use of instrumental variables as in equilibrium both quantities and prices are determined simultaneously. Roberts et al use two such instruments in their article to measure elasticities for corn, yield shocks and weather related data. We use their calculated weather shocks for our estimation. We rely on weather based instruments instead of yield shocks because the yield shock data they have calculated is specific for each commodity and relies on the availability of yield data; this data is not available for all the commodities that we consider.

The two climate related variables that Roberts et al employ are average temperature in Celsius in the calendar year and average precipitation in the calendar year, both observed in the growing region of the crop. Both averages are found using the University of Delaware's Climactic Research Centre data. They identify geographical regions where the commodities in their study are grown and using the database they find average annual temperature and average annual precipitation; these are then combined into annual global measures using area based weights. Since the growing regions for their commodities and ours tend to coincide (most production takes place in China, India, Brazil and the

United States), we use their average temperature and precipitation variables as instruments for our study. While this measure may not be perfect it is nevertheless likely to be highly co-related with actual weather data for our commodities and thus, fulfil the relevance criterion for instrumental variables. Note that yield shocks may themselves be endogenous as Roberts et al point out in their article. Weather related variables on the other hand are more likely to be exogenous as well as relevant.

Thus, we use two stage least squares (TSLS) to estimate the following equation:

$$lnQ_t = \alpha + \epsilon_d lnP_t + \eta_t \qquad (5.9)$$

Q_t is consumption demand in year 't', P_t is average real price of the commodity in year 't', ϵ_d is the price elasticity of demand and η_t is an i.i.d error term. Note that (5.9) is an equilibrium relationship; thus the error term includes the impact of a number of variables that were part of the structural model; this necessitates the use of instrumental variables. The list of instruments includes average temperature in Celsius, the squared of average temperature in Celsius, annual average precipitation in metres and squared average precipitation in metres. Data on the instruments are available from 1961-2008.

We use the same real price data that were used to derive the results in Table 5.2 (Grilli and Yang). Data on world consumption for 4 of the 5 commodities were obtained from the United States Department of Agriculture's Foreign Agriculture Service database. The database contains data on production and inventories. These data are only available from 1960, so equation (5.9) is estimated using data from 1961-2008. All data are annual. For Maize we rely on data from the Food and Agriculture Organization (FAO) of the United Nations. For each commodity, consumption was found by subtracting the change in inventories from production during the year i.e. $Q_t = h_t + S_{t-1} - S_t$.

Equation (5.9) is estimated for all 5 commodities. The estimation results for the TSLS regressions are available upon request. In summary, our qualitative findings are similar to Roberts et al. While the first stage coefficients are often individually insignificant, the regression itself is jointly significant and the second stage results are always significant, both individually and jointly. Roberts et al rely on futures data while we use measures of spot price; our primary data source is FAS while Roberts et al primarily rely upon FAO data. Column 2 in table 5.5 reports our point estimates for the price elasticity of demand for each commodity, which are obtained through estimating equation (5.9). The elasticity of supply ϵ_s is fixed at 0.3.

Using the estimated price elasticity of demand we find the solution of our augmented storage model for each commodity. Note from Table 5.4(c) that changing the elasticity of supply can have an impact on the solution which in turn could impact the average auto-correlation and moments. For simplicity,

we keep the elasticity of supply fixed for all commodities although we recommend using appropriate measures of supply elasticity if the model is to be used for policy recommendations. The numerical solution is then used to simulate 2000 price series (each being 700 periods long) and relevant moments are found as averages. Columns 3 and 4 report actual and estimated first order auto-correlations, columns 5 and 6 report actual and estimated measures of skewness while columns 7 and 8 report actual and estimated measures of kurtosis respectively.

Despite the fact that we are relying on only one parameter for our estimates, the model tends to perform well, particularly when first-order auto-correlation measures are considered. As seen from table 5.5, actual and estimated auto-correlation are much closer than have been observed in previous studies. The largest error occurs for Sugar where the estimated first order auto-correlation is 0.878 compared to the actual correlation of 0.689. Based on our study the results for sugar could be weaker due to sugar having a higher price elasticity of supply or the instrumental variable being inappropriate (as sugar may be grown in different geographical areas). We do not attempt to reconstruct the instrumental variable for each commodity as that would be beyond the scope of this chapter. Correlations for Coffee, Sugar, Maize and Wheat are overestimated while those for Cotton are under-estimated.

Measures of skewness and kurtosis on the other hand appear to be much more sensitive to changes in parameters. The errors are of larger magnitude in absolute as well as relative terms which highlights the shortcomings of our approach of using only one parameter. We believe that a more thorough study using more accurate parameter estimates will be able to match empirical estimates more closely. Measures of skewness in particular are underestimated and are often negative. Kurtosis measures on the other hand are much better and closer together with Cotton being the exception. Nevertheless, our brief foray into the empirical literature highlights the viability of our augmented model. Using just one empirically calculated parameter we are able to add a lot of respectability to the results obtained through the commodity storage model. With more precisely calibrated parameters the model would be able to better replicate the distribution of real commodity prices and also allow the researcher to measure impulse responses to changes in parameters of interest, such as the volatility of weather.

5.7 Conclusion

We have attempted to address the shortcomings of the commodity model with storage developed by William and Wright by augmenting the original model. This has involved modelling demand as iso-elastic, incorporation of consumption smoothing/saving behaviour in the quantity demanded function and the inclusion of a multiplicative convenience yield. As opposed to the original model, the

augmented model exhibits characteristics typically found in real commodity pricing data; this includes a high first order auto-correlation, positive skewness and excess kurtosis.

In addition to these distributional features our research is also the first to formally test for explosiveness in commodity prices in this literature. We have used Philips, Shi and Yu's GSADF statistics for this purpose. We note that when these tests are conducted on real as opposed to nominal prices, the GSADF test is often unable to detect explosiveness. The reason for this is the presence of a mean reverting state in the real price series which is usually not significantly long in a nominal price series. Using the solution to the augmented commodity storage model we have shown simulated prices tend to resemble real as opposed to nominal commodity prices and thus, do not often exhibit statistically significant explosiveness even though high peaks and troughs continue to be present due to excess kurtosis.

Impulse Response analysis shows asymmetric effects of negative and positive shocks when negative shocks lead to a stock-out. If quantity demanded and planned production are inelastic functions of price, most of the impact of the weather shock is absorbed by inventory owners. Price tends to respond by a greater magnitude if the weather shock leads to a stock-out. Irrespective of the size and direction of the shock, the impact of a weather shock on storage tends to be persistent. We also conducted a sensitivity analysis of the model solution and showed that the first order auto-correlation tends to be most sensitive to the variance of the weather shock, and the price elasticities of demand and supply respectively.

Empirical analysis further supports the augmented commodity storage model. We use a measure of price elasticity of demand to find the first order auto-correlation for 5 storable commodities. Our empirical results are encouraging and while skewness and kurtosis are not always accurately captured, the first-order auto-correlation is estimated well. We believe that correctly calibrated parameters will further improve results.

Future research areas are also highlighted. The model itself may be extended and incorporated in a broader General Equilibrium framework ala Arseneau et al. The augmented model itself can be used to analyse the impact of different types of weather shocks or different policies. Finally, the empirical validity of the model may be further scrutinized by using more precise calibration.

Thus, this chapter shows that the commodity storage model still has traction and the augmentations we have added turn it into a useful tool to understand the impact of different policies or different shocks to storable commodities.

