ROBUST NUMERICAL BAYESIAN UNIT ROOT TEST FOR MODEL UNCERTAINTY

by

XUEDONG WU

(Under the Direction of Jeffrey H. Dorfman)

ABSTRACT

Unit root testing is an important procedure when performing time series analysis, since all the succeeding inferences should be performed based on a stationary series. Thus, it is crucial to test the stationarity of the time series in hand accurately and efficiently. One issue among the existing popular unit root testing methods is that they all require certain assumptions about the model specification, whether on the functional form or the stochastic term distribution; then all the analyses are performed based on the pre-determined model. However, various circumstances such as data incompleteness, variable selection, and distribution misspecification which may lead to an inappropriate model specification that will produce an erroneous conclusion since the test result depends on the particular model considered. This dissertation focuses on confronting this issue by proposing a new numerical Bayesian unit root test incorporating model averaging which can take model uncertainty as well as variable transformation into account.

The first chapter introduces a broad literature review of all the building blocks need for the development of the new methods, including traditional frequentist unit root tests, Bayesian unit root tests, and Bayesian model averaging. Following chapter II

elaborates the mathematical derivation of the proposed methods, Monte Carlo simulation study and results, as well as testing conclusions on the benchmark Nelson and Plosser (1982) macroeconomic time series. Chapter III applies the method to investigate the effect of data frequency on unit root test results particularly for financial data. We perform our proposed Bayesian Model Averaging method to five commodity futures price series by averaging GARCH and ARCH models with different mean functions and distribution specifications to demonstrate the robustness and usefulness of our method especially when model specification uncertainty issue is presented. Overall, the proposed numerical Bayesian unit root test is a general approach to considering model uncertainty when performing the stationary tests, and it provides an alternative to researchers who are concerned this is a significant issue in their research.

INDEX WORDS: Robust numerical Bayesian unit root test, Model averaging, Model uncertainty, Data frequency

ROBUST NUMERICAL BAYESIAN UNIT ROOT TEST FOR MODEL UNCERTAINTY

by

XUEDONG WU

B.Eng., The University of Science and Technology Beijing, China, 2008

B.S., The University of Georgia, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

ROBUST NUMERICAL BAYESIAN UNIT ROOT TEST FOR MODEL UNCERTAINTY

by

XUEDONG WU

Major Professor:	Jeffrey H. Dorfman
Committee:	Gregory J. Colson
	Berna Karali

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
May 2015

DEDICATION

I dedicate this dissertation to my loving wife, Ting Meng, for her great love, selfless support and encouragement during all these years we have been through together; and to my dear parents, who have been giving me the courage and confidence to keep exploring the new world ever since my childhood.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Jeffrey Dorfman, for his excellent and valuable guidance, considerate caring and patience on a day to day basis from start to finish. You not only offered me the professional advises and mentorship, but also set an excellent example that I want to make myself to be: a scholar with extensive knowledge, wisdom, humor, humility, passion, and thoughtfulness. Thank you for being the best mentor one can ask for.

I would also like to thank my committee members, Dr. Gregory Colson and Dr. Berna Karali, without whose help this work cannot be smoothly finished. Dr. Coslon is such a kind professor who is always ready to help others, and thanks for your help on the paper we work together. You demonstrated to me how a naive idea gradually developed into a rigorous academic paper which brought tremendous benefits to my later research. Dr. Karali has been offered valuable assistance to me ever since I took your class, and your nimble mind, serious scientific attitude, rigorous scholarship deeply inspired me to keep improving my work, as well as my attitudes toward life.

I also want to show my appreciation to Dr. Michael Wetzstein and Dr. Octavio Ramirez, who offered me this precious chance pursuing my PhD in this great department as well as the graduate assistantship. It is you that gave me a hand when I tried to move my career to an economist and supported me to focus on the academic research.

Time is fleeting. The life in UGA is a priceless treasure to me forever, and I want to thank all my friends, especially to the ones in the same class: Hui Xian, Nanying

Wang, Xiaosi Yang, Anne Mandich, Mingrong Song, Xiaofei Li, Shiyu Ye, Yi Chen and Padmanand Madhavan Nambiar. Thank you all for making my life in Georgia enjoyable. I will always remember the precious days of our lives: the discussion and debate that generated great ideas, the support and encouragement we gave to each other, and the happiness and joy from heart we shared. It is my luck to have you accompany all these years, and this pure, crystal-clear friendship will always be an inspiration and purification of my mind.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

LITERATURE REVIEW

1.1 Introduction

Nonstationarity is one of the important issues in empirical research of economics and finance since time series usually display trending behavior or nonstationarity in the mean. Series like asset prices, exchange rates and the original levels of macroeconomic aggregates (GDP, for instance) are examples of these kind of data series. It is very important to detect these two kinds of trend in the data in order to adopt an appropriate strategy for accurate data analysis and model estimation. Unit root testing is the approach to achieve this goal and decide whether the data series is nonstationary. Thus, the test is of great importance in empirical research.

This dissertation develops a robust numerical Bayesian unit root testing method which can take model specification uncertainty and variable transformation into account. This chapter will detail every element and technique that will be used to derive this method. It starts with an introduction of the mathematic background and development of frequentist unit root tests. Although the method discussed belongs to a Bayesian framework, it is necessary to better understand the discussion and comparison in the later chapters if commonly used frequentist methods are introduced first. Then a general review of the development of Bayesian unit root tests follows, with some content about the advantages of this approach. Detailed mathematical part is saved for Chapter II when the main method developed in this dissertation is exploited. The last past is a revise of the

history of Bayesian model averaging, another significant building block of the methods derived in this paper. Since this technique now is widely adopted in many fields in economics, several papers that employed Bayesian model averaged will be listed in different sub-fields of economics and their major conclusions will be introduced. Again, the formulation and mathematical detailed will be elaborated in Chapter 2.

1.2 Review of Frequentist Unit Root Tests

The fundamental paper that elaborated perhaps the most popular method of univariate unit root testing in empirical research is Dickey and Fuller (1979) . Consider OLS estimation of an AR(1) model:

$$y_t = \rho y_{t-1} + \varepsilon_t \tag{1}$$

where $\varepsilon_t \sim i.i.d.N(0,\sigma^2)$ and $y_0 = 0$. The interested hypothesis test is:

$$\begin{aligned} H_0 &: \rho = 1 \\ H_a &: |\rho| < 1 \end{aligned} \tag{2}$$

The estimation of ρ is expressed by (Hamilton 1994, 2:475):

$$\hat{\rho}_T = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} \tag{3}$$

Under the alternative hypothesis, according to the asymptotical theory:

$$\sqrt{T}(\hat{\rho}_T - \rho) \xrightarrow{L} N(0,(1-\rho^2)) . \tag{4}$$

This does not hold under the null hypothesis, under which the distribution collapses to a point mass at zero. A larger normalizing constant is needed to obtain a nondegenerate

asymptotic distribution for $\widehat{\rho}_T$ in the nonstationary case. Instead of normal, the aforementioned distribution will also be more complicated and involves standard Brownian motion. Based on Hamilton (1994), the DF test can be categorized into the following situations:

Case 1. No constant term or time trend in the regression; true process is a random walk.

If both the true and estimated model are following the form (1) above, the OLS estimate of $\widehat{\rho}_T$ is characterized by:

$$T(\widehat{\rho}_T - 1) \xrightarrow{L} \frac{1/2\{[W(1)]^2 - 1\}}{\int_0^1 [W(r)]^2 dr} \tag{5}$$

in which $W(\cdot)$ is a Standard Brownian motion.

Case 2. Constant term but no time trend included in the regression; true process is a random walk.

The data generating process is assumed to remain of the form (1), while the estimated OLS model includes a constant term:

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \tag{6}$$

In this case the asymptotical distribution would have the form:

$$T(\widehat{\rho}_T - 1) \xrightarrow{L} \frac{1/2\{[W(1)]^2 - 1\} - W(1) \cdot \int W(r) dr}{\int [W(r)]^2 dr - [\int W(r)\, dr]^2} \tag{7}$$

Case 3. Constant term but no time trend included in the regression; true process is random walk with drift.

In this case, the true process is assumed to be:

$$y_t = \alpha + y_{t-1} + \varepsilon_t \tag{8}$$

while the estimated model has the same form. The distribution will be:

$$\begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T^{3/2}(\hat{\rho}_T - 1) \end{bmatrix} \underset{\rightarrow}{L} \mathrm{N}(\mathbf{0}, \sigma^2 \mathbf{Q}^{-1}) . \tag{9}$$

So in this case, the estimated coefficients follow a regular Gaussian distribution, thus the related hypothesis tests can be performed using standard OLS T or F tests.

Case 4. Constant term and time trend included in the regression; true process is random walk with or without drift.

The true model has the same form as (8) this time, while the estimated model includes a time trend:

$$y_t = \alpha + \rho y_{t-1} + \delta t + \varepsilon_t \tag{10}$$

The inclusion of time trend will cause collinear problem in larger samples when $\alpha \neq 0$, thus the model need to be rearranged to the following form:

$$\begin{aligned} y_t &= (1-\rho)\alpha + \rho[\mathrm{y}_{t-1} - \alpha(\mathrm{t}-1)] + (\delta + \rho\alpha)\,\mathrm{t} + \varepsilon_t \\ &\equiv \alpha^* + \rho^* \xi_{t-1} + \delta^* t + \varepsilon_t \end{aligned} . \tag{11}$$

Under this parameterization, the asymptotical distribution of the estimated coefficients would be:

$$\begin{bmatrix} T^{1/2}\alpha_T^* \\ T(\hat{\rho}_T^* - 1) \\ T^{3/2}(\hat{\delta}_T^* - \alpha_0) \end{bmatrix} \xrightarrow{L} \begin{bmatrix} \sigma & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sigma \end{bmatrix} \begin{bmatrix} 1 & \int W(\mathrm{r})\,\mathrm{dr} & \frac{1}{2} \\ \int W(\mathrm{r})\,\mathrm{dr} & \int [W(\mathrm{r})]^2\,\mathrm{dr} & \int rW(\mathrm{r})\,\mathrm{dr} \\ \frac{1}{2} & \int rW(\mathrm{r})\,\mathrm{dr} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} W(1) \\ \frac{1}{2}\{[\mathrm{W}(1)]^2 - 1\} \\ W(1) - \int W(\mathrm{r})\,\mathrm{dr} \end{bmatrix} \quad (12)$$

The above described the theory of unit root test for a first order autoregression. However, more complicated dynamic structures are required in empirical research work than only AR(1). Two different classes of unit root tests were proposed to consider higher order serial correlation. Said and Dickey (1984) generalized the basic autoregressive unit root test to accommodate general ARMA (p, q) as well as unknown order, which is referred to as the augmented Dickey-Fuller (ADF) test. The null hypothesis of the ADF test is a time series is I (1) against the alternative that it is I (0). By introducing higher-order autoregressive terms in the regression, ADF tests controls for serial correlation. More details about this test could be found in Dickey and Fuller (1979), Said and Dickey (1984), Hamilton (1994, 2:516).

One of the disadvantages of the ADF test is that a researcher has to determine the lag length p before performing the test. A small p will cause the remaining serial correlation in the error to bias the test; too large p will affect the power of the test (Zivot and Wang 2007). Some variable selection criterions like AIC of BIC can be adopted to help choose the lag length. Ng and Perron (1995) proposed another data dependent rule for lag length selection which can result in stable size of the test and minimal power loss. Our method proposed later in this dissertation can mitigate this problem without having to choose a single lag length.

Another famous unit root test was proposed by P. C. B. Phillips and Perron (1988), the PP test, which takes serial correlation of the differenced data into account by adjusting the statistics calculated from a simple first order autoregression. Under the null hypothesis, the PP test statistics have the same asymptotic distribution as the ADF t-statistic. The advantages of PP test include that the researchers do not need to specify the lag length used in the autoregressive model; meanwhile, it is also robust to general forms of heteroskedasticity in the error term (Zivot and Wang 2007).

Beside the aforementioned two classes of popular unit root tests, tremendous literatures exist on many other unit root tests methods derived from different aspects of underlying theory as well as data characteristics. Cochrane (1988) proposed a variance ratio unit root test based on the fact that if a series is stationary, or stationary with a deterministic time trend, the innovation should have no permanent effect on the series. Kwiatkowski, et al. (1992) suggested a test of the null hypothesis that an observable series is stationary around a deterministic trend, which is based on the LM test of the hypothesis that the random walk part of the series has zero variance. Stock (1994) provides a nice summary and survey to lots of major methods of unit root tests and related problems which is a great source to refer to. Recent developments about frequentist unit root test focus on deeper and more detailed problems of the existing methods, such as power and size problem under difference occasions (Schwert 1989; Perron and Ng 1996; Lopez 1997), selection of optimal truncation lags (Ng and Perron 1995; Lopez 1997; Ng and Perron 2001), and so on other topics, which will be skipped in this review.

1.3 Review of Bayesian Unit Root Test

The previous section gave a brief review of unit root test under classical statistical approach and provided details of the most popular methods. The key idea is calculating the distribution of $\hat{\rho}$ conditional on a particular assumed true value such as ρ=1 (Hamilton 1994). However, this classical distribution of $\hat{\rho}$ can be strongly skewed and does not follow standard distributions even in large samples. Sims (1988) and Sims and Uhlig (1991) discuss this problem in detail and pushed a new wave of research and application of Bayesian unit root tests. According to Sims and Uhlig (1991), compared to the classical test of an unit root with the null ρ=1 based only on the distribution of $\hat{\rho}$, the Bayesian unit root test focused on the distribution of $\hat{\rho}|\rho$ for all the possible values of ρ, together with the information provided on the prior probability of ρ. The key idea is that if the distribution of $\hat{\rho}|\rho$ had the same skew and dispersion for every ρ as it does at ρ=1, then we would conclude having observed any particular $\hat{\rho}$ the true value of ρ has a larger probability to be large (Hamilton 1994).

Since then, extensive literature has been devoted into the analysis of Bayesian unit root tests from various aspects. One of the directions is specifying the priors. P. C. Phillips (1991) noted that the adoption of flat priors in the previous two papers unwittingly biases inferences towards stationarity, and suggested using a Jeffreys prior instead, although this would not be appropriate if the prior distribution is intended to present the actual information available to the analysis before seeing the data (Hamilton 1994). Schotman and Van Dijk (1991) fully explained that although a Jeffreys prior has an advantage in the analysis of autoregressive time-series models since it is invariant with

to the parameterization, researchers do pay a price for the invariance because it depends on the sample size, the data and the complete structure of the model. Meanwhile, they concluded that improper priors like the uniform and the Jeffreys prior are less suited for Bayesian inference on a sharp null hypothesis like a unit root. Schotman (1994) shown that the posterior of the autoregressive parameter can be very sensitive to the degree of prior dependence between the unconditional mean and the autocorrelation parameter through the analysis of an AR(1) model. The weaker the prior dependence between these two parameters, the more the posterior of the autoregressive parameter will be shifted toward the unit root. Recently, Griffiths (2012) investigated the effects of different choices of priors on Bayesian unit root test results. Four types of priors are considered: diffuse (unifrom), Jeffreys, Lubrano (1995), and Berger and Yang (1994). They concluded that both testing procedures (posterior odds or posterior credible intervals) as well as the choice of prior have a significant impact on performances; meanwhile, under their Monte Carlo settings and model specification, the prior which led to the best performance in terms of test size and power was the Lubrano prior.

The previous research all focused on designing the priors on the dominant root and the other lag coefficients. An alternative and innovated method is specifying the priors directly on all the roots of the underlying dynamic process. Dorfman (1993) proposed this method, and by augmenting the AR process to state space form, he directly specified prior distributions on the eigenvalues of the coefficient matrix of the state vector for the dynamic system. This technique makes the connections between priors and underlying dynamic system explicit and provides straightforward understanding of the priors. So in the present paper, this mode is adopted as the fundamental idea on which the

new method will be built. The detailed description and mathematical derivations will be elaborated in Chapter II.

Besides the discussion of prior specifications, another direction of Bayesian unit root test literature focused on the model specification, including deciding the number of autoregressive lags, inclusion of deterministic trend and drift, as well as specifying the distributional assumption of the stochastic component. These considerations are not new since they are widely discussed under the classical unit root test regime. Different specifications will directly affect the asymptotically distribution of the dominant root in classical unit root test theory, as well as the power and size of the test. It is of the same importance under the Bayesian framework. Phillips and Schmidt (1989) pointed out that the Dickey-Fuller tests are not well designed for testing trend reversion in the presence of a deterministic trend. They proposed methods, although not Bayesian, to handle this problem, which are essentially LM tests. DeJong and Whiteman (1991) reconsidered the macroeconomic time series from Nelson and Plosser (1982). Compared to the results that most (13 out of 14) series are consistent with the random walk hypothesis in the latter paper, DeJong and Whiteman adopted a likelihood principle approach and explicit Bayesian methods to investigate the problem and indicated that for most series the trend-stationarity hypothesis is much more likely. Distribution of the stochastic component, generally speaking, is also related to model specification. Although in asymptotical theory it is not required of the stochastic term to have any specific distribution as long as it satisfies random walk conditions, certain ad hoc density functions do have significant effects on the empirical results of unit root tests. Furthermore, there is a branch of classical unit root test theory based on the likelihood function (Dickey and Fuller 1981;

Dufour and King 1991; Elliott, Rothenberg, and Stock 1992) that requires certain assumption about the stochastic term distribution. Of course, the Bayesian framework depends on the posterior distribution that is derived from likelihood function and prior distributions, so the distributional assumptions are central to Bayesian approaches. Since this part is closely related to the Chapter III of this thesis, the related review will be introduced in that chapter.

So far a brief review of Bayesian unit root test methods has been given and several problems and research directions summarized, the detailed theory and mathematical derivation will be introduced in Chapter II.

1.4 Review of Bayesian Model Averaging

The regular modeling approach in economic empirical research work is that first researchers select a single model form from possible model space which is assumed to be the “true model” that “generates” the data. Then all the analyses are performed based on this pre-determined model and conclusions will also be drawn from it. This typical practice is concise and effective when the chosen model is somewhat close to the real underlying “data generating processes”. However, due to various circumstances such as data incompleteness, variable selection, distribution misspecification, etc., this procedure may lead to inappropriate model assumption which will produce erroneous conclusions since estimates may well depend on the particular model considered (Moral-Benito 2013). For example, Leamer (1983) wrote a quite interesting paper to point out the problem of regression modeling analysis depending on arbitrary decisions about the choice of independent variables.

Researchers have developed ways to confront this problem, and model averaging is one of them which is popular in many fields. The basic idea is to estimate all the possible models from a certain model space and assign weight to each model. Inferences can be made from this weighted average model and by doing so researchers actually considered not only the uncertainty to the variables conditional on a given model, but also the uncertainty associated to the parameter estimates across different models (Moral-Benito 2013). This approach generally includes two different directions in the literature: frequentist model averaging and Bayesian model averaging. The method proposed in this research is under the Bayesian framework so the following reviews as well as theory introduction will all be related to the Bayesian model averaging. The research paper by Moral-Benito (2013) provided one chapter of good review on frequentist model averaging.

Historically, the early mention of the idea of model combination could trace to Barnard (1963) about quality control in airline industry. Poirier (1991) is one of the key early paper which applied this technique to investigate the empirical evidence on the effects of unanticipated changes in nominal money on real output in 47 countries. Clemen (1989) provided a detailed review regarding the combination of models in forecasting field. There are several good reviews about Bayesian model averaging from different disciplines and perspectives, including data mining and statistical inference (Chatfield 1995), social research (Raftery 1995). Hoeting et al. (1999) provided a great tutorial on Bayesian model averaging focused on implementation and practical matters. Recently, Clyde and George (2004) highlighted some of the new applications and methods related to Bayesian model averaging including tree models, graphical models, etc.

It was a long time before Bayesian Model Averaging technique was widely used in different sub-fields of economics. The literature before was either focused on statistical method development, or solely on applications about forecasting. In recent decades, due to new methods and the availability of more powerful computing resources, BMA is spreading to almost all sub-fields of economic research works (Moral-Benito 2013). For instance, in policy analysis, Brock, Durlauf, and West (2007) explored ways to integrate model uncertainty into policy evaluation and provided an excellent description of both theory and empirics. Besides introducing the theory of model averaging, he also compared BMA with traditional frequentist model averaging. The empirical application they considered is in the field of monetary policy. By taking 25,000 models differentiated in lags of interest rates, inflation, and output gap in the IS and Philips curve equations, they assess simple monetary policy rules for some standard New Kaynesian specifications.

There are also many other applications of model averaging in the macroeconomic field. Min and Zellner (1993) analyzed forecasts of output growth rates for eighteen countries from 1974-1987 using Bayesian Model averaging. They also reported the comparison of performance in forecasting one-year ahead output growth rates of these eighteen countries between Bayesian pooling techniques and non-Bayesian forecast combining techniques. Koop and Potter (2004) considered the problem of forecasting in dynamic factor models using Bayesian model averaging. They applied the proposed methods to the topic of forecasting GDP and inflation using quarterly U.S. data on 162 time series, and they claimed that relative to the small forecasting gains provided by including more factors, the gains provided by using Bayesian model averaging over

forecasting methods based on a singly model are appreciable. More recently, Eicher, Henn, and Papageorgiou (2012) argued that the effect of preferential trade agreements (PTAS) on trade flows is subject to major model uncertainty problems which caused the existing empirical literature to produce remarkably disparate results. They performed Bayesian model averaging techniques on this topic and found strong evidence of trade creation, trade diversion and open bloc effects.

Another important field in economics where model averaging research is active is economic growth. The influential work by Sala, Martin and Miller (2004) analyzed long run economic growth by a Bayesian averaging of classical estimates approach which averaged OLS coefficients across models. Fernandez, Ley, and Steel (2001) investigated the issue of model uncertainty in cross country growth and proposed the advantage of using a Bayesian model averaging approach over choosing a single model in such kind of topics.

Besides the aforementioned areas that BMA has actively involved, there are also many others where BMA techniques can be found as tools of analysis. For example, Koske and Wanner (2013) used Bayesian model averaging techniques to study the determinants of labor income inequality among OECD countries and suggested that some specific labor tax policies and education have important effects on labor income. In finance, Cremers (2002) investigated the predictability of excess stock returns using Bayesian model averaging and claims that the averaged model results provided superior performance to individual models. All these studies indicate that Bayesian model averaging techniques have been widely used in fields of economics and play an important role as an econometric approach in empirical research work. One interested in more

thorough review can refer to Moral-Benito (2013) which also provides a detailed review on frequentist model averaging techniques.

1.5 Conclusion

So far a brief review has been given on major blocks that will be needed to build the robust numerical Bayesian unit root test for model uncertainty in next chapters, including the famous (augmented) Dickey-Fuller unit root test as well as other methods that belong to the frequentist unit root test regime; the development of Bayesian unit root test methods and the main debated directions in the literature which is related to our method in next chapters; at the end the history and broad applications of Bayesian model averaging techniques are introduced. All these reviews concentrate on ideas and development of the corresponding theory, and the detailed mathematical derivation will be discussed in the Chapter II and Chapter III where these techniques are adopted. Next in Chapter II, the new proposed numerical Bayesian unit root test for model uncertainty will be elaborated with simulation results as well as an application to the famous benchmark data set from Nelson and Plosser (1982). In Chapter III, we propose to test the effect of data frequency on the results of unit root tests, specifically in the financial research field and apply our proposed BMA method to five commodity futures price series. The results demonstrate the robustness and usefulness of our method when model specification uncertainty issue is present.

References

Avramov, Doron. 2002. “Stock Return Predictability and Model Uncertainty.” *Journal of Financial Economics* 64 (3): 423–58.

Barnard, G. A. 1963. “New Methods of Quality Control.” *Journal of the Royal Statistical Society. Series A (General)* 126 (2): 255–58. doi:10.2307/2982365.

Bauwens, Luc, Michel Lubrano, and Jean-Francois Richard. 2000a. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press. http://books.google.com/books?hl=en&lr=&id=hRV3XiQVWLUC&oi=fnd&pg=PA1&dq=Bayesian+inference+in+dynamic+econometric+models&ots=mJMR-mN8hn&sig=Yq1euDu-iE8DAzNeoBWaA0DIsAA.

———. 2000b. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press.

Berger, James O., and Ruo-yong Yang. 1994. “Noninformative Priors and Bayesian Testing for the AR(1) Model.” *Econometric Theory* 10 (3/4): 461–82.

Bhargava, Alok. 1986. “On the Theory of Testing for Unit Roots in Observed Time Series.” *Review of Economic Studies* 53 (3): 369.

Bollerslev, Tim. 1986. “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics* 31 (3): 307–27.

Boswijk, H. Peter. 2001. *Testing for a Unit Root with near-Integrated Volatility*. Tinbergen Institute Discussion Paper. http://www.econstor.eu/handle/10419/85890.

Boswijk, H Peter, and Franc Klaassen. 2012. “Why Frequency Matters for Unit Root Testing in Financial Time Series.” *Journal of Business & Economic Statistics* 30 (3): 351–57.

Box, George EP, and David R Cox. 1964. “An Analysis of Transformations.” *Journal of the Royal Statistical Society. Series B (Methodological)*, 211–52.

Brock, William A., Steven N. Durlauf, and Kenneth D. West. 2007. “Model Uncertainty and Policy Evaluation: Some Theory and Empirics.” *Journal of Econometrics* 136 (2): 629–64.

Campbell, John Y., and Pierre Perron. 1991. “Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots.” In *NBER Macroeconomics Annual 1991, Volume 6*, 141–220. MIT Press. http://www.nber.org/chapters/c10983.pdf.

Chambers, Marcus J. 3. “Testing for Unit Roots with Flow Data and Varying Sampling Frequency.” *Journal of Econometrics* 119 (1): 1–18. doi:10.1016/S0304-4076(03)00152-0.

Charitidou, Efstratia, Dimitris Fouskakis, and Ioannis Ntzoufras. 2013. “Bayesian Transformation Family Selection: Moving towards a Transformed Gaussian Universe.” *arXiv Preprint arXiv:1312.3482*. http://arxiv.org/abs/1312.3482.

Chatfield, Chris. 1995. “Model Uncertainty, Data Mining and Statistical Inference.” *Jorunal of the Royal Statistical Society* 158 (3): 419–66.

Choi, In. 1992. “Effects of Data Aggregation on the Power of Tests for a Unit Root: A Simulation Study.” *Economics Letters* 40 (4): 397–401.

Choi, In, and Bhum Suk Chung. 1995. “Sampling Frequency and the Power of Tests for a Unit Root: A Simulation Study.” *Economics Letters* 49 (2): 131–36.

Claeskens, Gerda, and Nils Lid Hjort. 2008. *Model Selection and Model Averaging*. 1 edition. Cambridge ; New York: Cambridge University Press.

Clemen, Robert T. 1989. “Combining Forecasts: A Review and Annotated Bibliography.” *International Journal of Forecasting* 5 (4): 559–83. doi:10.1016/0169-2070(89)90012-5.

Clyde, Merlise, and Edward I. George. 2004. “Model Uncertainty.” *Statistical Science* 19 (1): 81–94.

Cochrane, John H. 1988. “How Big Is the Random Walk in GNP?” *Journal of Political Economy* 96 (5): 893–920.

Cremers, K. J. Martijn. 2002. “Stock Return Predictability: A Bayesian Model Selection Perspective.” *Review of Financial Studies* 15 (4): 1223–49. doi:10.1093/rfs/15.4.1223.

DeJong, David N., John C. Nankervis, N. E. Savin, and Charles H. Whiteman. 1992. “The Power Problems of Unit Root Test in Time Series with Autoregressive Errors.” *Journal of Econometrics* 53 (1–3): 323–43. doi:10.1016/0304-4076(92)90090-E.

DeJong, David N., and Charles H. Whiteman. 1991. “Reconsidering ‘Trends and Random Walks in Macroeconomic Time Series.’” *Journal of Monetary Economics* 28 (2): 221–54.

Dickey, David A., and Wayne A. Fuller. 1979a. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association* 74 (366a): 427–31. doi:10.1080/01621459.1979.10482531.

———. 1979b. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association* 74 (366a): 427–31. doi:10.1080/01621459.1979.10482531.

———. 1981. “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root.” *Econometrica* 49 (4): 1057–72. doi:10.2307/1912517.

Dorfman, Jeffrey H. 1993. “Bayesian Efficiency Tests for Commodity Futures Markets.” *American Journal of Agricultural Economics* 75 (5): 1206–10.

Dorfman, Jeffrey H., and Myung D. Park. 2011. “Estimating the Risk-Return Tradeoff in Agribusiness Stocks: Linkages with the Broader Stock Market.” *American Journal of Agricultural Economics* 93 (2): 426–33. doi:10.1093/ajae/aaq158.

Dufour, Jean-Marie, and Maxwell L. King. 1991. “Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary AR (1) Errors.” *Journal of Econometrics* 47 (1): 115–43.

Eicher, Theo S., Christian Henn, and Chris Papageorgiou. 2012. “Trade Creation and Diversion Revisited: Accounting for Model Uncertainty and Natural Trading Partner Effects.” *Journal of Applied Econometrics* 27 (2): 296–321. doi:10.1002/jae.1198.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. 1992. *Efficient Tests for an Autoregressive Unit Root*. National Bureau of Economic Research Cambridge, Mass., USA. http://www.nber.org/papers/t0130.

Engle, Robert F. 1982. “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.” *Econometrica* 50 (4): 987–1007. doi:10.2307/1912773.

Engle, Robert F., and Victor K. Ng. 1993. “Measuring and Testing the Impact of News on Volatility.” *The Journal of Finance* 48 (5): 1749–78. doi:10.1111/j.1540-6261.1993.tb05127.x.

Fan, Jianqing, and Runze Li. 2001. “Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties.” *Journal of the American Statistical Association* 96 (456): 1348–60.

Fernandez, Carmen, Eduardo Ley, and Mark FJ Steel. 2001. “Model Uncertainty in Cross-Country Growth Regressions.” *Journal of Applied Econometrics* 16 (5): 563–76.

Fernández-Villaverde, Jesús, and Juan Francisco Rubio-Ramírez. 2004. “Comparing Dynamic Equilibrium Models to Data: A Bayesian Approach.” *Journal of Econometrics* 123 (1): 153–87. doi:10.1016/j.jeconom.2003.10.031.

Gelfand, Alan E., and Adrian FM Smith. 1990. “Sampling-Based Approaches to Calculating Marginal Densities.” *Journal of the American Statistical Association* 85 (410): 398–409.

Geman, Stuart, and Donald Geman. 1984. “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images.” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6: 721–41.

Geweke, John. 1992. “Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments.” In *Bayesian Statistics 4*, edited by Jose M.

Bernardo, James O. Berger, A.P. dawid, and J.F.M Smith, 169–93. Oxford University Press.

Griffiths, Charley Xia. 2012. *Bayesian Unit Root Testing: The Effect of Choice of Prior On Test Outcomes*. The University of Melbourne.

Hamilton, James Douglas. 1994. *Time Series Analysis*. Vol. 2. Princeton university press Princeton.

Hansen, Peter R., and Asger Lunde. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics* 20 (7): 873–89. doi:10.1002/jae.800.

Harris, R. I. D. 1992. "Testing for Unit Roots Using the Augmented Dickey-Fuller Test: Some Issues Relating to the Size, Power and the Lag Structure of the Test." *Economics Letters* 38 (4): 381–86. doi:10.1016/0165-1765(92)90022-Q.

Harvey, Andrew C, and AC Harvey. 1993. *Time Series Models*. Vol. 2. Harvester Wheatsheaf New York.

Hastings, W. K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57 (1): 97–109. doi:10.2307/2334940.

Hoeting, Jennifer A., David Madigan, Adrian E. Raftery, and Chris T. Volinsky. 1999. "Bayesian Model Averaging: A Tutorial." *Statistical Science* 14 (4): 382–401.

Jing, Liang. 2010. "Hastings-within-Gibbs Algorithm: Introduction and Application on Hierarchical Model." University of Texas at San Antonio. http://georglsm.r-forge.r-project.org/site-projects/pdf/Hastings_within_Gibbs.pdf.

Koop, Gary, Eduardo Ley, Jacek Osiewalski, and Mark FJ Steel. 1997. “Bayesian Analysis of Long Memory and Persistence Using ARFIMA Models.” *Journal of Econometrics* 76 (1): 149–69.

Koop, Gary, and Simon Potter. 2004a. “Forecasting in Dynamic Factor Models Using Bayesian Model Averaging.” *The Econometrics Journal* 7 (2): 550–65.

———. 2004b. “Forecasting in Dynamic Factor Models Using Bayesian Model Averaging.” *The Econometrics Journal* 7 (2): 550–65.

Koske, I., and I. Wanner. 2013. “The Drivers of Labour Income Inequality – an Analysis Based on Bayesian Model Averaging.” *Applied Economics Letters* 20 (2): 123–26. doi:10.1080/13504851.2012.683164.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin. 1992. “Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” *Journal of Econometrics* 54 (1): 159–78.

Leamer, Edward E. 1983. “Let’s Take the Con Out of Econometrics.” *The American Economic Review* 73 (1): 31–43.

Lee, Peter M. 2012. *Bayesian Statistics: An Introduction*. John Wiley & Sons.

Ling, Shiqing, and WK Li. 2003. “Asymptotic Inference for Unit Root Processes with GARCH (1, 1) Errors.” *Econometric Theory* 19 (04): 541–64.

Lopez, J. Humberto. 1997. “The Power of the ADF Test.” *Economics Letters* 57 (1): 5–10. doi:10.1016/S0165-1765(97)81872-1.

Lubrano, Michel. 1995. “Testing for Unit Roots in a Bayesian Framework.” *Journal of Econometrics* 69 (1): 81–109.

Lucas, André. 1995. “Unit Root Tests Based on M Estimators.” *Econometric Theory* 11 (02): 331–46.

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. “Equation of State Calculations by Fast Computing Machines.” *The Journal of Chemical Physics* 21 (6): 1087–92.