APPENDIX 5A – Tables and Figures

Table 5.1: Calibrated values

Parameter	σ	$\bar{h}$	δ	r	k	ϵ_s	ϵ_d	C	L
Calibration	0.10	1	0.02	0.05	0	0.2	-0.3	0.98	50

Table 5.2a – Moments of Nominal Price series

	Coffee	Cotton	Maize	Sugar	Wheat
Auto-correlation	0.854	0.837	0.809	0.721	0.851
Skewness	1.263	0.911	1.573	2.073	1.740
Kurtosis	4.026	4.375	6.800	7.837	6.909
PSY Test Result	Yes	No	Yes	Yes	Yes

Table 5.2b – Moments of Real Price series

	Coffee	Cotton	Maize	Sugar	Wheat
Auto-correlation	0.829	0.936	0.855	0.689	0.904
Skewness	1.608	0.203	0.839	1.617	0.823
Kurtosis	6.886	2.390	4.220	6.448	3.393
PSY Test Result	No	No	No	No	No

Table 5.3 – Properties of simulated series

	Augmented Model	Iso-elastic Demand	Linear Demand	No CY
Correlation	0.7960	0.6443	0.4986	0.1889
Skewness	1.4411	5.1331	4.231	2.0812
kurtosis	15.3942	57.2949	34.6041	9.2586
stockouts	1.61%	.99%	2.79%	36.29%
Explosive series	14.12%	62.8%	15%	0%

Table 5.4 (a) Depreciation

	$\delta = 0.02$	$\delta = 0.00$	$\delta = 0.05$	$\delta = 0.10$
Correlation	0.7960	0.8030	0.7712	0.7302
Skewness	1.4411	1.4006	1.7929	2.1952
kurtosis	15.3942	14.3836	19.6657	24.8560
stockouts	1.61%	1.65%	1.70%	1.32%
Bubbles	14.12%	16%	18.0%	15.2%

Table 5.4 (b) Demand Elasticity

	$\epsilon_d = -0.3$	$\epsilon_d = -0.2$	$\epsilon_d = -0.5$	$\epsilon_d = -1.0$
Correlation	0.7960	0.7024	0.8810	0.9472
Skewness	1.4411	3.4439	0.0661	-0.5111
kurtosis	15.3942	42.5758	4.4387	2.5124
stockouts	1.61%	1.17%	2.51%	3.63%
Bubbles	14.12%	30.0%	2.6%	0.6%

Table 5.4 (c) Supply Elasticity

	$\epsilon_s = 0.2$	$\epsilon_s = 0.0$	$\epsilon_s = 0.5$	$\epsilon_s = 1.0$
Correlation	0.7960	0.9251	0.7243	0.6579
Skewness	1.4411	0.4191	2.3217	3.1658
kurtosis	15.3942	4.9359	26.8832	40.0015
stockouts	1.61%	1.09%	1.57%	1.14%
Bubbles	14.12%	38.8%	15.8%	16.4%

Table 5.4 (d) Sensitivity of weather shock

	$\sigma = 0.10$	$\sigma = 0.05$	$\sigma = 0.11$	$\sigma = 0.15$
Correlation	0.7960	0.9235	0.7451	0.5175
Skewness	1.4411	-0.3498	2.1271	5.2858
kurtosis	15.3942	3.3257	21.6595	61.7512
stockouts	1.61%	0.18%	2.19%	4.64%
Bubbles	14.12%	0.2%	17.65%	34.8%

Table 5.4 (e) Interest Rate

	$r = 0.05$	$r = 0.0$	$r = 0.03$	$r = 0.10$
Correlation	0.7960	0.8021	0.7990	0.7870
Skewness	1.4411	1.3991	1.4356	1.5478
kurtosis	15.3942	14.0282	15.7072	16.5785
stockouts	1.61%	1.55%	1.55%	1.73%
Bubbles	14.12%	13.8%	12.2%	13.8%

Table 5.4 (f) Consumption's sensitivity to stocks

	$C = 0.98$	$C = 0.95$	$C = 0.99$	$C = 1.0$
Correlation	0.7960	0.8043	0.7880	0.7861
Skewness	1.4411	1.2822	1.5547	1.5647
kurtosis	15.3942	13.4020	17.2622	16.2360
stockouts	1.61%	1.66%	1.65%	1.67%
Bubbles	14.12%	12.2%	16%	15.2%

Table 5.5 – Emprical estimates

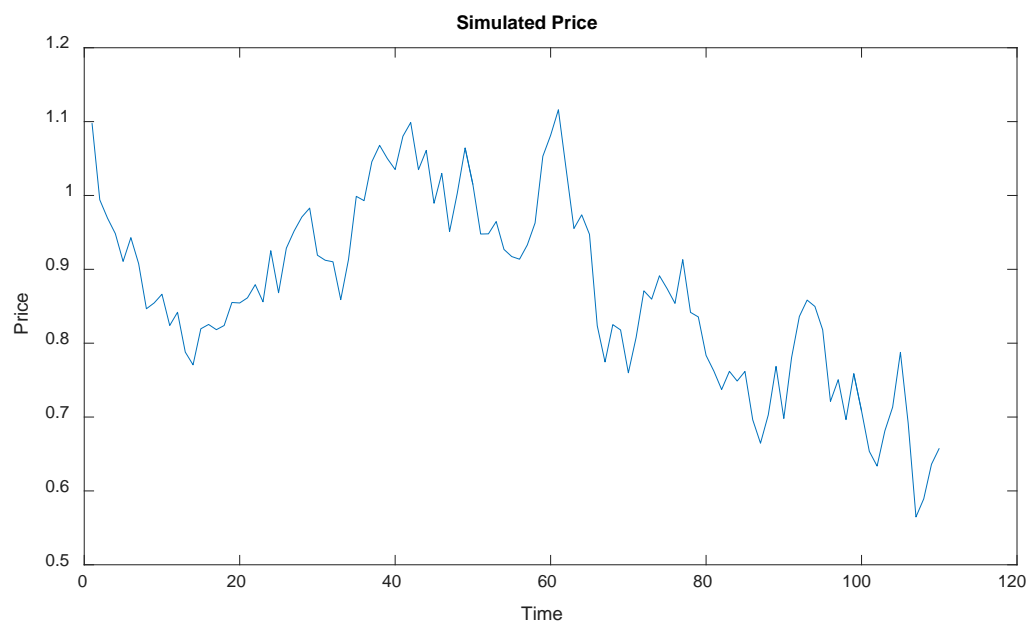| Commodity	ϵ_d	Actual Correlation	Estimated Correlation	Actual Skewness	Estimated Skewness	Actual Kurtosis	Estimated Kurtosis
Coffee	-0.484	0.829	0.8540	1.608	0.2601	6.886	5.7091
Cotton	-0.497	0.936	0.8580	0.203	0.2088	2.390	5.5223
Maize	-0.743	0.855	0.9058	0.839	-0.3757	4.220	3.1788
Sugar	-0.573	0.689	0.8777	1.617	-0.0515	6.448	4.2348
Wheat	-0.746	0.904	0.9087	0.823	-0.3808	3.393	3.1379

Figure 5.1: Equilibrium Functions
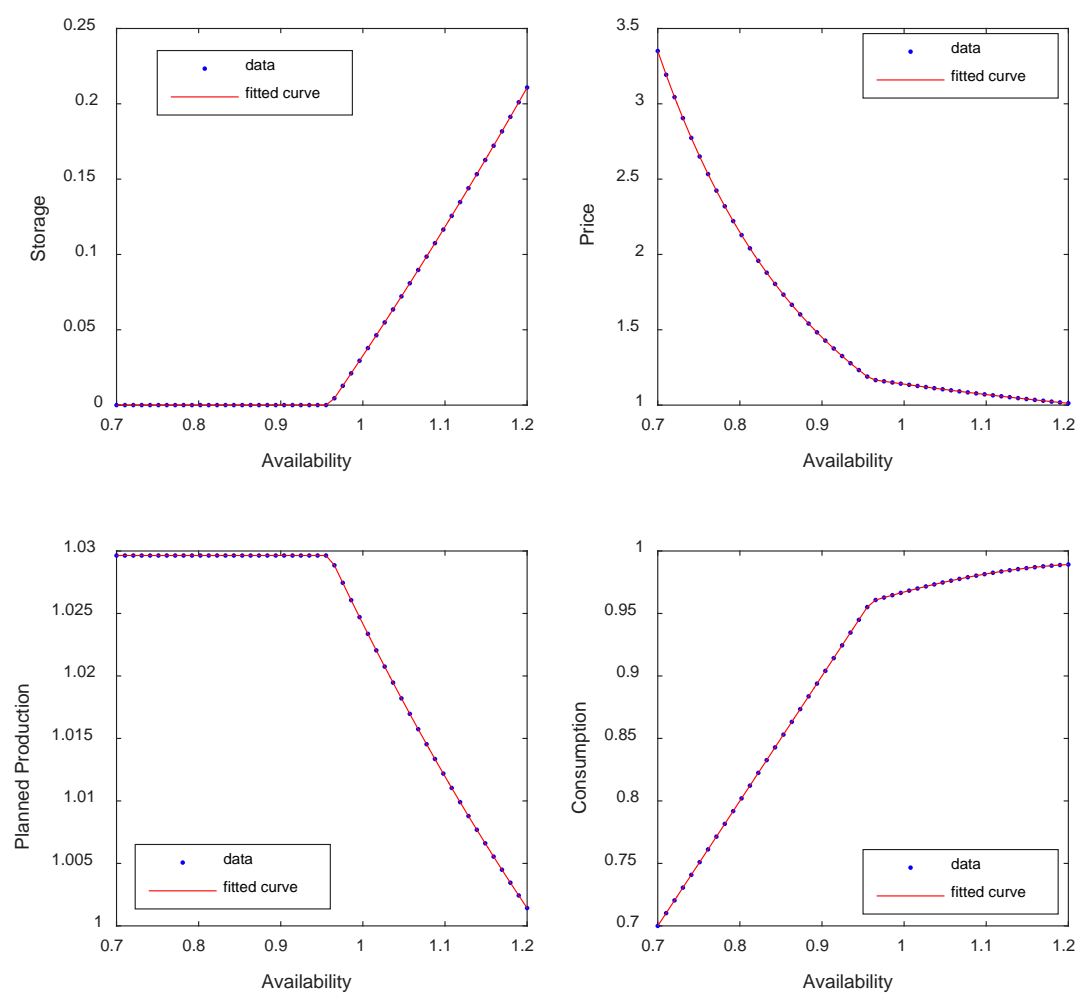
Figure 5.2: Nominal and Real Price Series - Sugar
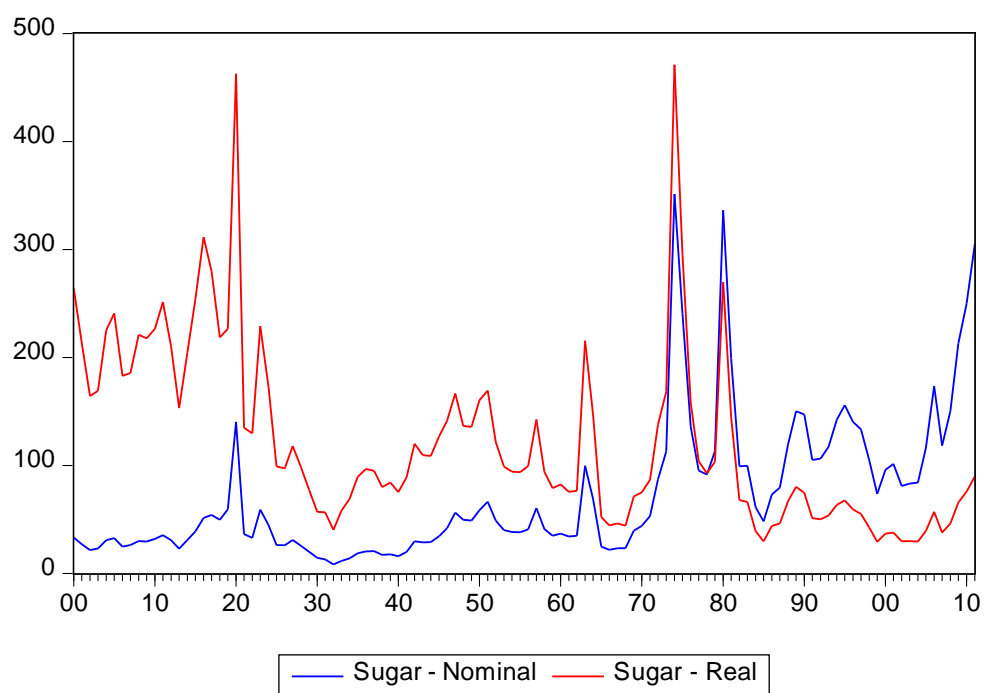
Figure 5.3: Simulated Price using Augmented Model Solution

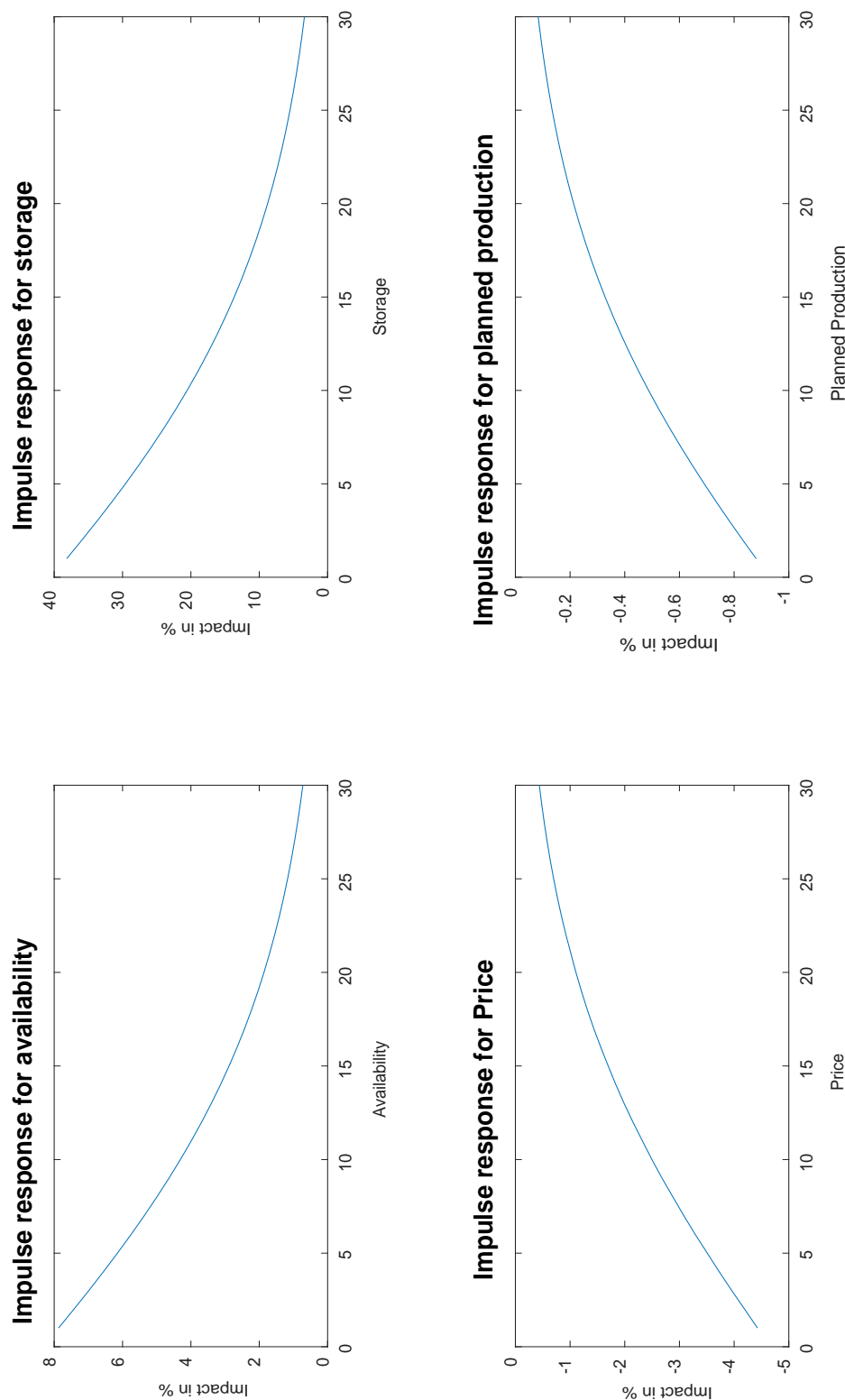# Figure 5.4: Impulse Responses for a Positive 1 s.d. shock