Min, Chung-ki, and Arnold Zellner. 1993. “Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates.” *Journal of Econometrics* 56 (1–2): 89–118. doi:10.1016/0304-4076(93)90102-B.

Mitchell, James, Nigel Pain, and Rebecca Riley. 2011. “The Drivers of International Migration to the UK: A Panel-Based Bayesian Model Averaging Approach*.” *The Economic Journal* 121 (557): 1398–1444.

Moral-Benito, Enrique. 2013. “Model Averaging in Economics: An Overview.” *Journal of Economic Surveys*.

Morales, Knashawn H., Joseph G. Ibrahim, Chien-Jen Chen, and Louise M. Ryan. 2006. “Bayesian Model Averaging with Applications to Benchmark Dose Estimation for Arsenic in Drinking Water.” *Journal of the American Statistical Association* 101 (473): 9–17.

Nelson, Charles R, and Charles R Plosser. 1982. “Trends and Random Walks in Macroeconmic Time Series: Some Evidence and Implications.” *Journal of Monetary Economics* 10 (2): 139–62.

Nelson, Daniel B. 1991. “Conditional Heteroskedasticity in Asset Returns: A New Approach.” *Econometrica: Journal of the Econometric Society*, 347–70.

Ng, Serena, and Pierre Perron. 1995. “Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag.” *Journal of the American Statistical Association* 90 (429): 268–81. doi:10.2307/2291151.

———. 2001. “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power.” *Econometrica* 69 (6): 1519–54.

Pagano, Marcello. 1974. “Estimation of Models of Autoregressive Signal plus White Noise.” *The Annals of Statistics*, 99–108.

Perron, Pierre, and Serena Ng. 1996. “Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties.” *The Review of Economic Studies* 63 (3): 435–63.

Phillips, Peter CB. 1991. “To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends.” *Journal of Applied Econometrics* 6 (4): 333–64.

Phillips, Peter C. B., and Pierre Perron. 1988. “Testing for a Unit Root in Time Series Regression.” *Biometrika* 75 (2): 335–46. doi:10.2307/2336182.

Phillips, Peter C. B., and Peter Schmidt. 1989. *Testing for a Unit Root in the Presence of Deterministic Trends*. Cowles Foundation Discussion Paper 933. Cowles Foundation for Research in Economics, Yale University. https://ideas.repec.org/p/cwl/cwldpp/933.html.

Phillips, Peter CB, and Peter Schmidt. 1989. *Testing for a Unit Root in the Presence of Deterministic Trends*. Cowles Foundation for Research in Economics, Yale University. http://ideas.repec.org/p/cwl/cwldpp/933.html.

Poirier, Dale J. 1991. “A Bayesian View of Nominal Money and Real Output through a New Classical Macroeconomic Window.” *Journal of Business & Economic Statistics* 9 (2): 125–48. doi:10.2307/1391778.

Raftery, Adrian E. 1995. “Bayesian Model Selection in Social Research.” *Sociological Methodology* 25: 111–64.

Roberts, G. O., and A. F. M. Smith. 1994. “Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms.” *Stochastic Processes and Their Applications* 49 (2): 207–16. doi:10.1016/0304-4149(94)90134-1.

Rothenberg, Thomas J, and James H Stock. 1997. “Inference in a Nearly Integrated Autoregressive Model with Nonnormal Innovations.” *Journal of Econometrics* 80 (2): 269–86.

Said, Said E., and David A. Dickey. 1984. “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order.” *Biometrika* 71 (3): 599–607. doi:10.2307/2336570.

Sala, Xavier, Gernot Doppelhofer Martin, and Ronald I Miller. 2004. “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach.” *The American Economic Review*.

Schotman, Peter C. 1994. “Priors for the AR(1) Model: Parameterization Issues and Time Series Considerations.” *Econometric Theory* 10 (3/4): 579–95.

Schotman, Peter C, and Herman K Van Dijk. 1991. “On Bayesian Routes to Unit Roots.” *Journal of Applied Econometrics* 6 (4): 387–401.

Schwert, G. William. 1989a. “Tests for Unit Roots: A Monte Carlo Investigation.” *Journal of Business & Economic Statistics* 7 (2): 147–59. doi:10.2307/1391432.

———. 1989b. “Tests for Unit Roots: A Monte Carlo Investigation.” *Journal of Business & Economic Statistics* 7 (2): 147–59. doi:10.2307/1391432.

Seo, Byeongseon. 1999. “Distribution Theory for Unit Root Tests with Conditional Heteroskedasticity.” *Journal of Econometrics* 91 (1): 113–44.

Shiller, Robert J, and Pierre Perron. 1985. “Testing the Random Walk Hypothesis: Power versus Frequency of Observation.” *Economics Letters* 18 (4): 381–86.

Sims, Christopher A. 1988. “Bayesian Skepticism on Unit Root Econometrics.” *Journal of Economic Dynamics and Control* 12 (2): 463–74.

Sims, Christopher A, and Harald Uhlig. 1991. “Understanding Unit Rooters: A Helicopter Tour.” *Econometrica: Journal of the Econometric Society*, 1591–99.

Stock, James H. 1994. “Unit Roots, Structural Breaks and Trends.” *Handbook of Econometrics* 4: 2739–2841.

Tanner, Martin A. 1998. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. 3rd edition. New York: Springer.

Tibshirani, Robert. 1996. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)* 58 (1): 267–88.

Wright, Jonathan H. 2008. “Bayesian Model Averaging and Exchange Rate Forecasts.” *Journal of Econometrics* 146 (2): 329–41.

Zellner, Arnold. 1996. *An Introduction to Bayesian Inference in Econometrics*. 1 edition. New York; Chichester: Wiley-Interscience.

Zivot, Eric. 1994. “A Bayesian Analysis of the Unit Root Hypothesis within an Unobserved Components Model.” *Econometric Theory* 10 (3-4): 552–78.

Zivot, Eric, and Jiahui Wang. 2007. *Modeling Financial Time Series with S-PLUS®*. Vol. 191. Springer. http://books.google.com/books?hl=en&lr=&id=Mby9u8_WXXgC&oi=fnd&pg=PA1&dq=Modeling+Financial+Time+Series+with+S-PLUS&ots=ppvXOD137o&sig=MznbxY3r1ntQm8-obUvNSfui9CA.

CHAPTER 2

ROBUST NUMERICAL BAYESIAN UNIT ROOT TEST FOR MODEL UNCERTAINTY

2.1 Introduction

The regular modeling approach in applied economics empirical work is that first researchers select a single model specification which is assumed to be true, then all the analyses are performed based on this pre-determined model. Conclusions will also be drawn from it. This typical practice is concise and effective when the chosen model is somewhat close to the real underlying "data generating process." However, due to various circumstances such as data incompleteness, variable selection, distribution misspecification, etc., this procedure may lead to inappropriate model assumptions which will produce erroneous conclusion since estimates may well depend on the particular model considered (Moral-Benito 2013). As long as empirical research involves model building and data analysis, this problem could potentially exist.

Researchers have developed ways to confront this problem. For example, by constructing several different criteria like Akaike's information criterion (AIC) or Bayesian information criterion (BIC), researchers can compare several candidate models and select one to perform their analysis (Claeskens and Hjort 2008); or use some discrete model selection techniques such as backwards, forwards and stepwise variable selection to filter the variables in hand and reach a parsimony model form. This is usually referred to as variable selection approach and in recently decades numerous new algorithms have

been proposed to continuous selection like Lasso (Tibshirani 1996) or SCAD (Fan and Li 2001). However, these ideas focus more on the uncertainty of variables included in the model while ignoring the uncertainty of model form as well as the stochastic term distribution. This may bring limitations for the application of these methods in economic research since the economic theory should play a more important role when deciding the inclusion of certain variables, rather than the statistical standard.

Another strategy is to consider all possible models and then try to "weighted average" them in some way then make inference based on the whole universe of candidate models (Moral-Benito 2013). This approach is referred as model averaging and generally has two directions: frequentist model averaging (FMA) and Bayesian model averaging (BMA). The method derived in this research fits under the Bayesian model averaging framework, which introduces a model prior distribution to reflect the uncertainty across the possible model space. This BMA approach has been widely adopted to different topics in the economics literature. For instance, in labor economics (Koske and Wanner 2013), health/environmental economics (Morales et al. 2006), migration (Mitchell, Pain, and Riley 2011), and so on. The most active field for application of BMA techniques might be empirical growth models (Moral-Benito 2013), including Sala et al. (2004); it is also being employed in macroeconomics and finance; for example, see Koop and Potter (2004),Wright (2008) and Avramov (2002).

Among the numerous empirical works which have applied BMA, time series (or panel) data sets are commonly utilized. An important issue associated with using such data is the (non)stationarity of the data, based on which the following analysis is valid. Traditional unit root testing methods like Dickey-Fuller test will encounter contradictions

when applied within the BMA framework, since they are not Bayesian and all require some aspects of determined restrictions, either on the functional form, or on the distribution of the error term. For instance, the Dickey-Fuller (DF) test requires the specification of the number of lags included in the model before performing the test; while the Phillips-Perron (PP) test relaxed this requirement, but when estimating using maximum likelihood method it requires certain distribution form of the random term. Model uncertainty should be taken into account in the unit root testing process, no matter the uncertainty existing in the functional form or stochastic term distribution; especially if the researcher believes this uncertainty is an issue in his time series analysis and plans to use model averaging to help. Little research has been done on this topic. Koop et al. (1997) addressed a method to treat a unit root when he analyzed the long memory and persistence properties of real U.S. GNP using model averaging. However this method is particularly focused on the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model he used in the paper. There is no method proposed to process unit root tests issue in general model averaging circumstances. This research hopefully could make contributions to this gap in the literature.

The purpose of this research is to develop a Bayesian unit root test robust to model averaging, which can consider not only the uncertainty of the functional form, but also various assumptions of the distribution of the random component. Meanwhile, we also hope to make the test robust to possible transformations of the data (like Box-Cox transformation).

The structure of this paper is as follows. Section 2 motivates the idea of the proposed unit root test method and details the methodology including mathematical

derivation of the method, prior distribution, posterior distribution as well as some generalizations like transformation. After the comprehensive introduction, the Monte Carlo simulation design designated to test the suggested method is elaborated in Section 3. The simulation results are shown together with some comparisons with existing popular unit root test methods. In Section 4, we apply our Bayesian unit root test to 14 macroeconomic series from the famous Nelson and Plosser (1982) paper and discuss the results compared to some previous work which also tested the same data sets, concentrating on the model uncertainly problem. Section 5 concludes the whole paper.

2.2 Methodology

2.2.1 Model Parameterization

As most researchers did in the unit root test literature, this paper will derive the theory based on the fundamental univariate autoregression formula which has the following form (take AR(1) as an example now):

$$y_t = \gamma + \delta t + x_t,\ x_t = \rho x_{t-1} + \varepsilon_t \tag{1}$$

And $\varepsilon_t \sim f(0,\sigma^2)$ has some distribution with mean 0 and variance σ^2 . This can be any distribution as long as it satisfies the white noise assumption. The unit root situation corresponds to ρ=1. This parameterization was first considered by Bhargava (1986), and discussed and adopted again by Phillips and Schmidt (1989), and following the terminology used in Lubrano (1995), this parameterization will be called a structural parameterization. For a unit root test we are not only interested in whether a series is difference stationary (i.e. ρ=1), but also commonly test if a series is trend stationary. That is, one wishes to test for a unit root in the presence of a deterministic trend (Phillips and

Schmidt 1989). Deterministic terms play a key role in the dynamic behavior of the process and this dynamic pattern can be very different according to the parameterization of the model. The model forms in the original work by Dickey and Fuller (1979) are not well designed for testing trend reversion since it is either based on model without trend, or the model with trend but whose existence as well as distribution depend on the null hypothesis we want to test (Phillips and Schmidt 1989). The parameterization in (1) mitigates these problems. In the model mentioned above, parameter γ represents level (drift) and δ takes the place of deterministic trend, no matter if ρ equals 1 or not. So it permits for trend under both the null and the alternative hypothesis, without introducing any parameters that are irrelevant under either (Phillips and Schmidt 1989). Lubrano(1995) also provided a detailed comparison between this structural parameterization with a traditional model form which he referred as "reduced form parameterization". This structural form will not only provide consistency under both hypotheses, but also will benefit to the method derived in this paper since in Bayesian framework priors need to be specified to every parameter, and it will be found during the following discussion that this consistency will allow both theoretically unambiguous and computationally tractable analysis of the hypotheses of interest.

While previous equation (1) is a fundamental AR(1) process, to incorporate the more complicated and general form which is needed in the method to be derived in this paper, the common AR(p) process is necessary. Introducing a lag polynomial to the second part of (1) gives the following form:

$$A(L)x_t = \varepsilon_t \tag{2}$$

where

$$A(L) = 1 - \rho_1 L - \rho_2 L^2 - ... - \rho_p L^p \tag{3}$$

The "L" is a lag operator which makes $Lx_t = x_{t-1}$ and in general $L^k x_t = x_{t-k}$. Then the model becomes:

$$y_t = \gamma + \delta t + x_t,\ A(L)x_t = \varepsilon_t,\ \varepsilon_t \sim f(0, \sigma^2) \tag{4}$$

It can be written in a form that is more familiar like regular a AR(P) model with some rearrangement:

$$\begin{aligned} & A(L)\,\mathrm{y}_t = \gamma + A(L)\delta t + A(L)x_t \\ & \Rightarrow y_t = \gamma + A(L)\delta t + \rho_1 y_{t-1} + \rho_2 y_{t-2} + ... + \rho_p y_{t-p} + \varepsilon_t \\ & \Rightarrow y_t = (1 - \sum_{i=1}^{p} \rho_i)\gamma + [\sum_{i=1}^{p} i\rho_i + (1 - \sum_{i=1}^{p} \rho_i)t]\delta + \rho_1 y_{t-1} + \rho_2 y_{t-2} + ... + \rho_p y_{t-p} + \varepsilon_t \end{aligned} \tag{5}$$

The main focus of a unit root test would be the coefficients of different lags, and from (5) it can be seen that all the other nuisance parameters are linear, so for convenience of derivation we denote $\boldsymbol{\theta} = (\gamma, \delta)'$, $\mathbf{x}_t = [(1 - \sum_{i=1}^{p} \rho_i), \sum_{i=1}^{p} i\rho_i + (1 - \sum_{i=1}^{p} \rho_i)\,t]'$, $\boldsymbol{\rho} = (\rho_1, \rho_2 ..., \rho_p)$, and $\mathbf{y}_{-1} = (y_{t-1}, y_{t-2}, ..., y_{t-p})'$. The model can be simplified to:

$$y_t = \mathbf{x}_t'\boldsymbol{\theta} + \mathbf{y}_{-1}'\boldsymbol{\rho} + \varepsilon_t, \varepsilon_t \sim f(0, \sigma^2) \tag{6}$$

2.2.2 Likelihood Function and the Initial Condition

Although the stochastic term allows any distribution which satisfies the white noise assumption, for the convenience of derivation and without loss of generality, right now we assume that it follows a normal distribution with parameters specified above. So the likelihood function is:

$$L(\mathrm{y} \mid \boldsymbol{\theta}, \boldsymbol{\rho}, \sigma^2) = (2\pi\sigma^2)^{-T/2} \exp\{-\frac{1}{2\sigma^2} \sum_{t=\mathrm{p}+1}^{T+p} (y_t - \mathbf{x}_t'\boldsymbol{\theta} - \mathbf{y}_{-1}'\boldsymbol{\rho})^2\} \tag{7}$$

Notice that there should be (T+p) observations in the data in order to support autoregression lag p.

It is important to notice that the time series data in hand for analysis is usually only a portion of a dynamic process that may have started either in the infinite past or in a distant past (Bauwens, Lubrano, and Richard 2000a). Basically the initial observation y_0 is crucial information we may not know about the past so it is important to model the initial condition as well as to measure its impact as ρ approaches one. In the literature there are two main methods to treat the initial observation: consider the initial condition y_0 as fixed and compute the likelihood function conditional on their first observation (Phillips 1991), or treat the initial conditions as random with some certain distribution derived from the assumption about the stochastic term (Zivot 1994; Lubrano 1995). Zellner (1996) proved that in the general stationary case and without a constant term, the consideration of fixed or random initial condition is not of importance. However, since mentioned before, it is of interest to include a constant as well as a deterministic trend in the test, the initial condition problem should be discussed here. Lubrano (1995) provided a detailed discussion and proof about the problems and consequences of treating y_0 differently in various circumstances which is highly recommended to refer to. According to his conclusion, treating y_0 as fixed will cause the derived posterior distribution to be unbounded at the unit root case when the constant term is present, which makes it is unsuitable for Bayesian analysis. In contrast, random initial observation with certain distribution will overcome the aforementioned problem and provide a finite likelihood function at ρ=1 even when there is a constant term (Lubrano 1995). Based on these results we will treat the initial condition as random in our analysis discussed below. To be

more specific, with normal distribution assumption mentioned above for the convenience of discussion, assume the process is AR(1) and starts from some past date at $t = -s$ at which $\mu_{-s} = \varphi$, Griffiths (2012) showed that the initial observation y_0 has the distribution

$$f(\mathrm{y}_0 \mid \mu, \sigma^2, \varphi, \mathrm{s}) = \mathrm{N}(\mu + \rho^s \varphi, \sigma^2 \, \mathrm{q}(\rho, \mathrm{s}))$$
$$\mathrm{q}(\rho, \mathrm{s}) = \sum_{i=0}^{s} \rho^{2i} = \frac{1-\rho^{2(\mathrm{s}+1)}}{1-\rho^2} \tag{8}$$

Under the stationary assumption, in the infinite time horizon the distribution can be shown as $y_0 \sim N(\mu, \sigma^2/(1-\rho^2))$. For the general AR(p) process, the exact likelihood function would require the first p observations to be random. Hamilton (1994, 2:125) derived this joint density which is fairly complicated. An alternative applied in this research is treating only y_0 as random while considering $(\mathrm{y}_1, \mathrm{y}_2, ..., \mathrm{y}_{P-1})$ as fixed (Schotman and Van Dijk 1991; Lubrano 1995). Meanwhile, Bauwens, Lubrano, and Richard (2000b, 182) provided a functional form which behaves similarly with $\mathrm{q}(\rho, \mathrm{s})$ in (8) and that will be used in this analysis:

$$q(\rho, \upsilon) = \frac{1+\upsilon}{1+\upsilon-\rho^2} \tag{9}$$

This function becomes usual modeling of y_0 under stationary assumption when υ=0, and when $\upsilon \to \infty$ the function equals to 1. In between, it is possible to find s and υ so that they have a similar behavior in the dominant root (Bauwens, Lubrano, and Richard 2000b). A prior could also be assigned to υ and will be discussed in detail in the following part. Then the likelihood function is calculated as:

$$
\begin{aligned}
L(\mathrm{y}\,|\,\boldsymbol{\theta},\boldsymbol{\rho},\sigma^2) &= (2\pi)^{-(\mathrm{T}+1)/2}(\sigma^2)^{-(T+1)/2}q(\rho,\mathrm{s})^{-1/2} \\
&\quad\times\exp\{-\frac{1}{2\sigma^2}[\sum_{t=\mathrm{p}+1}^{T+p}(y_t-\mathbf{x}'\boldsymbol{\theta}-\mathbf{y}'\boldsymbol{\rho})^2+\frac{1}{q(\rho,\mathrm{s})}(\mathrm{y}_0-\mu)^2]\} \\
&\propto q(\rho,\mathrm{s})^{-1/2}(\sigma^2)^{-(T+1)/2} \\
&\quad\times\exp\{-\frac{1}{2\sigma^2}[\sum_{t=\mathrm{p}+1}^{T+p}(y_t-\mathbf{x}'\boldsymbol{\theta}-\mathbf{y}'\boldsymbol{\rho})^2+\frac{1}{q(\rho,\mathrm{s})}(\mathrm{y}_0-\mu)^2]\}
\end{aligned}
\tag{10}
$$

in which $q(\rho,\mathrm{s})$ is defined in (9).

2.2.3 Bayesian Unit Root Test

The prevalence of Bayesian unit root tests can be traced back to Sims (1988) and Sims and Uhlig (1991) who detailed the differences and advantage of Bayesian unit root tests compared to classical unit root test methods. Since then, extensive literature has been developed focusing on the analysis of this topic from various aspects. The most common approach is specifying prior distributions to the lag coefficients and performing inferences based on the distribution of $\widehat{\rho}|\rho$. This research will adopt a slightly different approach – instead of specifying priors for all the coefficients, we assign priors to the roots which determine the stochastic dynamic behavior of the system. The method is generalized from Dorfman (1993) who presented some early Bayesian numerical unit root tests. In his paper, the dynamic process of AR(p) is reparameterized to a state space model form whose dynamic properties can be investigated by studying the following matrix:

$$
A=\begin{bmatrix}
\rho_1 & 1 & 0 & 0 & \cdots & 0 \\
\rho_2 & 0 & 1 & 0 & \cdots & 0 \\
\rho_3 & 0 & 0 & \ddots & & \vdots \\
\vdots & 0 & 0 & \cdots & 0 & 1 \\
\rho_p & 0 & 0 & \cdots & 0 & 0
\end{bmatrix}
\tag{11}
$$

The $\boldsymbol{\rho} = (\rho_1, \rho_2, ..., \rho_P)'$ are coefficients of lags in the autoregression model. The roots, or eigenvalues of this matrix, determine the behavior of the modeled process. If any eigenvalues of $\mathbf{A}$ have moduli greater or equal than one, the time series defined by these coefficients has a unit root and is nonstationary. So denote $\phi_i = \|\lambda_i\|$ which is the moduli of the eigenvalues of $\mathbf{A}$ sorted by its magnitude where $\|a\|$ stands for the Euclidian distance of a. Thus the statistical hypothesis of unit root can be expressed as:

$$H_0 : \phi_1 \geq 1 \text{ vs. } H_1 : \phi_1 < 1 \tag{12}$$

Unlike the classical unit root test method which focuses on deriving the asymptotically distribution of the coefficient on which the inference is to be made, in the Bayesian framework there is really no "test" concept similar to the classical world since the Bayesian method requires a posterior distribution which combines the likelihood function with prior distributions and usually does not have standard form to perform regular "test". Instead, two alternative methods are usually used in practice: the Bayes factor and credible interval. The credible interval method is similar to the confidence interval in the under classical statistical inference framework but is not quite the same. Usually, we reject H_0 if $\phi_1 \geq 1$ lies outside the 95% (or any other preferred range) credible interval defined by:

$$\int_{\lambda_{\text{inf}}}^{\lambda_{\text{sup}}} f(\lambda \mid \boldsymbol{\Omega}, \mathbf{y}) \, \mathrm{d}\lambda = 0.95 \tag{13}$$

Here λ is the dominant root and $f(\lambda \mid \boldsymbol{\Omega}, \mathbf{y})$ is the marginal posterior distribution with respect to λ (this will be discussed in detailed later) with parameter space $\boldsymbol{\Omega}$ and data $\mathbf{y}$.

An alternative choice is the Bayes factor approach. The Bayes factor B in favor of H_0 against H_1 is defined as (Lee 2012):

$$B = \frac{(p_0 / p_1)}{(\pi_0 / \pi_1)} = \frac{p_0 \pi_1}{p_1 \pi_0} \tag{14}$$

The elements are defined as follows:

$$\begin{aligned} p_0 &= P(\theta \in \Theta_0 \mid \mathbf{x}) \\ p_1 &= P(\theta \in \Theta_1 \mid \mathbf{x}) \\ \pi_0 &= P(\theta \in \Theta_0) \\ \pi_1 &= P(\theta \in \Theta_1) \end{aligned} \tag{15}$$

in which θ is the unknown parameter from a set Θ ($\Theta_0 \bigcup \Theta_1 = \Theta, \Theta_0 \bigcap \Theta_1 = \varnothing$) we want to test, and $\mathbf{x}$ is the data available whose density $f(\mathbf{x} \mid \theta)$ depends on θ. Here p_0 and p_1 are posterior probabilities of parameter supports of null (Θ_0) and alternative (Θ_1) hypotheses, while π_0 and π_1 are prior probabilities of the two hypotheses parameters respectively. The Bayes factor is sometimes can be interpreted as "odds in favor of H_0 against H_1 that are given by the data" (Lee 2012). In this paper, we will use the Bayes factor approach since it is more convenient to numerical Bayesian technique. Let $\boldsymbol{\eta} = (\gamma, \delta, \sigma^2)'$, $\boldsymbol{\Phi} = (\phi_1, \phi_2, ..., \phi_p)'$ be the parameters with prior distribution $\pi(\boldsymbol{\Phi}) = \pi(\phi_1, \phi_2, ..., \phi_p)$ and $h(\boldsymbol{\eta}) = h(\gamma, \delta, \sigma^2)$ (the detailed prior distribution discussion will be given in next section). With the likelihood function denoted as $f(\mathrm{y}; \boldsymbol{\eta}, \boldsymbol{\Phi})$, the marginal posterior of the dominant root would be:

$$p(\phi_1 \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathrm{y}) = \int_{\phi_2} \cdots \int_{\phi_p} \int_{(\boldsymbol{\Phi}, \boldsymbol{\eta})} f(\mathrm{y}; \boldsymbol{\eta}, \boldsymbol{\Phi}) \pi(\boldsymbol{\Phi}) h(\boldsymbol{\eta}) d\boldsymbol{\eta} d\phi_2 \cdots d\phi_p \tag{16}$$

Using numerical methods (to be discussed later) samples can be drawn from this marginal posterior distribution. Define an indicator function:

$$D(\phi_{(1)}) = \begin{cases} 1 & \phi_{(1)} < 1 \text{ in the sample (support } \mathrm{H}_1) \\ 0 & \text{otherwise (support } \mathrm{H}_0) \end{cases} \tag{17}$$

The posterior probability in support H_1 is given by:

$$K_1 = p(\mathrm{H}_1 \mid \mathbf{y}) = \frac{\sum_{i=1}^{B} D(\phi_{(1)})\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{y})}{\sum_{i=1}^{B} p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{y})} \tag{18}$$

where B is the number of valid iterations used in the sampler. Defining K_0 by analogous way to (18):

$$K_0 = P(\mathrm{H}_0 \mid \mathbf{y}) = \frac{\sum_{i=1}^{B} [1 - D(\phi_{(1)})]\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{y})}{\sum_{i=1}^{B} p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{y})} \tag{19}$$

and the posterior odds ratio in favor of nonstationarity is:

$$K_{01} = \frac{\sum_{i=1}^{B} [1 - D(\phi_{(1)})]\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathrm{y})}{\sum_{i=1}^{B} D(\phi_{(1)})\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathrm{y})} = \frac{K_0}{K_1} \tag{20}$$

This statistic is the center of the numerical Bayesian unit test we used in our method. And one can declare the series to have a nonstationary root if $K_{01} > 1$ (or using different threshold depending on loss function specified).

2.2.4 Prior Distributions, Posterior Distributions and Sampling Methods

The choice of priors to the parameters is always a crucial issue in Bayesian econometric analysis. Sims (1988) adopted flat uninformative priors in his paper, while later Phillips (1991) criticized the flat priors by evidence from simulation and comparison with reference Jeffreys prior, claiming that it is not truly informative and will bias the results. Berger and Yang (1994) compared various approaches to the development of a noninformative prior for the AR(1) model and pointed out the Jeffreys prior has two features: depends on the initial observation and favor of explosive case when sample size is large, which bring problems when performing unit root tests. They derived a reference prior approach which worked well for the stationary case but less satisfactory in the explosive case. Recently, Griffiths (2012) performed a Monte Carlo experiment to examine the size and power properties of Bayesian unit root tests with four different prior distributions. In light of this continuing debates, in this research we do not spent much time on judging the priors since this is not the main point of this research and it is still hard or impossible to find a perfect priors that works well under all circumstances due to the fact that it may well depend on the data and topic (preknowledge). We just applied priors similar to those in Dorfman (1993): independent Beta priors were assigned to the moduli of roots of (11), while maintaining an uninformative priors to the other parameters. Meanwhile, we wish to allow slightly explosive values of the dominant root since that incorporates the possibility of upward bias due to sampling error from estimation of the root from the posterior distribution (Dorfman 1993). To sum up, these informative priors on all the roots are:

$$\begin{aligned} &\pi(\mathbf{\Phi}) = \pi(\phi_1, \phi_2, ..., \phi_p) = \pi_1(\phi_1)\pi_2(\phi_2)...\pi_p(\phi_p) \\ &\pi_1(\phi_1 - \upsilon) \sim \text{Beta}(30, 2); \pi_2(\phi_2) = .\pi_3(\phi_3) = ... = \pi_p(\phi_p) \sim \text{Beta}(1.1,1.1) \end{aligned} \tag{21}$$

These priors are informative on the moduli of dominant root with severely skewed to the right ranging from υ to $(1+\upsilon)$ with mode around 1. The priors on all the other roots are quite weakly informative with very flat shape, round curves ranging over [0,1]. For the issue of choosing υ, by the recommendation of Bauwens, Lubrano, and Richard (2000) on υ=1/3, we follows Griffiths (2012) to adopt a hierarchical approach since his work suggests that assigning a prior to υ is a good strategy. We assign an exponential distribution with mean $E(\upsilon)=\kappa^{-1}$ as a prior to υ, i.e.,

$$f(\upsilon \mid \kappa)=\kappa \mathrm{e}^{-\kappa\upsilon} \tag{22}$$

where κ is chosen such that $P(\upsilon > 0.15)=0.05$.

For the drift and coefficient of deterministic trend, we assign bivariate normal distributions as their priors, which has the following form:

$$\begin{pmatrix}\gamma\\ \delta\end{pmatrix} \sim \mathrm{MVN}(\boldsymbol{\omega}, \boldsymbol{\Sigma}), \boldsymbol{\omega}=\begin{pmatrix}\gamma_0\\ \delta_0\end{pmatrix}, \boldsymbol{\Sigma}=\begin{pmatrix}\omega_1^2 & 0\\ 0 & \omega_2^2\end{pmatrix} \tag{23}$$

Without loss of generality, we assume $\gamma_0=\delta_0=0$ for computational simplicity. The variance of the error term is assumed to follow an inverse-gamma distribution with shape parameter α and scale parameter β which is the conjugate prior of the normal likelihood function:

$$\sigma^2 \sim IG(\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}[\sigma^2]^{-\alpha-1}\exp(-\frac{\beta}{\sigma^2}) \tag{24}$$

Since these three parameters are assumed to be independent, the joint prior of all parameters except the roots shown in (15) is:

$$h(\boldsymbol{\eta}) = h(\gamma, \delta, \sigma^2) = f(\gamma, \delta) f(\sigma^2) \tag{25}$$

So with all these settings, the joint posterior distribution of the parameters is:

$$f(\boldsymbol{\Phi}, \boldsymbol{\eta}, \upsilon \mid \mathbf{y}) \propto L(\mathrm{y} \mid \boldsymbol{\eta}, \boldsymbol{\rho}) \times \pi(\boldsymbol{\Phi}) \times h(\boldsymbol{\eta}) \times f(\upsilon \mid \kappa) \tag{26}$$

where $L(\mathrm{y} \mid \boldsymbol{\eta}, \boldsymbol{\rho})$, $\pi(\boldsymbol{\Phi})$, $h(\boldsymbol{\eta})$ and $f(\upsilon \mid \kappa)$ are defined in (10), (21), (25) and (22) respectively. The analytical form of the marginal posterior distribution of the nuisance parameters can be derived as follows:

(1). The posterior for variance σ^2 :

$$\begin{aligned} & f(\sigma^2 \mid \boldsymbol{\Phi}, \gamma, \delta, \upsilon, \mathbf{y}) \propto f(\boldsymbol{\eta}, \boldsymbol{\rho}) \bullet f(\sigma^2) \\ & \propto (\sigma^2)^{-\alpha - T/2 - 1} \exp\{-\frac{1}{2\sigma^2}[\sum_{t=\mathrm{p}+1}^{T+p} (y_t - \mathbf{x}_{\mathrm{t}}^{'}\boldsymbol{\theta} - \mathbf{y}_{-1}^{'}\boldsymbol{\rho})^2 + 2\beta]\} \\ & \sim IG(\alpha + \frac{T}{2}, \frac{1}{2}[\sum_{t=\mathrm{p}+1}^{T+p} (y_t - \mathbf{x}_{\mathrm{t}}^{'}\boldsymbol{\theta} - \mathbf{y}_{-1}^{'}\boldsymbol{\rho})^2 + 2\beta]) \end{aligned} \tag{27}$$

which also follows inverse-gamma distribution.