Impulse response for availability

Impulse response for storage

Impulse response for Price

Impulse response for planned production

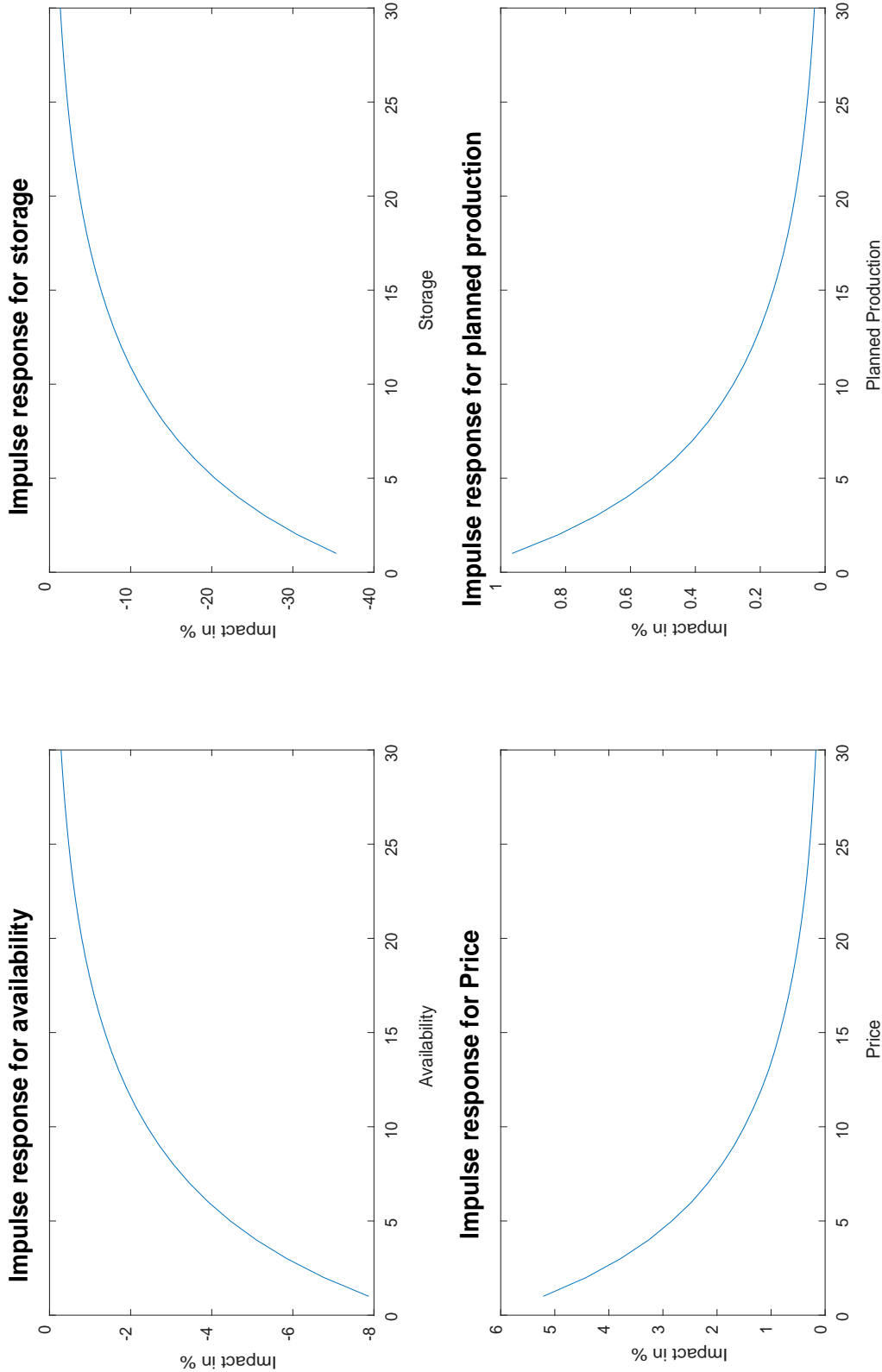**Figure 5.5: Impulse Responses for a Negative 1 s.d. shock**

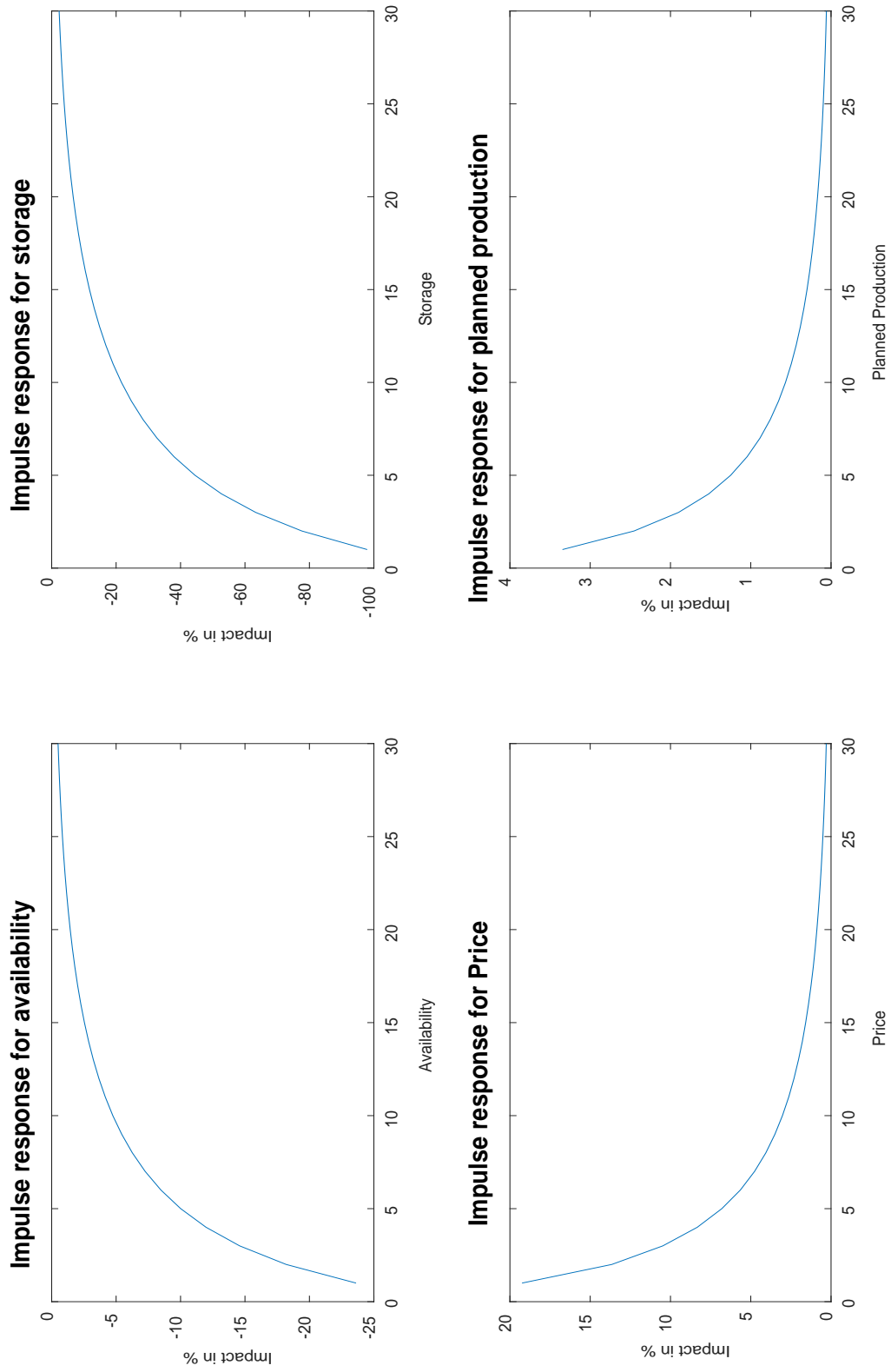Figure 5.6: Impulse Responses for a Negative 3 s.d. shock

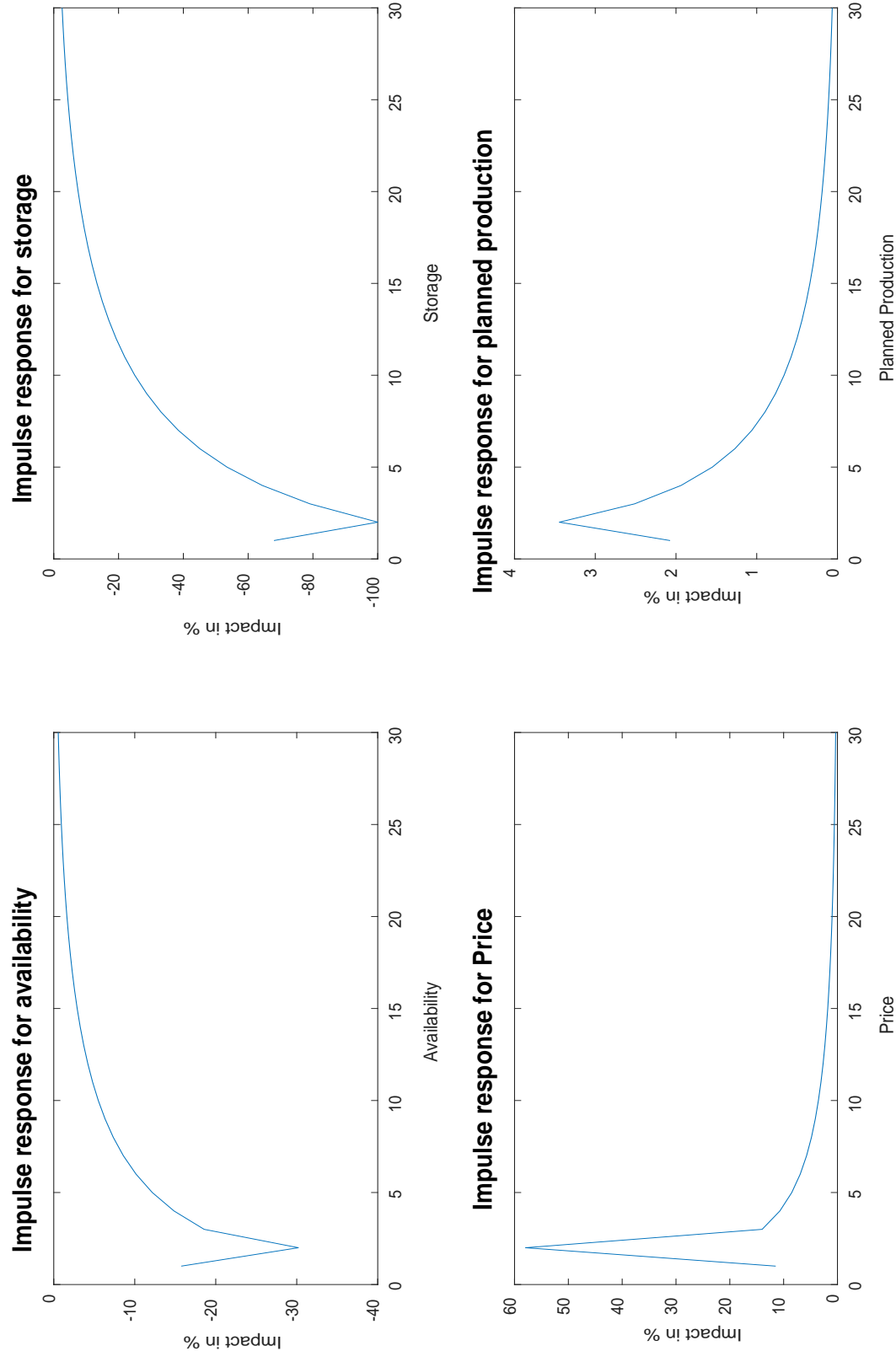**Figure 5.7: Impulse Responses for a 2 period 2 s.d shock shock**

Impulse response for availability

Impulse response for storage

Impulse response for Price

Impulse response for planned production

APPENDIX 5B: IMPULSE RESPONSES

This appendix explains the impulse responses that we generate in section 5.4.3.

We define an impulse response to a shock of size δ in period $t - k$ as

$$\Delta y(t, k, \delta) = y(t, k, \delta) - y(t, k, 0) \qquad\qquad \mathbf{5B - 1}$$

Where $y(t, k, \delta)$ is the shocked variable (e.g. price) at time t that results from the shock $\delta + \epsilon_{t-k}$ at time $t - k$ (ϵ refers to the iid shock in the shocked variable's process) whilst $y(t, k, 0)$ is the value the variable takes on if no additional shock is introduced.

If the data generating process of variable y can be represented as an $MA(\infty)$

$$y(t, k, 0) = y(t) = \sum_{j=0}^{\infty} \phi_j \epsilon_{t-j} \qquad\qquad \mathbf{5B - 2}$$

And

$$y(t, k, \delta) = \sum_{j=0}^{k-1} \phi_j \epsilon_{t-j} + \phi_k(\epsilon_{t-k} + \delta) + \sum_{j=k+1}^{\infty} \phi_j \epsilon_{t-j} \qquad \mathbf{5B - 3}$$

This implies that $\Delta y(t, k, \delta) = \phi_k \delta$ $\qquad\qquad\qquad\qquad$ $\mathbf{5B - 4}$

If the data generating process cannot be represented as an $MA(\infty)$ the value of the response will be different. The Impulse responses we generate in section 5.4.3 tend to impact the process for over 30 years and likely have a different analytical form; however, the MA representation will suffice as an explanation.

If instead of a 1 period shock, we have shocks in multiple periods e.g. δ in period $t - k$ and δ in period t-k-1, we get a similar result. We define a 2 period impulse response in the following way:

$$\Delta y(t, k, k + 1, \delta, \delta) = y(t, k, k + 1, \delta, \delta) - y(t, k, k + 1, 0, 0) \qquad\qquad \mathbf{5B - 5}$$

Where $y(t, k, k + 1, \delta, \delta)$ is the value of the shocked variable at time t that results from the shock $\delta + \epsilon_{t-k}$ in period $t - k$ and $\delta + \epsilon_{t-k-1}$ in period $t - k - 1$. The definition of $y(t, k, k + 1, 0, 0)$ follows as before.

Analogous to 5A-4 above, this implies that:

$$\Delta y(t, k, k + 1, \delta, \delta) = \phi_k \delta + \phi_{k+1} \delta \qquad\qquad \mathbf{5B - 6}$$

Thus, if have 'm' shocks $(\delta_1, \delta_2, \ldots, \delta_m)$ in 'm' consecutive periods $(t - k, t - (k + 1), \ldots, t - (k + m))$, the impulse response will be:

$$\Delta y(t, k, \ldots, k + m, \delta_1, \ldots, \delta_m) = \sum_{j=k}^{k+s} \phi_j \delta_{j-k+1} \qquad 5A - 7$$

APPENDIX 5C: DATA

Table 5C-1: Data units for Commodity Consumption

Commodity	Units
Coffee	1000 60 KG bags
Cotton	1000 480 lb. Bales
Maize	1000 Metric Tonnes
Sugar	1000 Metric Tonnes
Wheat	1,000,000 Metric Tonnes

CHAPTER 6: *Conclusions*

By treating market efficiency as one of several states, we have provided different methodologies which allow us to estimate asset prices and measure the frequency of market inefficiency. This is an important contribution to the market efficiency literature as previous attempts have treated market efficiency as a binary concept. Market efficiency is desirable for an asset market as it negates the possibility of excess returns being earned by some market agents at the expense of others. Thus, our efficiency metrics could enable market experts to identify situations in which a market may be tending towards inefficiency. For some markets this may help policy makers identify when a policy change or intervention may become necessary. In addition to showing the utility of our methodologies using simulations we have also shown that they hold empirical relevance and in Chapter 4 we have also made theoretical contributions. Thus, this dissertation has complemented applied work with empirical and theoretical contributions.

Our analysis began with a detailed overview of the literature on market efficiency with an emphasis on explosiveness. We presented an econometric test that enabled us to identify periods where commodity markets may have been explosive. By contrasting this test with existing tests we showed that it was a viable alternative for detecting market efficiency. Its main shortcoming was its inability to precisely date periods of inefficiency; however, the test performs better than alternatives when small samples our considered. Thus, Chapter 2 provided a brief introduction and insight into the crux of this dissertation. We found that for commodities most episodes of explosiveness were concentrated around the oil price crises in the 1970s and the financial crisis of 2007-08.

We then showed that for a storable commodity, an auto-regressive switching state reduced form is possible under rational expectations. This allows us to estimate asset prices through state-switching autoregressive models. The first approach that we introduce uses Markov-state switching to estimate asset prices. Our contribution to this approach is to use a transition matrix derived from an adaptation of the PSY procedure (explained in Chapter 3). By using information on explosiveness and mean reversion in the sample, we are able to estimate a transition matrix. If this transition matrix is used in the maximum likelihood procedure for estimating a Markov switching autoregressive model we get much better results than the case where the transition matrix is directly estimated. We show that we gain a high probability for explosiveness in time periods which were typically associated with periods of explosiveness (such as the financial crisis from 2007-08) and get a high probability for mean reversion in time periods where prices were unusually low (such as the post financial crisis period).