(2). The posterior for the drift term γ :

From the assumption of (23) the marginal prior distribution of γ is $N(\gamma_0, \sigma_\gamma^2)$. Assume the following notation: $A = y_t - [\sum_{i=1}^{p} i\rho_i + (1 - \sum_{i=1}^{p} \rho_i)t]\delta - \mathbf{y}_{\text{-}1}^{'}\boldsymbol{\rho}$, the posterior distribution of γ is (several intermediate steps are omitted):

$$
\begin{aligned}
f(\gamma|\mathbf{\Phi},\delta,\sigma^2,\upsilon,\mathbf{y}) &\propto \exp\Big\{-\frac{1}{2\sigma^2}\sum_{t=p+1}^{T+p}[\mathrm{A}-\gamma(1-\sum_{i=1}^{p}\rho_i)]^2\Big\}\exp(-\frac{\gamma^2}{2\omega_1^2}) \\
&\propto \exp\Big\{\frac{1}{\sigma^2}(1-\sum_{i=1}^{p}\rho_i)\sum_{i=p+1}^{T+p}A\cdot\gamma-(\frac{T(1-\sum_{i=1}^{p}\rho_i)^2}{2\sigma^2}+\frac{1}{2\omega_1^2})\gamma^2\Big\} \\
&\propto \exp\Big\{-\frac{1}{2}[\frac{\omega_1^2T(1-\sum_{i=1}^{p}\rho_i)^2+\sigma^2}{\sigma^2\omega_1^2}\gamma^2-\frac{2}{\sigma^2}(1-\sum_{i=1}^{p}\rho_i)\sum_{i=p+1}^{T+p}A\cdot\gamma]\Big\} \\
&\propto \exp\Big\{-\frac{1}{2}\cdot\frac{\omega_1^2T(1-\sum_{i=1}^{p}\rho_i)^2+\sigma^2}{\sigma^2\omega_1^2}[\gamma-(\frac{\omega_1^2(1-\sum_{i=1}^{p}\rho_i)\sum_{i=p+1}^{T+p}A}{\omega_1^2T(1-\sum_{i=1}^{p}\rho_i)^2+\sigma^2})]^2\Big\}
\end{aligned}
\tag{28}
$$

Thus, the marginal posterior distribution of drift term γ is:

$$
\gamma\sim N\left([\frac{\omega_1^2(1-\sum_{i=1}^{p}\rho_i)\sum_{i=p+1}^{T+p}A}{\omega_1^2T(1-\sum_{i=1}^{p}\rho_i)^2+\sigma^2}],\frac{\sigma^2\omega_1^2}{\omega_1^2T(1-\sum_{i=1}^{p}\rho_i)^2+\sigma^2}\right) \tag{29}
$$

(3). The posterior for the deterministic trend coefficient δ:

The marginal prior distribution of δ is also normal of $N(\delta_0,\sigma_\delta^2)$ from (23). Similarly denote $B=y_t-(1-\sum_{i=1}^{p}\rho_i)\gamma-\mathbf{y}_{-1}'\boldsymbol{\rho}$ and $S=(1-\sum_{i=1}^{p}\rho_i)t+\sum_{i=1}^{p}i\rho_i$ in model (5), the posterior distribution of δ can be derived as (several intermediate steps are omitted):

$$f(\delta|\mathbf{\Phi},\gamma,\sigma^2,\upsilon,\mathbf{y}) \propto \exp\left\{-\frac{1}{2\sigma^2}[\sum_{t=p+1}^{T+p}(y_t-(1-\sum_{i=1}^{p}\rho_i)\gamma-\mathbf{y}_{-1}'\boldsymbol{\rho}-\delta(1-\sum_{i=1}^{p}\rho_i)t+\delta\sum_{i=1}^{p}i\rho_i)^2]-\frac{\delta^2}{2\omega_2^2}\right\}$$

$$\propto \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=p+1}^{T+p}[\mathrm{B}-S\delta)]^2-\frac{\delta^2}{2\omega_2^2}\right\}$$

$$\propto \exp\left\{-\left(\frac{\sum_{t=p+1}^{T+p}S^2\omega_2^2+\sigma^2}{2\sigma^2\omega_2^2}\right)\delta^2+\frac{1}{\sigma^2}\mathrm{B}\sum_{t=p+1}^{T+p}S^2\cdot\delta\right\} \tag{30}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\frac{\sum_{t=p+1}^{T+p}S^2\omega_2^2+\sigma^2}{\sigma^2\omega_2^2}\right)\left[\delta-\frac{\omega_2^2\,\mathrm{B}\sum_{t=p+1}^{T+p}S}{\omega_2^2\sum_{t=p+1}^{T+p}S^2+\sigma^2}\right]^2\right\}$$

So the marginal posterior distribution of the deterministic trend coefficient δ is:

$$\delta \sim N\left(\frac{\omega_2^2\,\mathrm{B}\sum_{t=p+1}^{T+p}S}{\sigma^2+\omega_2^2\sum_{t=p+1}^{T+p}S^2},\frac{\sigma^2\omega_2^2}{\sigma^2+\omega_2^2\sum_{t=p+1}^{T+p}S^2}\right) \tag{31}$$

The posterior distribution of the eigenvalues $\mathbf{\Phi}$ is fairly difficult to be analytically derived since the model parameterization using lag coefficients $\boldsymbol{\rho}$ whereas the prior distribution is specified to $\mathbf{\Phi}$. Meanwhile it is also very difficult to derive the posterior distribution of υ. A numerical method has to be adopted to sample from the marginal posterior distribution of these parameters and derive measures of their posterior empirically.

From the above derivation, some parameter posterior distributions can be analytically calculated while others cannot. To better incorporate this property and achieve more accurate and faster sampling result, Metropolis–Hastings within Gibbs (MH within Gibbs) algorithm will be adopted in our computation. It is one of the Markov

Chain Monte Carlo (MCMC) methods that samples from a conditional distributions of subsets of parameters in order to constructing a Markov chain whose convergence distribution is the desired distribution. There are basically two fundamental mechanisms among MCMC sampling algorithms (Jing 2010). The first approach sets up an "accept/reject" rule to keep revising the chain so that it can reach the aimed convergence distribution, which is the spirit of the Metropolis-Hastings (MH) algorithm. It was first developed by Metropolis et al. (1953) and later generalized by Hastings (1970) and now is widely used in all areas which need numerical sampling including statistics, econometrics, physics, engineering. Assume that random variables $\mathbf{X}$ have joint distribution $\psi(\mathbf{X})$ from which the sample need to be drawn. A MH algorithm requires choosing a proposal distribution $q(\cdot \mid \mathbf{X})$ which generates a candidate chain based on the current values of random variables, then decides to accept or reject according to the detailed balance rule. To be more specific, the algorithm can be described as below:

(1). Assign initial values to $\mathbf{X}: \mathbf{X^{(0)}} = (\mathrm{x}_1^{(0)}, \mathrm{x}_2^{(0)}, ..., \mathrm{x}_n^{(0)})$;

(2). For iteration m from 1 to N, repeat the following procedures:

a. generate a candidate value from proposal density, $\mathbf{X}^* \sim q(\cdot \mid \mathbf{X}^{\mathrm{m}})$;

b. generate a random variable $u \sim Uniform(0,1)$;

c. decide using the following rule:

$$\mathbf{X}^{(m+1)} = \begin{cases} \mathbf{X}^* & \text{if } u \le \alpha(\mathbf{X}^{(m)}, \mathbf{X}^*) \\ \mathbf{X}^{(m)} & \text{otherwise} \end{cases}, \text{ in which}$$

$$\alpha(\mathbf{X}^{(m)}, \mathbf{X}^*) = \begin{cases} \min\{\dfrac{\psi(\mathbf{X}^*)\, q(\mathbf{X}^{(m)} \mid \mathbf{X}^*)}{\psi(\mathbf{X}^{(m)})\, q(\mathbf{X}^* \mid \mathbf{X}^{(m)})}, 1\} & \text{if } \psi(\mathbf{X}^{(m)})\, q(\mathbf{X}^* \mid \mathbf{X}^{(m)}) > 0 \\ 1 & \text{otherwise} \end{cases}$$

(3). Return the final values $(\mathbf{X}^{(0)}, \mathbf{X}^{(1)}, ..., \mathbf{X}^{(N)})$.

The MH algorithm is flexible since the proposal density can be arbitrary, but there is a price for this convenience (Jing 2010). If the proposal density is not chosen appropriately, the acceptance rate is low which means that the sample we create may include lots of noise compared to the convergence distribution we actually want. Also the Markov Chain may not cover the whole support or (at least a large part) of the convergence distribution; sometimes it even can stick to a small area which leads to very poor sampling efficiency.

The second approach to a MCMC mechanism simplifies high dimensional situations by successively generating each variable or subset of variables from probability distributions conditional on the previously drawn samples, which is the general idea of a Gibbs sampler. The Gibbs sampling algorithm is one of the simplest but most powerful Markov chain Monte Carlo algorithms which was introduced by Geman and Geman (1984) in the context of image processing. Today Gibbs sampling is commonly used as a means of statistical inference, especially Bayesian inference. Gelfand and Smith (1990) demonstrate the importance of the Gibbs algorithm for a range of problems in Bayesian analysis with comparison to other sampling methods. Unlike the MH algorithm which requires a proposal density, the Gibbs sampler generates samples from full conditional distributions and is guaranteed to converge to the full joint probability distributions under mild regularity conditions introduced by Roberts and Smith (1994). In this paper we use the simplest but most widely used version of the Gibbs sampler which can be described below:

(1). Starting with some initial values $\mathbf{X}^{(0)} = (x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)})$;

(2). For iteration m from 1 to N, repeat the following procedures:

For variable i from 1 to n, generate sample x_i^m from the conditional probability density $f(x_i \mid x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)$;

(3). Return the final values $(\mathbf{X^{(0)}}, \mathbf{X^{(1)}}, ..., \mathbf{X^{(N)}})$.

The price that the Gibbs sampler pays for reducing the dimensionality is slow convergence and high correlation when the components of sampled random variables exhibit heavy dependence (Jing 2010).

In this paper, we combined these two methods since the marginal posterior distribution of part of the parameters can be analytical derived conditional on all others while others cannot, which requires MH algorithm. To be more specific, we use the multivariate normal distribution as the proposal distribution density of the MH algorithm to sample the roots $\mathbf{\Phi}$; after that, conditional on all the other random variables draws of nuisance parameters $\boldsymbol{\eta}$ can be achieved one by one using a Gibbs sampler based on (27), (29) and (31). The details and results will be given in the next section.

2.2.5 Allowing Variable Transformations

In economic empirical work, data transformation is usually needed either for achieving some assumption (like normality of the distribution) or for the convenience of interpretation of the coefficients (e.g., coefficients can be interpreted as elasticity for a linear model with log transformed data). It is also the case in time series analysis. Meanwhile, model transformation can also be treated as one type of model specification uncertainty which can be considered in our Bayesian unit root test method. So in this paper we also want to incorporate variable transformation into our method. Specifically, we consider the famous and widely used Box-Cox transformation (Box and Cox 1964) which is:

$$y_t = \begin{cases} \frac{y_t^{\lambda} - 1}{\lambda}, if \ \ \lambda \neq 0 \\ \log(y_t), if \ \ \lambda = 0 \end{cases} \tag{32}$$

There are lots of other kinds of transformations that can be used in our general framework, and one can refer to Charitidou, Fouskakis, and Ntzoufras (2013) for comparison of different transformation from a Bayesian perspective if interested in transformations.

In order to average together with other models which use the un-transformed variables, the likelihood function needs to modified and becomes:

$$f(y^{\lambda} \mid \boldsymbol{\theta}, \boldsymbol{\rho}, \sigma^2) = (2\pi\sigma^2)^{-T/2} \exp\{-\frac{1}{2\sigma^2} \sum_{t=p+1}^{T+p} (y_t^{\lambda} - \mathbf{x}'\boldsymbol{\theta} - \mathbf{y}^{\lambda}{}'\boldsymbol{\rho})^2\} \times \prod_{t=P+1}^{T+P} |\frac{\partial y_t^{\lambda}}{\partial y_t}| \tag{33}$$

in which $|\frac{\partial y_t^{\lambda}}{\partial y_t}|$ is the Jacobian for the corresponding transformation and $\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\rho}$ are defined the same as in (10). Then similar to before, we can derive the posterior distribution by multiplying the likelihood function by the prior distribution and integrating out the irrelevant term. Notice here to incorporate the Box-Cox transformation the only major difference is the likelihood function (the Jacobian part) while other components like initial conditions and prior distributions do not need to be adjusted so the detailed derivation will not be repeated here.

2.2.6 Unit Root Test Using Bayesian Model Averaging

After all these derivations and interpretations, the last step is considering the model averaging which is the key contribution of this research. Suppose there are k possible candidate models which can hold differences in various aspects, for instance, different numbers of time lags can be included in the model, and the model could be with

or without deterministic trend, also there could be some models using the transformed data. We assign a model prior probability distribution across the possible model space, so each potential model has prior probability $pr(\mathrm{M}_k)$,

$$pr(\mathrm{M}_k \mid \mathrm{y}) = \frac{pr(\mathrm{y} \mid \mathrm{M}_k)\, pr(\mathrm{M}_k)}{\sum_{l=1}^{k} pr(\mathrm{y} \mid \mathrm{M}_l) pr(\mathrm{M}_l)} \tag{34}$$

and the marginal likelihood function under M_k is

$$pr(\mathrm{y} \mid \mathrm{M}_k) = \int_S pr(\mathrm{y} \mid \boldsymbol{\Omega}_{\mathbf{k}}, \mathrm{M}_k) pr(\boldsymbol{\Omega}_{\mathbf{k}} \mid \mathrm{M}_k)\, \mathrm{d}\, \boldsymbol{\Omega}_{\mathbf{k}} \tag{35}$$

$\boldsymbol{\Omega}_{\mathbf{k}}$ here is all the parameters in the model M_k (in our case, $\Omega_k = (\boldsymbol{\Phi}_k, \boldsymbol{\eta}_k, \upsilon_k)'$) and $pr(\boldsymbol{\Omega}_{\mathbf{k}} \mid \mathrm{M}_k)$ is the prior distribution of parameters $\boldsymbol{\Omega}$ under model M_k , and S is the support of $\boldsymbol{\Omega}$.Then the final comprehensive probability of possible unit root across the whole model space is:

$$pr(\phi_1 \geq 1 \mid \mathrm{y}) = \sum_{i=1}^{k} pr(\phi_1 \geq 1 \mid \mathrm{M}_k, \mathrm{y}) pr(\mathrm{M}_k \mid \mathrm{y}) \tag{36}$$

in which $pr(\phi_1 \geq 1 \mid \mathrm{M}_k, \mathrm{y})$ is the posterior probability of unit root given model M_k derived in (16).

Finally we can compute the probability of the dominant root larger than 1 considering all these circumstances including model uncertainty on functional form, random component distribution and transformation and draw the conclusion according to the decision rule defined in (20).

2.3 Simulation and Results

2.3.1 Monte Carlo Simulation Design

Monte Carlo simulations are adopted to investigate the performance of the method we derived. In general, a desirable test should display accurate test size and maximize test power (Griffiths 2012) so it would be a good strategy to compare the proposed method with some existing unit root test methods. However, one thing to remember is that Bayesian and frequentist tests are completely different logically. For instance, Bayesian tests do not conform to the conventional 0.05 size from the traditional sampling theory (Griffiths 2012). Also one fundamental difference between these two frameworks is that the traditional frequentist framework treats unknown parameters as fixed and data as random, while the Bayesian technique believes data is fixed and treats parameters as random. In this case it might be inappropriate to evaluate a "pure" Bayesian test which makes decisions conditional on the fixed observed data using criteria constructed by sampling theory whose estimates based on data at hand plus hypothetical repeated sampling in the future with similar data. Regardless of these potential debates, we will compare the performances of the method proposed in this paper with two famous unit root tests: the Dickey-Fuller (DF) test and the Phillips-Perron (PP) test. One should bear in mind that no matter how the comparison results turn out, the method proposed in this paper provided a solution to performing unit root tests under model specification uncertainty, which cannot be achieved by other traditional test methods since they all require some kind of model restrictions.

In the simulation, we adopted a "autoregressive signal plus white noise" model from Pagano (1974) as the underlying data generating process. This approach will not

only generate autoregressive time series but also provides a convenient way to control the signal to noise ratio of the generated series, which is helpful to evaluate the performance of different test methods under various data circumstances, e.g., very clean data or noisier data. The fundamental data generating process can be described as:

$$x_t = s_t + \eta_t, \text{ and } \sum_{j=0}^{q} \beta_j s_{t-j} = \varepsilon_t \tag{37}$$

Restrictions on this system are:

(1). s_t and η_t are independent;

(2). η_t is independent identically distributed as $N(0, \sigma_\eta^2)$, and $\varepsilon_t \sim iid\ N(0, \sigma_\varepsilon^2)$;

(3). β_0=1, $\beta_q \neq 0$

According to this setting, s_t is actually a AR(q) process guaranteed by condition (3), and the final data process x_t is composed of signal series s_t and noise component η_t. Under these assumptions, the signal to noise ratio is defined as:

$$r = \sigma_\varepsilon^2 / \sigma_\eta^2 \tag{38}$$

Furthermore, researchers often face a question about whether the series should have drift and/or a deterministic trend term. If we consider this then two terms could be added to data generating process. For example, trend can be incorporated as

$$s_t = \beta_0 + \sum_{j=1}^{q} \beta_j s_{t-j} + \varepsilon_t \tag{39}$$

with all other assumptions unchanged, and a drift term can be specified by letting some initial value of $s_0 = \alpha$. This process can also be found in Harvey and Harvey (1993, 2:122).

In order to incorporate the model uncertainty problem in the data generating process, we adopt the following mechanism:

(1). Generate two binomial distributed random variables with probability 0.5 to decide whether the DGP has drift or deterministic trend term;

(2). Generate an uniformly distributed random variable ranging from 1 to 6 to set number of lags included in the DGP; i.e., the process might be randomly AR(1) to AR(6). Notice that any degree of lags could be included, what we did here is for discussion convenience.