The second methodology for detecting market inefficiency uses threshold autoregressions instead of Markov-switching. Since market efficiency implies a random walk process for log asset prices, we first derived conditions which need to be satisfied for the existence of a mean and a variance in the

steady state. Building upon the results of these formulae we show that if asset returns or asset prices can be estimated using a threshold autoregression where states switch based on an exogenous trigger variable, we can identify periods in which asset markets exhibit inefficiency. This serves as a metric and enables us to calculate proportional efficiency for asset markets. To depict our methodology we empirically estimated the S&P500 and the FTSE100. Our results show that with the most general threshold model (i.e. one with a switching drift), the S&P500 and the FTSE100 are mostly efficient with period of inefficiency concentrated around the dot-com bubble and the oil price crises. We also highlighted how these results may change if model specification is not considered appropriately. Additionally, we also provided a limitation of tests of explosiveness and showed that when an asset market exhibits mean reversion, explosiveness may be difficult to detect using econometric tests.

Finally, we modified a structural model for the real price of a storable commodity. We made 3 changes to the original model and solved it using numerical procedures. The 3 changes introduced a convenience yield, incorporated saving behaviour and used an iso-elastic demand curve instead of a linear one. The resulting solution is able to capture autocorrelations of real commodities when calibrated appropriately. In addition, it is also does a reasonable job in capturing other distributional properties such as skewness and kurtosis. Calibrating the model for real data is not a trivial exercise and requires the identification of elasticities using data on price and quantities. We believe that the augmented commodity storage model could prove useful for policy analysis particularly in the light of climate change which is likely to be become a growing concern in the future and which is also the main mechanism through which commodity prices go through periods of explosiveness and mean reversion. Through simulation analysis we have also shown that periods of adverse weather, particularly those resulting from persistent negative shocks can have long lasting impacts and can cause the kind of explosive behaviour in commodity markets that is often referred to as a bubble.

Thus, this dissertation provides a new perspective on market efficiency by treating it as a state. Not only have we provided objective methodologies which can indicate when markets are efficient or inefficient but have also sought to explain why these periods occur by augmenting the commodity storage model. A number of research streams open up as a result. We wish to apply our methodologies to more asset markets, particularly the market for exchange rates. From an econometric perspective additional theoretical results may be derived for the case where the threshold autoregressive model has a Markov-chain trigger variable instead of an exogenous trigger variable. The commodity storage model can also be extended or modified to make it more relevant to analyse other commodities such as metals.

Bibliography

'A History of Iron Ore Pricing,' *The Mining Journal,* Published online 28 Sep 2012. <http://www.mining-journal.com/reports/iron-ore-a-history-of-iron-ore-pricing>

Abreu, D and M. K. Brunnermeier (2003). Bubbles and Crashes, *Econometrica,* Vol. 71(1), pp. 173-204

Albert, J. H. and S. Chib (1993). Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts, *Journal of Business and Economic Statistics,* Vol. 11 (1), pp. 1-15

Alessandri, P. (2006). Bubbles and fads in the stock market: another look at the experience of the US, *International Journal of Finance and Economics*, Vol 11(3), pp. 195-203

Arseneau, D. M. and S. Leduc (2013). Commodity Price Movements in a General Equilibrium Model of Storage. *IMF Economic Review,* Vol. 61(1), pp. 199-224.

Baffes, John. (2004). Cotton: Market Setting, Trade Policies, and Issues. World Bank, Washington, D.C.. © World Bank. Retrieved from <https://openknowledge.worldbank.org/handle/10986/14735>

Bhargava, A. (1986), On the Theory of Testing for Unit Roots in Observed Time Series. *The Review of Economic Studies,* Vol 53(3), pp. 369-384

Blanchard, O. (1979), Speculative Bubbles, Crash and Rational Expectations. *Economic Letters,* Vol 3, pp. 387-389

Blanchard, O. J. and Mark W. Watson (1982). Bubbles, rational expectations, and financial markets. *NBER Working paper number 945*

Bohl, M. (2003). Periodically Collapsing Bubbles in the US Stock Market. *International Review of Economics and Finance,* Vol. 12, pp. 385-397

Boudoukh, J., Matthew Richardson, YuQing Shen and Robert Whitelaw (2007). Do asset prices reflect fundamentals? Freshly squeezed evidence from the OJ market. *Journal of Financial Economics,* 83, pp. 397-412

Breitung, J. and U. Homm, (2010). Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods. *Journal of Financial Econometrics,* Vol 10(1), pp. 198-231

Brooks, C and M. Prokopczuk (2013). The dynamics of commodity prices. *Quantitative Finance,* Vol. 13(4), pp. 527-542, DOI: 10.1080/14697688.2013.769689

Brooks, C., M. Prokopczuk and Yingying Wu (2015). Booms and Busts in Commodity Markets: Bubbles or Fundamentals? *Journal of Futures Markets,* Forthcoming.

Bureau of Labor Statistics. 1890-2014, "Consumer Price Index, All Urban Consumers," *US Department of Labour,Retrieved from* <https://www.minneapolisfed.org/community/teaching-aids/cpi-calculator-information/consumer-price-index-1800>.

Caballero, R J, Emmanuel Farhi and Pierr-Olivier Gourinchas (2008). Financial Crash, Commodity Prices and Global Imbalances. *Brookings Papers on Economic Activity,* pp. 1-55.

Camerer, C., (1989). Bubbles and Fads in Asset Prices. *Journal of Economic Surveys,* Vol. 3 (1), pp. 3-41.

Campbell, J. Y and R. J. Shiller, 1987. Co-integration and test of present value models. *Journal of Political Economy,* Vol 95 (5), pp. 1062-1088.

Carvalho, V. A. Martin and J. Ventura (2012). Understanding Bubbly Episodes. *American Economic Review,* Vol. 102(3), pp. 95-100.

Caspi, I. (2013). Rtadf: Testing for Bubbles with Eviews. MPRA Paper 58791, *University Library of Munich, Germany*, revised 06 Sep 2014.

Chambers, M and Bailey R (1996). A theory of commodity price fluctuations. *Journal of Political Economy,* Vol. 104(5), pp. 924-957.

Chang, E. C (1985). Returns to Speculators and the Theory of Normal Backwardation. *The Journal of Finance,* Vol. 40(1), pp. 193-208.

Charles, A. and O. Darné (2009). Variance-ratio Tests of Random Walk: An Overview. *Journal of Economic Surveys*, Vol. 23, pp. 503–527.

De Bondt, W. F. M. and R. Thaler (1985). Does the stock market overreact? *The Journal of Finance* Vol 40(3), pp. 793-805.

Deaton A., 1999. Commodity Prices and Growth in Africa. *Journal of Economics Perspectives,* Vol. 13(3), pp. 23-40.

Deaton, A. and G. Laroque (1992). On the behaviour of commodity prices. *Review of Economic Studies,* Vol 59(1), pp. 1-23.

---- (1995). Estimating a nonlinear rational expectations commodity price model with unobservable state variables. *Journal of Applied Econometrics,* Vol. 10, pp. S9-S40.

---- (1996). Competitive Storage and Commodity Price Dynamics. *Journal of Political Economy,* Vol. 104 (5), pp. 896-923.

Diba, B T., and Herschel I. Grossman (1984). Rational bubbles in the price of gold. *NBER Working paper 1300.*

------ (1988). Explosive bubbles in stock prices. *American Economic Review,* Vol. 78 (3), pp. 520-530.

Edelstein, D. (1999). Copper, *US Geological Survey*, Retrieved from <minerals.usgs.gov/minerals/pubs/commodity/copper/240798.pdf>.

Ellen, Saskia ter and C. J. Zwinkels (2010). Oil Price Dynamics: A behavioural finance approach with heterogenous agents. *Energy Economics,* Vol 32 (6), pp. 1427-1434.

Enders, W. and P. L. Siklos (2001). Co-integration and threshold adjustments. *Journal of Business and Economic Statistics,* Vol. 19(2), pp. 166-176.