Meanwhile, to completely measure and judge the performance of proposed method, numerous data sets were generated with different properties and used to test and compare the result. The dominant roots of the series were set to be 0.80, 0.85, 0.9, 0.95, 0.96, 0.97, 0.98, 0.99 and 1 since more samples and results need to be collected near unit root to achieve a meticulous performance comparison. The signal to noise (SN) ratio was regulated to be 0.25, 0.5, 1, 5, 10, 50 and 100 in order to measure the capability of testing under various data circumstances. With each of these settings, data sets were generated in four different sample sizes: 50, 100, 300 and 500 so that both small sample properties and the asymptotical behavior of these methods can be viewed. Overall, 252 different combinations of DGP parameters were analyzed with either 200 or 500 data sets generated in order to test performance for each of the 252 settings. For dominant root equals to 1, 0.99, 0.98, 0.97, and 0.96, 500 data sets are generated for each value to evaluate performance more precisely and reliable; while for dominant root equals to 0.95, 0.90, 0.85, 0.80, 200 replications are used for computational convenience. All the detailed DGP information can be found in appendix A. Also, in order to remove the

initial condition effect, the first several generated observations (the specific number of observations removed depending on sample size) were eliminated (Schwert 1989).

Since in the DGP, trends and number of lags are randomly assigned, each simulation data set has an unknown true model specification. For our proposed method, 24 models will be averaged together to draw a conclusion on each data set:

(1). $y_t = \sum_{i=1}^{P} \rho_i y_{t-i} + \xi_t,\ p = 1, 2, ..., 6$

(2). $y_t = \mu + \beta t + \sum_{i=1}^{P} \rho_i y_{t-i} + \xi_t,\ p = 1, 2, ..., 6$

(3). $\log y_t = \sum_{i=1}^{P} \rho_i \log y_{t-i} + \xi_t,\ p = 1, 2, ..., 6$

(4). $\log y_t = \mu + \beta t + \sum_{i=1}^{P} \rho_i \log y_{t-i} + \xi_t,\ p = 1, 2, ..., 6$

The last 12 models employ a logarithm form to incorporate variable transformation, and one can always add other transformation formats into this frame. Here we only consider log transformation since it is not only one case of Box-Cox transformation which is commonly used in economic literature, but also helpful to demonstrate our method without introducing too much computational complexity.

2.3.2 Monte Carlo Simulation Results

Eight tests are performed to each data set generated according to the rules in previous section: the test method proposed in this paper by averaging 24 models, 6 Augmented Dickey-Fuller tests with different lag specification 1 to 6, and the Phillips-Perron test. For the ADF and the PP tests, the decision rule is do not reject the null hypothesis (nonstationary) if the p-value is larger or equal than 10%, which is a pretty loose standard. The results are compared and discussed below.

Figure 2. 1 displays the percentage of unit roots detected in 500 replicated sample simulations by each testing methods under different signal to noise ratios, which is connected to the idea of test size. The data generating process is I (1) series. First notice that the BMA test did fairly well in all these four samples especially when the sample size is small. Among all four panels, the BMA curve almost always stays the top of all others which indicates a strong ability to detect the existence of a unit root. When the sample size is small, BMA also keeps this consistency which shows good small sample property. This can also be demonstrated by viewing plots when sample size is 50 and 100, where the BMA method detects most of the unit roots, especially when the signal to noise ratio is small. This also shows good ability of handling messy data when the underlying process is nonstationary.

For our simulation data generating process (37), with a little transformation one can get the following formula:

$$\begin{aligned} & x_t = s_t + \eta_t,\ \sum_{j=0}^{q} \beta_j s_{t-j} = \varepsilon_t, x_{t-1} = s_{t-1} + \eta_{t-1} \\ & \Rightarrow x_t = x_{t-1} - \sum_{j=2}^{q} \beta_j s_{t-j} + \varepsilon_t + \eta_t - \eta_{t-1} \end{aligned} \tag{40}$$

The last part introduces moving average (MA) component to the data series. According to Schwert (1989), the ADF and PP tests will have severe size distortion when testing data series that have a large negative moving average root, and in his paper, simulation shows that when the dominant of MA root is -0.8, the size is almost 100%. In our simulation formula (40), the MA root is -1, which directly explains the bad performance of ADF and PP test. Although overparameterizaiton in the ADF test could mitigate this problem on some level, it does not seem to have a significant impact and will bring problems to the

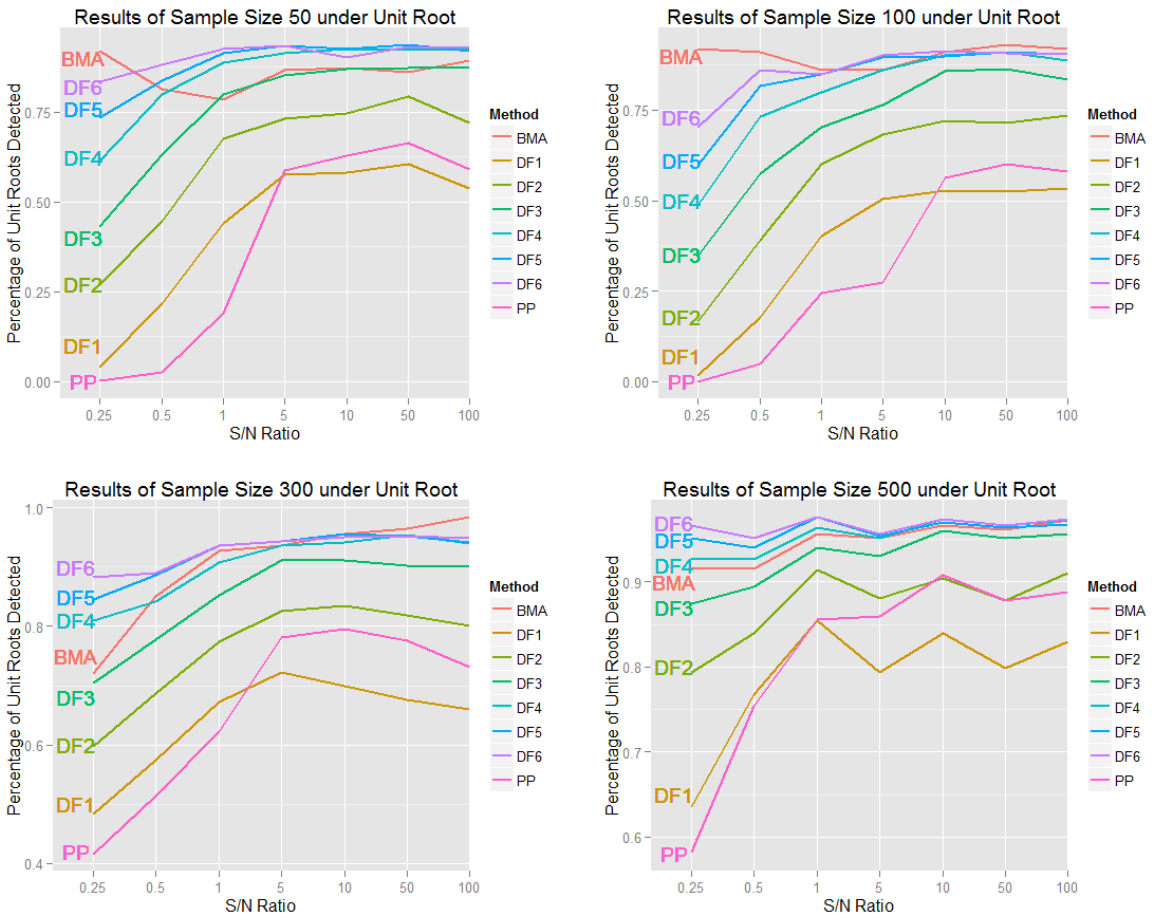


Figure 2.1 Percentage of Unit Root Detected with Different SN Ratio and Sample Size when the DGP is I (1) Process

test power which will be introduced in next few paragraphs. Furthermore, Schwert (1989) also indicates that the size of the PP test will suffer more in this situation and do not come close to their asymptotic distribution even when the sample size is fairly large. This shows in Figure 2. 1 since the PP test generally stays at the bottom of all curves showing quite unsatisfactory performance and does not move up as the sample size increases. On the contrary, the BMA approach can handle this problem well judging from the curves in Figure 2. 1, and behaves pretty consistent under all circumstances. Next we see that the BMA method keeps this good performance when the DGP is stationary compared to other methods.

The second part of the simulation measures the performance when the underlying data generating process is stationary, i.e., I (0) process. Notice that in this case it is also possible for the generated data to have deterministic trend since it is randomly assigned, and the specific functional form is still unknown to each test method. This measurement is similar to test power in frequentist statistics and traditionally a power curve could be drawn. However, the method we derive belongs to the Bayesian framework where debate still exists about using traditional frequentist performance judgments for performance evaluation. Meanwhile, since we have many parameters combinations in the simulation, 28 graphs are needed to fully show all the curves under different circumstances which is redundant and unclear. Instead of drawing the power curve graphs, an optimal method is chosen among all 8 tests for each combination of parameter settings and shown as a form of heat map. Figure 2.2 lists such graphs for the 4 different sample sizes. For each subgraph, the vertical axis stands for different dominant roots of the data generating process while the horizontal axis shows the various signal to noise ratios. Each color

demonstrates the testing method which achieved the optimal result (rejecting the unit root null hypothesis at the highest rate, since the underlying data is stationary). First notice is that among the 8 testing methods only 3 were proven to be best in different situations: the BMA test proposed in this paper, the ADF test with lag truncated at 1 and the PP test. The ADF test with lag 1 dominated over all other ADF tests with different lag specification in all four sample sizes. Several previous researches have shown that the power of the ADF test will suffer if the lag length is chosen too large (Harris 1992; Ng and Perron 1995) which could explain this dominance. Meanwhile, too small lag length may bias the test since there are remaining serial correlation in the errors, which provides insight into the better performance of the PP test compared to the ADF test since it is a nonparameteric approach which adjusts any serial correlation and heteroskedasticity in the errors by modifying the test statistics. However, the most significant result is that in each graph, the BMA test proposed in this paper occupies a larger area which indicates better performance in lots of the situations. The model specification uncertainty of the underlying data plays important role here, but in real unit root testing, the model specification is uncertain. Roughly 50% of the testing data actually contains a deterministic trend but when is unknown to the testing methods. It is well known that the power of the ADF and the PP tests is substantially reduced if the time trend term is inappropriately excluded from the testing model (Campbell and Perron 1991). On the contrary, the better performance of the BMA in general showed a good ability to handle unit root tests when model uncertainty is present. Meanwhile, notice that the advantage of the BMA test becomes more obvious as sample size increases since more information is provided by a larger sample. The better performance of the BMA test in larger sample

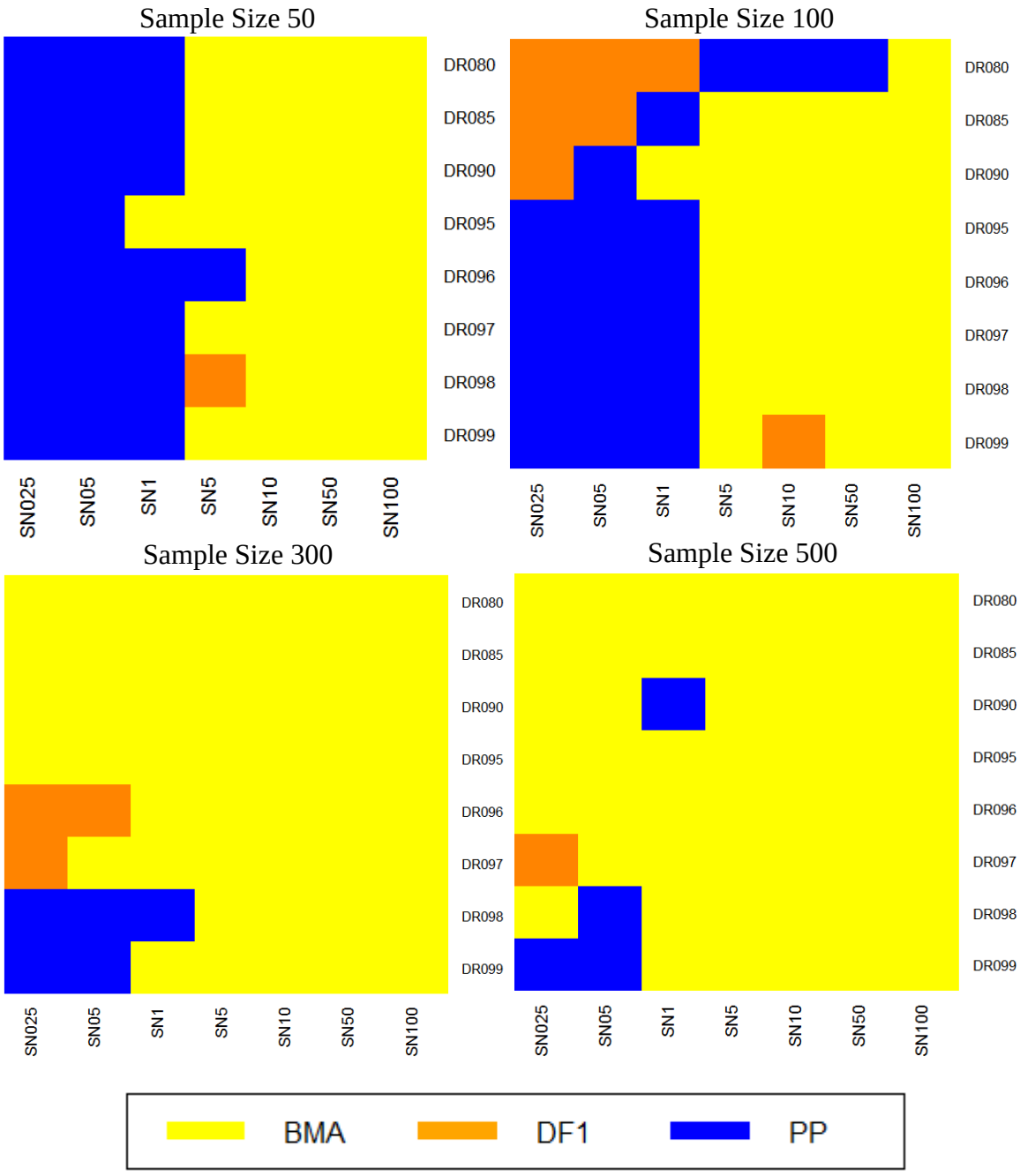


Figure 2.2 Optimal Testing Method with Different Sample Size and Parameter Settings when Underlying DGP is Stationary

sizes is surprising since ADF and PP should have larger power under large sample sizes because they are based on asymptotical properties.

Secondly, examine each individual graph in more detail. From the view of the magnitude of dominant root, the BMA method works extremely well when the dominant root is close to 1, especially when the signal to noise ratio is larger than 1. The ADF and PP tests are known to have low power when testing series with a large autoregressive root (DeJong et al. 1992), which could be caused by the discontinuity in the asymptotic distribution near the unit root; while Bayesian framework can handle this problem perfectly by introducing priors on the parameters (Sims 1988).

Also notice that in the simulation under all circumstances, the prior used for dominant root is always the same highly left skewed Beta distribution with quite large density around 1; so in practice if one have some belief of stationary on the data he analyzes he can assign prior less skewed with more density to the stationary region, then the result of BMA test should gain more power. Nevertheless, with the consistent left skewed prior, the result is strong enough here to illustrate the advantages of BMA test approach when the dominant root is close to 1 and this advantage increases when sample size increases.

It seems that the BMA approach works better when the signal to noise ratio is relatively high since one can observe that most dominance happens when this ratio is larger or equal to 1. One possible reason is that when the signal to noise ratio is quite low, like 0.25 and 0.5, according to the data generating process defined in (37), more than half of the data would be noise information which will cancel the real underlying AR process and in this case the highly skewed prior used in BMA approach may dominate the

posterior density which will produce more unit root conclusions that will lower the "power" of the test. As mentioned before, researchers usually should have some beliefs on whether their data is stationary, especially when the real dominant root is far away from unity, and thus assign more informative and reasonable priors to the parameters which will bring better results. Though increasing the prior variance (flattening the prior) to include more possible candidates can be helpful in this situation, more cost could happen by doing so (like lower the acceptance rate and increasing the computational burden) and it is beyond the discussion scope of this research.

2.4 Empirical Application to the Nelson-Plosser Data

Nelson and Plosser (1982) investigated the stochastic behavior of 14 historical U.S. macroeconomic time series and concluded that all of them are non-stationary stochastic processes with no tendency to return to a trend line. Since then the debate about the behavior of these 14 series has been continuing and the data they used in the analysis has become benchmark data for unit root test methods. Therefore, we apply the methodology derived in this paper to evaluate the empirical performance of our method, and compare the results with that from Nelson and Plosser (1982) as well as other three famous paper which also analyzed the same data using Bayesian unit root test methods, Phillips (1991), DeJong and Whiteman (1991) and Lubrano (1995). For each of the 14 series, we applied our method by averaging 24 models as we did in the simulation: models with and without deterministic trend, with original levels and logarithms of the data, and for each of the above settings with autoregressive lags from 1 to 6. The number of lags included in the formula and the existence of a deterministic trend are evidences of

model uncertainty and even the three papers listed here have different opinions: Nelson and Plosser (1982) chose one lag value for each series from 1 to 6 suggested by autocorrelations of first differences and by the partial autocorrelations of the deviations from the trend, and they included time trend in all the series and test them. DeJong and Whiteman (1991) chose 3 lags for all the series and considered the trend in the series. Phillips (1991) considered both 1 lag and 3 lags together with trend, and he also compared two different priors to confront his main goal of the paper. Lubrano (1995) using 2 lags for all series except lag 4 for unemployment rate, consumer prices and bond yield which has chosen according to the computed Schwarz criterion. Our method will not have this inconsistency since it is designed to handle the issue of model specification uncertainty, and by averaging 24 models together, all the situations are considered at one time to achieve a comprehensive conclusion.