Eo, Y. and Chang-Jin Kim (2015). Markov-Switching Models with Evolving Regime-Specific Parameters: Are Post-War Booms or Recessions All Alike? *Working Paper 2015-04, University of Sydney, School of Economics.*

Etienne, X. L, S. H. Irwin and P. Garcia (2014). Bubbles in food commodity markets: Four decades of evidence. *Journal of International Money and Finance,* Vol. 42, pp. 129-155.

Evans, G.W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review,* Vol 81, pp. 1189-1214.

Fama, E. F. (1965). The behaviour of stock-market prices. *Journal of Business,* Vol 38(1), pp. 34-105.

Fama E., (1970). Efficient Capital Markets: A review of theory and empirical work. *Journal of Finance,* Vol. 25 (2), pp. 383-417.

Fama, E and Kenneth R. French (1987). Commodity Futures Prices: Some Evidence on Forecast Power, Premiums and the Theory of Storage. *The Journal of Business,* Vol. 60 (1), pp. 55-73.

Fama, E. F, and K. R. French (1988). Permanent and Temporary Components of stock prices. *Journal of Political Economy,* Vol. 96(2), pp. 246-273.

Financial Times, Market Data. Retrieved from www.ft.com.

Fischer, B (1976). The Pricing of Commodity Contracts. *Journal of Financial Economics,* Vol. 3, pp. 167-179.

Flood, R P and Peter M. Garber (1980). Market funadmentals versus price-level bubbles: The first tests. *Journal of Political Economy,* Vol. 88, pp. 745-770.

Food and Agricultural Organization (FAO) Statistics Division. 1960-2010. Production – Crops; Food Balance Sheets – Commodity Balances – Crops Primary Equivalent; Resource – Fertilizer Archive. Retrieved from <http://faostat.fao.org>.

Foreign Agricultural Service (FAS). 1960-2010. "Downloadable Data Sets – All Commodities." *United States Department of Agriculture.* Retrieved from http://www.fas.usda.gov/psdonline/psdDownload.aspx.

Frankel, J. A. (1986). Expectations and Commodity Price Dynamics: The Overshooting Model. *American Agricultural Economics association,* Vol. 68(2), pp. 344-348.

----- (2007), The Effect of Monetary Policy on Real Commodity Prices, *NBER Working Paper Series.*

Frankel, J and Andrew Rose (2010). Determinants of Agricultural and Mineral Commodity Prices. *Harvard Working Paper Series.*

Franklin A and Gary Gorton, (1993). Churning Bubbles. *The Review of Economic Studies,* Vol. 60(4), pp. 813-836.

Friedman, D and Masanao Aoki (1986). Asset Price Bubbles from Poorly Aggregated Information: A parametric example. *Economics Letters,* Vol. 21, pp. 49-52.

Froot, K A. and Maruice Obstefeld (1991). Intrinsic Bubbles: The Case of Stock Prices. *The American Economic review,* Vol. 81(5), pp. 1189-1214.

Goldfeld, S.M. and R.E. Quandt (1973). A Markov Model for Switching Regressions. *Journal of Econometrics,* Vol 1(1), pp. 3-16.

Google Finance (2015). *S&P 500 index.* Retrieved from Google Finance database https://www.google.co.uk/finance.

Google Finance (2015). *FTSE 100 Index.* Retrieved from Google Finance database https://www.google.co.uk/finance.

Gorton, G B, Fumio Hayashi and Geert Rouwenhorst (2012). The Fundamentals of Commodity Futures Returns. *Review of Finance,* Vol. 17(1), pp. 35-105.

Gouel, C. (2013). *C*omparing Numerical Methods for Solving the Competitive Storage Model. *Computational Economics,* Vol. 41 (2), pp. 267-295.

Grilli, E., and M. C. Yang (1988). Primary Commodity Price, Manufactured Good Prices, and Terms of Trade of Developing Countries: What the Long Run Shows. *The World Bank Economics Review,* Vol 2(1), pp. 1-47.

Grossman, S J and Joseph E. Stiglitz (1981). The determinants of the variability of stock market prices. *American Economic Review,* Vol. 71, pp. 222-227.

Guerra, V. E. A., Bobenrieth, E. S. A.H., Bobenrieth, J. R. A. H. and C. Cafiero (2014). Empirical commodity storage model: the challenge of matching data and theory. *European Review of Agriculturlal Economics,* Vol. 42(4), pp. 607-623.

Gustafson, R. L. (1958). Carryover levels for gains: A method for determining amounts that are optimal under specified conditions. *USDA Technical Bulletin 1778,* 1958a.

Gutierrez, L., Piras, F. and P. P. Roggero (2015). A Global Vector Autoregression Model for the Analysis of Wheat Export Prices. *American Journal of Agricultural Economics,* Vol. 97(5), pp. 1494-1511 .

Hall, S.G., Z. Psaradakis and M. Sola (1999). Detecting periodically collapsing bubbles: a Markov-switching unit root test. *Journal of Applied Econometrics,* Vol. 14(2), pp. 143-154.

Hamilton, J. D and C. H. Whiteman (1985). The Observable Implications of self-fulfilling expectations. *Journal of Monetary Economics,* Vol 16, pp. 353-373.

Hamilton, J.D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica,* Vol. 57, pp. 319-352.

----- (1990). Analysis of Time Series Subject to Regime Changes. *Journal of Econometrics,* Vol. 45, 39-70.

----- (2009), Causes and Consequences of the Oil Shock of 2007-08. *Brookings papers on Economic activity,* 2009, pp. 215-261.

Hansen, B. E. (2011). Threshold autoregression in economics, *Statistics and Its Inference,* Vol. 4(2), pp. 123-127.

Harvey, A. C. (1981). *Time Series Models.* Oxford: Philip Allan and Humanities Press.

Hirshleifer, D (1988). Residual risk, trading costs and commodity futures risk premia. *The Review of Financial Studies,* Vol. 1(2), pp. 173-193.

Hooker, M. A. (2000). Misspecification vs bubbles in hyperinflation data: Monte Carlo and interwar European evidence. *Journal of International Money and Finance,* Vol 19, pp. 583-600.

International Copper Study Group. *'The World Copper Factbook 2014,'* Retrieved from <www.icsg.org/index.php/component/jdownloads/finish/170/1997>.

International Monetary Fund (2014). International Financial Statistics. Mimas, University of Manchester. DOI: <http://dx.doi.org/10.5257/imf/ifs/2014-05>.

Itmar, C. (2014). Rtadf: Testing for Bubbles with Eviews. *MPRA Paper 58791,* University Library of Munich, Germany.

Jagannathan, R (1985). An investigation of Commodity Futures Prices Using the Consumption-Based Intertemporal Capital Asset Pricing Model. *The Journal of Finance,* Vol. 40(1), pp. 175-191.

Jovanovic, B. (2013). Bubbles in prices of Exhaustible Resources. *International Economic Review,* Vol. 54(1), pp. 1-34.

Kaldor, N. (1939). Speculation and economic stability. *The Review of Economic Studies,* Vol 7(1), pp. 1-27.

Kilian, L (2009). Not all price shocks are alike: disentangling demand and supply shocks in the crude market. *American Economic Review,* Vol. 99 (3), pp. 1053-1069.

Kim, Chang-Jin (1993a). Unobserved-Component Time Series Models with Markov-Switching Heteroskedasticity: Changes in Regime and the Link Between Inflation Rates and Inflation Uncertainty. *Journal of Business and Economic Statistics,* Vol 11, pp. 341-349.

Kim, Chang-Jin (1993b). Sources of Monetary Growth Uncertainty and Economic Activity: The Time-Varying-Parameter Model with Heteroskedastic Disturbances. *Reviews of Economics and Statistics,* Vol. 75, pp. 483-492.

Kim, Chang-Jin (1994). Dynamic Linear Models with Markov-Switching. *Journal of Econometrics,* Vol. 60, pp. 1-22.

Kim, C-J and C. R. Nelson (1999). *State-space models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications,* MIT Press.

Kim J.H. (2006). Wild bootstrapping variance ratio tests. *Economics Letters*, Vol. 92, pp. 38-43.
Kim, C-J and C. R. Nelson (1999). *State-space models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications,* MIT Press.

Kindleberger, C. P. (1989). *Manias, panics, and crashes: a history of financial crisis.* Basic Books, New York.

Knight, J. and S.E. Satchell (2011). Some new results for threshold AR(1) models. *Journal of Time Series Econometrics,* Vol (3), pp. 1-42.