The model parameterization and priors assigned are identical to those in the previous simulation. The acceptance rate is controlled to around 40% for univariate draws and 25%~35% for multivariate situations to ensure the good mixture of the Markov chain. Meanwhile, 11,000 draws are taken for each series and the first 1,000 draws are eliminated in order to reduce the initial condition effects as well as to achieve good convergence.

Figure 2.3 (i)-(xiv) displays the robust model averaged posterior densities of the dominant root for each of the Nelson& Plosser series. One may find the shape of all the posterior distribution are quite different from some other work using the similar data. The major reason is that instead of showing the density of only one model specification, the posterior distribution of the dominant root showed here is the average result of all 24

models so some part might be smoothed by different model result but most of them are still far away from any standard distribution density. Each density has some probability that the root exceeds 1 but the majority mass will decide whether the series has stochastic trend. Table 2.1 gives the results of averaged posterior probabilities of nonstationarity $P(\phi_1 \geq 1)$, and the 95% credible interval for the deterministic trend coefficient. To compare results, we also list the test statistics or probability from the aforementioned four papers. The first two columns $\tau(\widehat{\rho}_1)$ and $t(\hat{\gamma}_1)$ are from Nelson and Plosser (1982) Table 5 which are the test statistics for the root and trend term respectively. The following two columns are part of Table IV in Phillips (1991), who used the reduced form parameterization and the Jeffreys prior, and here both the results of AR(1) and AR(3) are listed. The results from DeJong and Whiteman (1991) are shown in column 5; in their paper they employed a truncated flat prior on the autoregressive coefficients but perform inference based on the root Λ. The next column provides the results from Table 1 in Lubrano (1995) who adopted a Beta prior and employed a structural parameterization including a special treatment of the first observation which we followed in this paper. The last two columns show result of our modeling averaging approach. Comparing all these results, one will find some different conclusions although most of them are broadly consistent with each other. These differences show that time series specification can have a significant influence on the posterior distribution as well as the unit root decision. Nelson and Plosser (1982) claimed that except for the marginal stationarity of unemployment rate, the other 13 macroeconomic series are difference stationary, i.e., contain a unit root. Meanwhile none of them has a significant deterministic trend.

In general, the results of all five studies are qualitatively consistent in spite of some discrepancies on certain series, as well as differences in probability levels estimated which is reasonable because of the different priors, methods and specifications adopted. Our BMA approach finds 6 series contains stochastic trend: Nominal GNP, Consumer prices, Nominal wages, Velocity, Interest rate (Bond yield) and Common stock prices. Lubrano (1995) found two more besides the previous 6: GNP deflator and real wages. In Phillips (1991), 5 of 14 series show evidence of stochastic trends (if we adopt 0.15 as significant level mentioned in his paper) under the Jeffreys prior with AR(3) plus trend specification (except for unemployment AR(4) was used): Industrial production, Consumer prices, Velocity, Bond yields and Stock prices. Compared to these results, DeJong and Whiteman (1991) believed all series are trend stationary except two: Velocity and Bond yields, with marginal nonstationary in the case of Consumer prices. This is criticized by Phillips (1991) due to the fragile flat priors which are always biased away from stochastic trend alternatives that made the result quite different from others.

In spite of mostly consistent results between our result and the two other Bayesian papers (excepting DeJong and Whiteman (1991)), there are differences concerning whether some series contains unit root. One of the differences is on industrial production, which is considered to have a unit root in Phillips (1991) when using Jeffreys prior on AR(3) with trend specification but stationary in our method as well as Lubrano (1995). Another one is wages which is considered stationary in Phillips (1991) but strictly nonstationary in our paper and Lubrano (1995). One possible reason for this kind of disagreement could be model uncertainty, i.e., the uncertainty over the number of lags included in the formula. For example, for wages, Phillips using both AR(1) and AR(3)

with trend, and concluded opposite results – with AR(1), he detected the existence of unit root but with AR (3) indicated stationarity. In contrast, Lubrano, using AR(2), concluded nonstationarity, which is the same result as our conclusion. We examined the component and proportion of our final comprehensive posterior probability, and found that for testing this wages series, overall the models using logarithms of the data with a trend specification (6 models, with from 1 to 6 lags) had almost 97% of the model posterior weight. Further, we when dig into the 6 models using logarithm data with trend and we compute the weights of each model with different lags; we find: AR(1), 23.2%; AR(2), 24.7%; AR(3), 15.2%; AR(4), 14.3%; AR(5), 11.0%; AR(6), 11.6%. It is clear that specifications of AR(1) and AR(3) take the majority of the support and results conditional on those two models, AR(1) to AR(3), are that a unit root does exist in the series. Given that these three specifications gather about 63.1% of the support among all 24 models, it is not surprising that the model-averaged results support the nonstationarity conclusion. Meanwhile, it is sometimes hard to decide the appropriate number of lags included in the model according to criteria like AIC or BIC since their values may not vary significant to strongly support one lag specification. Table 2.2 lists AIC and BIC value for each model specification of all the 14 series. Take common stock as an example, the AIC values are quite flat over all the 6 model specifications which may not provide enough evidence to fully support one model; meanwhile, for this series AIC and BIC will give different lag specification result if we choose lag numbers based on the minimum value. These all demonstrate the importance of accounting for model specification uncertainty issues in the empirical work, and also proves the effectiveness of our Bayesian model averaging approach in handling such situations.

2.5 Conclusion

Model uncertainty sometimes can be a significant issue in empirical economic research, which without a doubt, includes unit root tests in time series analysis. Although a tremendous literature has been devoted to unit root testing, both in the frequentist domain and the Bayesian framework, not much work has been done to propose a general approach to testing for unit roots under model uncertainty. This paper tries to contribute to filling this gap by introducing an approach which can test for unit roots by averaging all potential model candidates in a Bayesian framework. These model candidates can vary in different ways. For example, for autoregressive models, they may differ in the number of lags included in the model since lag length choice can have a major impact on unit root inference. It can also be uncertainty about the format of the response variable, like original level data or some function of the data since one might believe the underlying data generating process is transformed, or simply it is easier to interpret if working on some transformed data, like natural logs. Finally, it is often difficult to accurately decide the distribution of the stochastic term included in the model. We can simply average a set of potential assumptions to mitigate this uncertainty. This is not explicitly done in this paper, but it is a natural generalization of our method and will be discussed in the next chapter.

Provided the general approach developed in the first several subsection of this paper, a Monte Carlo simulation was performed to examine our method. By randomly assigning the number of lags, the existence of drift, and the inclusion of a deterministic trend, we generated the testing data sets have a specification unknown to the unit root test methods. Then 24 possible models were averaged to achieve the final conclusion: models

with/without deterministic trend, and using original data and log-transformed data, and for each of these four settings, models with from 1 to 6 autoregressive lags were considered. After analyzing the results and comparing them with two existing popular methods, the Augmented Dickey-Fuller Test and the Phillips-Perron test, we conclude that the proposed BMA method performed fairly well. In testing the nonstationary series, it has relatively smaller "size" (if we use terminology from frequentist statistics) compared to the PP test and some of the ADF tests with short lag length specifications. For stationary data sets, we found that BMA test proposed in this paper had higher power in a large percentage of scenarios tested compared to the other tests which indicates better performance in lots of the situations. This can be explained by better handling the model uncertainty of the underlying data generating process, since traditional test methods power will suffer under incorrect specification. Meanwhile, it is also noticed that the BMA method works extremely better when the sample size are large, as well as when the signal to noise ratio is relatively large (usually larger than one).

Following the simulation, we also applied our proposed approach to the famous Nelson and Plosser (1982) data of 14 macroeconomic time series, and compared the results with those from four other papers which also tested the same data using different methods: the original paper which used the Augmented Dickey-Fuller test, and other three papers which adopted various Bayesian unit root test frameworks. Generally speaking, we achieved similar decisions on most of the series with Phillips (1991) and Lubrano (1995), which indicates the empirically suitability, as well as the parameterization and prior validity of our method. We carefully examined the posterior odds of 24 models in the set of considered models and found that most posterior support

belonged to 3 models with autoregressive lags 4 to 6; whereas the two papers which achieved different result with ours happened to only use 3 or 4 lags. This again not only shows the robustness of our method but also proved the importance of considering model uncertainty when conducting unit root tests and the ability of the proposed method to successfully address this issue.

This paper developed a numerical Bayesian unit root test method which can incorporate the model specification uncertainty issues in empirical research work. By both simulation and analysis to real world data set, we showed the effectiveness of the proposed method. Though the method is not uniformly dominant, it works well in many circumstances and there are many future directions to improve the performance of the method, like introducing the structural breaks. As Lubrano (1995) said “intuitively it is very hazardous to discriminate between competing economic theories”, or Phillips (1991) indicated “we see no reason why empirical researchers should not judiciously pursue this approach (Bayesian unit root test, author’s note) as well as classical methods”. The most important thing is that it is a general approach to considering model uncertainty when performing unit root tests, and provides an alternative to researchers who are concerned this is a significant issue in their work.

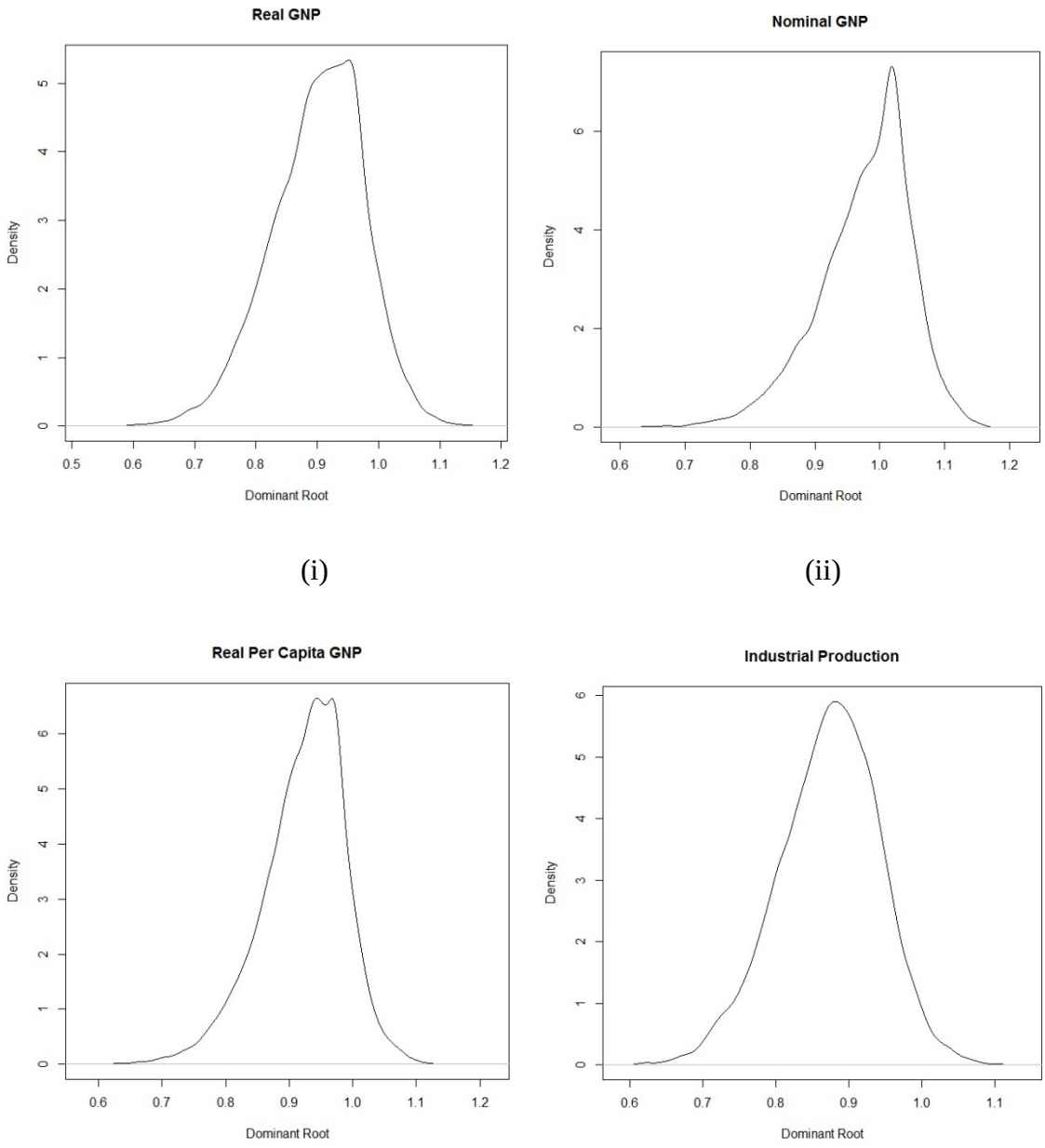
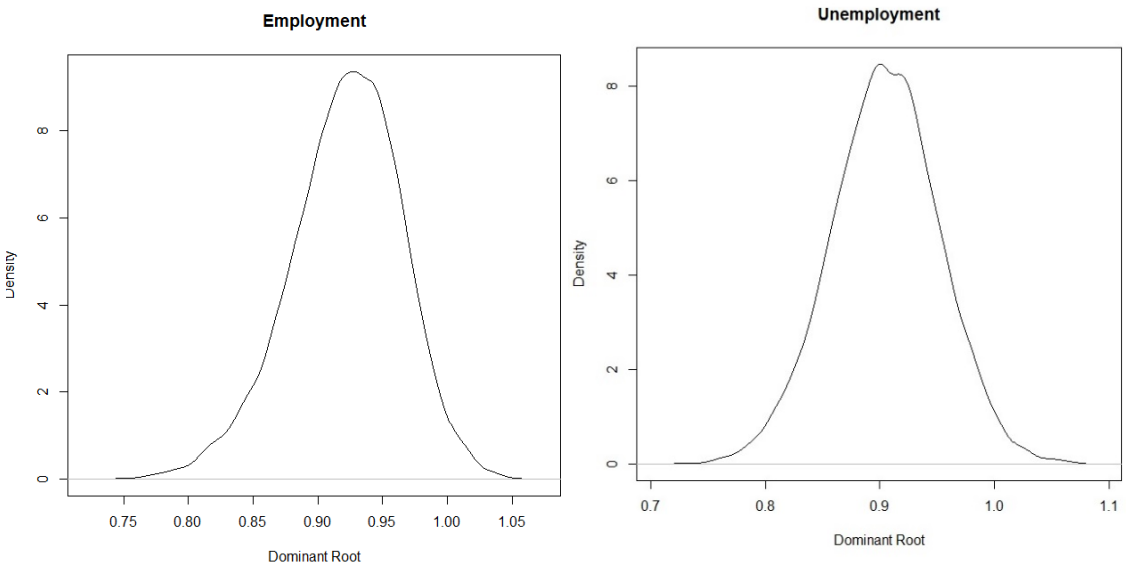
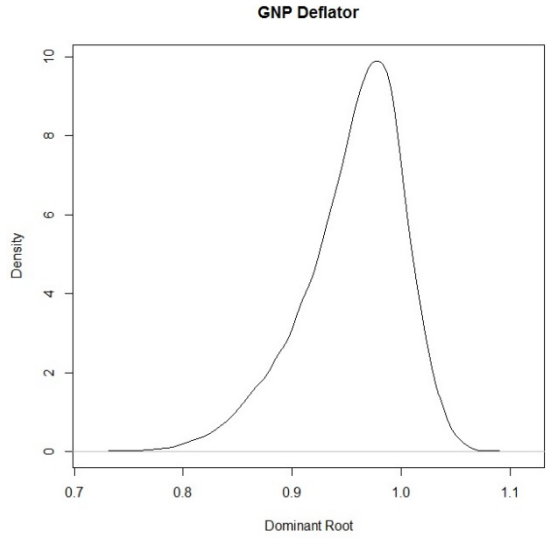


Figure 2.3 Averaged Posterior Density Plot of Dominant Root

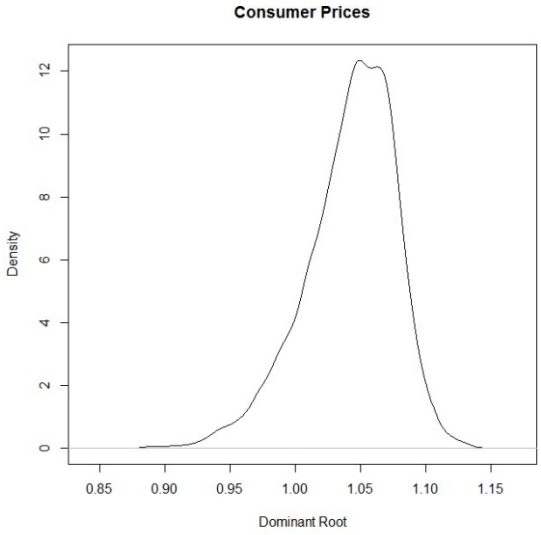


(v)