Knight, J., S.E. Satchell and N. Srivastava (2014). Steady state distributions for models of locally explosive regimes: Existence and econometric implications. *Economic Modelling,* Vol. 41, pp. 281-288.

Kuch, P. H. (1999). 'Nickel', *US Geological Survey,* Retrieved from <minerals.usgs.gov/minerals/pubs/commodity/nickel/500798.pdf>.

Leroy, S. F and R. D. Porter (1981). The Present-Value Relation: Tests Based on Implied Variance Bounds. *Econometrica,* Vol 49(3), pp. 555-574.

London Metal Exchange (2016), '2016 trading volumes,' < https://www.lme.com/metals/reports/monthly-volumes/annual/2016/>

Lo A.W. and MacKinlay A.C. (1988). Stock market prices do not follow random walk: Evidence from a simple specification test. *The Review of Financial Studies*, Vol. 1, pp. 41-66.

------ (1989). The size and power of the variance ratio test in finite samples: A Monte Carlo investigation. *Journal of Econometrics*, Vol. 40, pp.203-238.

Lyons, R. (2001). *The Microstructure Approach to Foreign Exchange*, MIT Press.

Mallory, I. A. (1990). Conduct Unbecoming: The Collapse of the International Tin Agreement. *American University International Law Review,* Vol. 5(3) pp. 835-892.

Mankiw N. G, D. Romer and M. D. Shapiro (1985). An Unbiased Re-examination of Stock Market Volatility. *The Journal of Finance,* Vol. 40 (3), pp. 677-687.

Martin, A. and J. Ventura (2012). Economic Growth with Bubbles. *American Economic Review,* Vol. 102(6), pp. 3033-3058.

Marsh, T. A. and R. C. Merton (1986). Dividend variability and variance bounds tests for the rationality of stock market prices. *The American Economic Review,* Vol 76(3), pp. 483-498.

Miao, Y., Wu, W. and N. Funke (2011). Reviving the Competitive Storage Model: A Holistic Approach to Food Commodity Prices. *IMF WP/11/64.*

Miranda, M. J. and J. W. Glauber (1993). Estimation of Dynamic Nonlinear Rational Expectations Models of Primary Commodity Markets with Private and Government Stockholding. *The Review of Economics and Statistics,* Vol. 75(3), pp. 463-470.

Miranda, M. J. (1997). Numerical strategies for solving the nonlinear rational expectations commodity market model. *Computational Economics,* Vol 11(1-2), pp. 71-87.

Miranda, M. J. and X. Rui (1999). An empirical reassessment of the commodity storage model. *Unpublished Working Paper, Ohio State University.*

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica,* Vol. 29, pp. 3-22.

Newbery, D and Stiglitz, J (1979). The theory of commodity price stabilization rules: Welfare impacts and Supply responses. *The Economic Journal,* Vol. 89 (356), pp. 799-817.

--- (1981), *The theory of commodity price stabilization.* Oxford University Press.

Ng, S. and F. J. Ruge-Murcia (2000). Explaining the Persistence of Commodity Prices. *Computational Economics,* Vol. 16, pp. 149-171.

Papp, J.F., Bray, E.L., Edelstein, D.L., Fenton, M.D., Guberman, D.E., Hedrick, J.B., Jorgenson, J.D., Kuck, P.H., Shedd, K.B., and Tolcin, A.C. (2008). Factors that influence the price of Al, Cd, Co, Cu, Fe, Ni, Pb, Rare Earth Elements, and Zn. *U.S. Geological Survey Open-File Report 2008-1356.*

Perlin, M. (2014). MS_Regress – The MATALB Package for Markov Regime Switching Models. SSRN: http://ssrn.com/abstract=1714016.

Perron, P and Robert J. Shiller (1985). Testing the random walk hypothesis: Power versus frequency of observation. *Economics Letters,* Vol. 18, pp. 381-386.

Peter, C. B., and J. Yu (2010). Dating the timeline of Financial bubbles during the subprime crisis. *Cowles Foundation for research in Economics Discussion Paper 1770,* pp. 1-41.

Pfaffenzeller, S., Newbold, P. and A. Rayner (2007). A Short Note on Updating the Grilli and Yang Commodity Price Index. *The World Bank Economic Review,* Vol. 21(1), pp. 151-163.

Phillips, P. C. B., S. Shi and J. Yu (2013). Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500, *Cowles Foundation Discussion Papers 1914,* Cowles Foundation for Research in Economics, Yale University.

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

Quandt, R. E. (1972). A New Approach to Estimating Switching Regressions. *Journal of the American Statistical Association,* Vol. 67, pp. 306-310.

Quinn, B.G. (1982). A Note on the Existence of Strictly Stationary Solutions to Bilinear Equations. *Journal of Time Series Analysis,* Vol 3, pp. 249-252.

Ribeiro, D. (2004). Models for the price of a storable commodity. *PhD Thesis, University of Warwick.*

Rios, I. and R. J-B. West (2013). Modeling and Estimating Commodity Prices: Copper Prices. *Technical Report, Universidad de Chile.*

Roberts, H. (1967). *Statistical versus clinical prediction of the stock market.* Unpublished manuscript.

Roberts, M. J. and W. Schlenker (2013). Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate. *American Economic Review,* Vol. 103(6), pp. 2265-2295.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review,* Vol 6(2), pp. 41-49.

---- (1971), Stochastic Speculative Price. *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 68 (2), pp. 335-337.

Schnittker, J. A. (1973). The 1972-73 Food Price Spiral. *Brookings Papers on Economic Activity,* Vol. 1973 (2), pp. 498-507.

Sewel, M. (2011). A History of the Efficient Market Hypothesis. *Research Note: UCL Department of Computer Science.*

Sherfin, H. (2010). Behavioralizing Finance. *Foundations and Trends in Finance,* Vol. 4(1-2), pp. 1-184.

Shiller, R. J.(1981). Do Stock Prices Move too much to be justified by subsequent changes in dividends? *American Economic review,* Vol. 71, pp. 421-436.

------ (1984) Stock Prices and Social Dynamics. *Brooking Papers on Economic Activity,* Vol. 2, pp. 457-498.

------ (2005) *Irrational Exuberance.* Princeton University Press; 2nd edition.

Stein, J (1961). The Simultaneous Determination of Spot and Futures Prices. *American Economic Review,* Vol. 51 (5), pp. 1012-1025.

Summers, L H, (1986). Does the stock market rationally reflect fundamental values? *Journal of Finance,* Vol. 41, pp. 591-601.

Summers, L. H and J. M. Poterba (1988). Mean reversion in stock prices: evidence and implications, *Journal of financial economics,* Vol. 22, pp. 27-59.

Turnovsky, S J (1983). The determination of spot and future prices with storable commodities. *Econometrica,* Vol. 51(5), pp. 1363-1387.

Tirole, J. (1982). On the Possibility of Speculation under Rational Expectations, *Econometrica,* Vol. 50 (5), pp. 1163-1181.

----- (1985). Asset Bubbles and Overlapping Generation. *Econometrica,* Vol. 53(6), pp. 1499-1528.

Tong, H. (1978). On a Threshold Model. *In Patter Recognition and Signal Processing* (ed. C.H. Chen), Sijthoff and Noordhoff, Amsterdam.

University of Michigan (2015). *The Index of Consumer Sentiment.* Retrieved from the survey of consumers University of Michigan database. https://data.sca.isr.umich.edu/tables.php.

U.S. Geological Survey. Mineral Commodity Summaries, January 2015. Retrieved from <minerals.usgs.gov/minerals/pubs/commodity/copper/mcs-2015-coppe.pdf>.

Van Nordern S. and H. Schaller (2002) Fads or Bubbles? *Empirical Economics,* Vol 27, pp. 335-362.

West, K. D. (1987). A specification test for speculative bubbles. *Quarterly Journal of Economics,* Vol 102, pp. 553-580.

West, K. D. (1988). Dividend Innovations and Stock Price Volatility. *Econometrica,* Vol 56 (1), pp. 37-61.

Wright, B. D. (2011), The Economics of Grain Price Volatility. *Applied Economic Perspectives and Policy,* Vol. 33(1), pp.32-58.

Williams, J. C and B. D. Wright (1991). *Storage and Commodity Markets*. Cambridge University Press, 1st ed.