(vi)

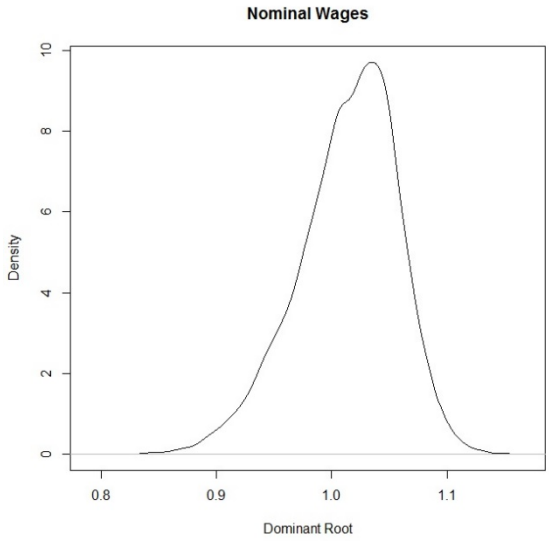


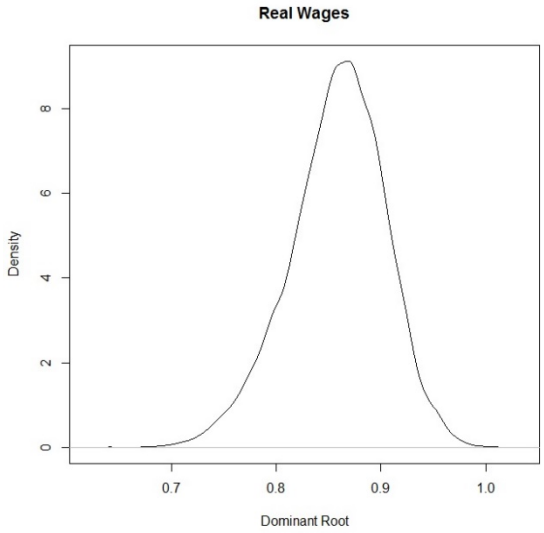
(vii)



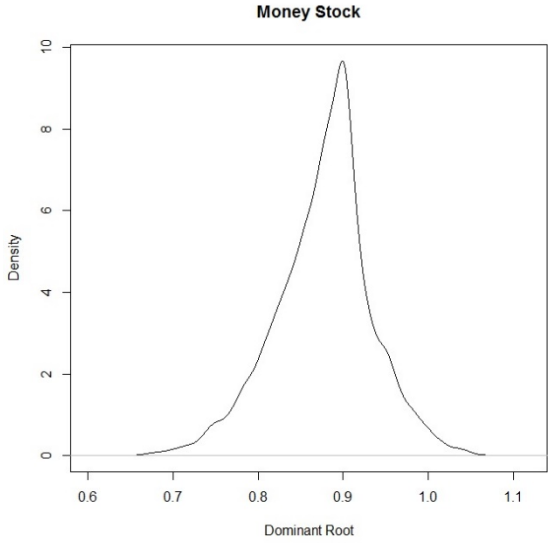
(viii)

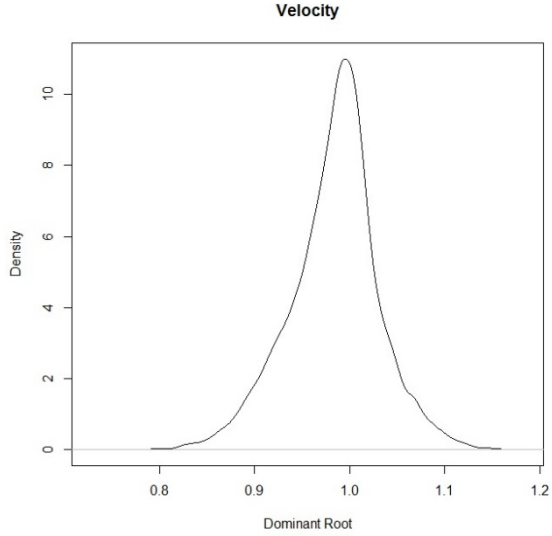
Figure 2.3 *continued*. Averaged Posterior Density Plot of Dominant Root





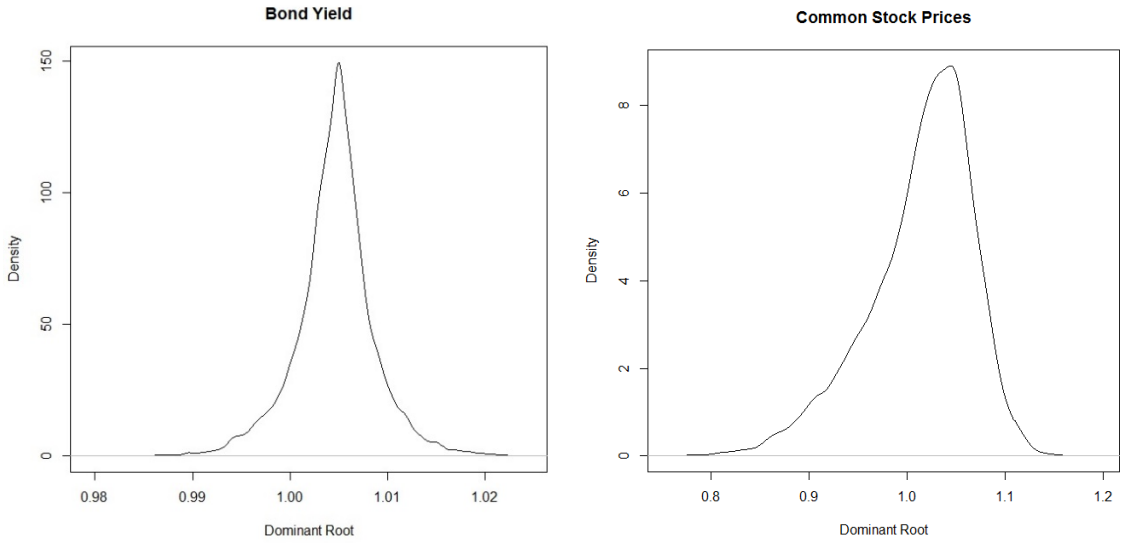
(ix) (x)





(xi) (xii)

Figure 2.3 *continued*. Averaged Posterior Density Plot of Dominant Root



(xiii) (xiv)

Figure 2.3 *continued*. Averaged Posterior Density Plot of Dominant Root

Table 2.1 Testing Result of Nelson-Plosser Data[a]

Series	Nelson Plosser (1982)[b]		Phillips (1991)		DeJong Whiteman (1991)	Lubrano (1995)	BMA Unit Root Test	
			AR1+trend	AR3+trend	$P_{DJW}(\Lambda \geq 0.975)$	$P(\rho \geq 1 \mid y)$		
	$\tau(\hat{\rho}_1)$	$t(\hat{\gamma})$	$P_J(\rho \geq 1)$	$P_J(\rho \geq 1)$			$P(\phi_1 \geq 1)$	95% credible interval of deterministic trend
Real GNP	**-2.99**	3.03	**0.193**	0.012	0.003	0.031	0.0318	(0.0845, 1.360)
Nominal GNP	**-2.32**	2.34	**0.361**	0.074	0.02	**0.149**	**0.582**	(-0.728, 4.388)
Real per capita GNP	**-3.04**	3.01	**0.163**	0.01	0.003	0.029	0.274	(0.140, 1.629)
Industrial production	**-2.53**	2.44	0.124	**0.188**	0.001	0.015	0.0915	(-0.0102,0.418)
Employment	**-2.66**	2.54	**0.190**	0.04	0.004	0.012	0.0553	(-0.0816, 0.941)
Unemployment rate	-3.55	-0.23	0.126	0.086	0.002	0.001	0.0475	(0.143, 0.576)
GNP deflator	**-2.52**	2.65	**0.162**	0.02	0.01	**0.267**	0.0564	(0.0225, 1.001)
Consumer prices	**-1.97**	2.84	**0.601**	**0.176**	0.196	**0.231**	**0.615**	(-0.0437, 1.332)
Wages	**-2.09**	2.30	**0.319**	0.045	0.018	**0.104**	**0.537**	(0.0238, 2.181)
Real wages	**-3.04**	3.14	0.103	0.014	0.003	**0.225**	0.0570	(0.0386, 0.944)
Money stock	**-3.08**	3.03	**0.315**	0.008	0.005	0.014	0.0520	(0.250, 2.377)
Velocity	**-1.66**	-0.65	**0.353**	**0.537**	**0.592**	**0.486**	**0.756**	(-0.742, 1.576)
Bond Yield	**0.686**	1.75	**0.999**	**0.996**	**0.617**	**0.208**	**0.944**	(-0.00110, 0.000477)
Common stock prices	**-2.05**	2.37	**0.301**	**0.215**	0.04	**0.214**	**0.785**	(-0.0166, 0.586)

a.The bold font stands for decision of nonstationary in each paper, since different methods and criteria are used the meaning of specific probability number is not consistent and comparable in general.

b.For Nelson and Plosser (1982) only Unemployment was detect to be marginally stationary.

Table 2.2 AIC/ BIC Value of Different Specifications of 14 Series

Series		AR (1)	AR (2)	AR (3)	AR (4)	AR (5)	AR (6)
Real GNP	AIC	-156.15	-157.84	-152.46	-149.04	-144.42	-140.37
	BIC	-149.82	-149.46	-142.08	-136.68	-130.11	-124.17
Nominal GNP	AIC	-105.13	-113.07	-109.30	-105.77	-102.78	-99.25
	BIC	-98.80	-104.69	-98.91	-93.41	-88.47	-83.05
Real Per Capita GNP	AIC	-154.79	-156.57	-151.09	-147.19	-142.49	-138.73
	BIC	-148.45	-148.19	-140.70	-134.83	-128.19	-122.52
Industrial Production	AIC	-187.31	-183.02	-180.01	-175.45	-172.81	-177.36
	BIC	-179.21	-172.25	-166.60	-159.41	-154.16	-156.13
Employment	AIC	-290.62	-292.51	-288.45	-291.23	-288.63	-286.42
	BIC	-283.48	-283.03	-276.67	-277.17	-272.32	-267.88
Unemployment Rate	AIC	107.01	103.78	101.01	85.81	87.69	88.82
	BIC	114.15	113.26	112.79	99.87	104.00	107.36
GNP Deflator	AIC	-244.07	-254.61	-248.63	-243.19	-238.03	-236.67
	BIC	-236.89	-245.08	-236.79	-229.05	-221.62	-218.02
Consumer Prices	AIC	-312.00	-349.69	-355.98	-359.16	-366.21	-376.11
	BIC	-303.89	-338.93	-342.57	-343.12	-347.56	-354.88
Wages	AIC	-172.70	-183.82	-179.58	-173.96	-169.20	-163.66
	BIC	-165.95	-174.88	-168.48	-160.73	-153.87	-146.26
Real Wages	AIC	-258.27	-254.68	-248.36	-242.48	-236.92	-232.63
	BIC	-251.52	-245.75	-237.26	-229.26	-221.59	-215.23
Money Stock	AIC	-216.10	-249.74	-245.87	-240.22	-242.46	-236.37
	BIC	-208.92	-240.21	-234.02	-226.08	-226.05	-217.72
Velocity	AIC	-255.60	-251.35	-247.00	-248.39	-244.84	-240.99
	BIC	-247.75	-240.93	-234.03	-232.88	-226.82	-220.47
Bond Yield	AIC	-172.49	-168.01	-165.35	-161.63	-156.47	-152.58
	BIC	-165.75	-159.07	-154.25	-148.40	-141.14	-135.18
Common Stock Prices	AIC	-75.51	-76.75	-76.53	-72.75	-73.09	-73.40
	BIC	-67.73	-66.41	-63.66	-57.37	-55.22	-53.06

References

Avramov, Doron. 2002. “Stock Return Predictability and Model Uncertainty.” *Journal of Financial Economics* 64 (3): 423–58.

Barnard, G. A. 1963. “New Methods of Quality Control.” *Journal of the Royal Statistical Society. Series A (General)* 126 (2): 255–58. doi:10.2307/2982365.

Bauwens, Luc, Michel Lubrano, and Jean-Francois Richard. 2000a. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press. http://books.google.com/books?hl=en&lr=&id=hRV3XiQVWLUC&oi=fnd&pg=PA1&dq=Bayesian+inference+in+dynamic+econometric+models&ots=mJMR-mN8hn&sig=Yq1euDu-iE8DAzNeoBWaA0DIsAA.

———. 2000b. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press.

Berger, James O., and Ruo-yong Yang. 1994. “Noninformative Priors and Bayesian Testing for the AR(1) Model.” *Econometric Theory* 10 (3/4): 461–82.

Bhargava, Alok. 1986. “On the Theory of Testing for Unit Roots in Observed Time Series.” *Review of Economic Studies* 53 (3): 369.

Bollerslev, Tim. 1986. “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics* 31 (3): 307–27.

Boswijk, H. Peter. 2001. *Testing for a Unit Root with near-Integrated Volatility*. Tinbergen Institute Discussion Paper. http://www.econstor.eu/handle/10419/85890.

Boswijk, H Peter, and Franc Klaassen. 2012. “Why Frequency Matters for Unit Root Testing in Financial Time Series.” *Journal of Business & Economic Statistics* 30 (3): 351–57.

Box, George EP, and David R Cox. 1964. “An Analysis of Transformations.” *Journal of the Royal Statistical Society. Series B (Methodological)*, 211–52.

Brock, William A., Steven N. Durlauf, and Kenneth D. West. 2007. “Model Uncertainty and Policy Evaluation: Some Theory and Empirics.” *Journal of Econometrics* 136 (2): 629–64.

Campbell, John Y., and Pierre Perron. 1991. “Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots.” In *NBER Macroeconomics Annual 1991, Volume 6*, 141–220. MIT Press. http://www.nber.org/chapters/c10983.pdf.

Chambers, Marcus J. 3. “Testing for Unit Roots with Flow Data and Varying Sampling Frequency.” *Journal of Econometrics* 119 (1): 1–18. doi:10.1016/S0304-4076(03)00152-0.

Charitidou, Efstratia, Dimitris Fouskakis, and Ioannis Ntzoufras. 2013. “Bayesian Transformation Family Selection: Moving towards a Transformed Gaussian Universe.” *arXiv Preprint arXiv:1312.3482*. http://arxiv.org/abs/1312.3482.

Chatfield, Chris. 1995. “Model Uncertainty, Data Mining and Statistical Inference.” *Jorunal of the Royal Statistical Society* 158 (3): 419–66.

Choi, In. 1992. “Effects of Data Aggregation on the Power of Tests for a Unit Root: A Simulation Study.” *Economics Letters* 40 (4): 397–401.

Choi, In, and Bhum Suk Chung. 1995. “Sampling Frequency and the Power of Tests for a Unit Root: A Simulation Study.” *Economics Letters* 49 (2): 131–36.

Claeskens, Gerda, and Nils Lid Hjort. 2008. *Model Selection and Model Averaging*. 1 edition. Cambridge ; New York: Cambridge University Press.

Clemen, Robert T. 1989. “Combining Forecasts: A Review and Annotated Bibliography.” *International Journal of Forecasting* 5 (4): 559–83. doi:10.1016/0169-2070(89)90012-5.

Clyde, Merlise, and Edward I. George. 2004. “Model Uncertainty.” *Statistical Science* 19 (1): 81–94.

Cochrane, John H. 1988. “How Big Is the Random Walk in GNP?” *Journal of Political Economy* 96 (5): 893–920.

Cremers, K. J. Martijn. 2002. “Stock Return Predictability: A Bayesian Model Selection Perspective.” *Review of Financial Studies* 15 (4): 1223–49. doi:10.1093/rfs/15.4.1223.

DeJong, David N., John C. Nankervis, N. E. Savin, and Charles H. Whiteman. 1992. “The Power Problems of Unit Root Test in Time Series with Autoregressive Errors.” *Journal of Econometrics* 53 (1–3): 323–43. doi:10.1016/0304-4076(92)90090-E.

DeJong, David N., and Charles H. Whiteman. 1991. “Reconsidering ‘Trends and Random Walks in Macroeconomic Time Series.’” *Journal of Monetary Economics* 28 (2): 221–54.

Dickey, David A., and Wayne A. Fuller. 1979a. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association* 74 (366a): 427–31. doi:10.1080/01621459.1979.10482531.

———. 1979b. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association* 74 (366a): 427–31. doi:10.1080/01621459.1979.10482531.

———. 1981. “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root.” *Econometrica* 49 (4): 1057–72. doi:10.2307/1912517.

Dorfman, Jeffrey H. 1993. “Bayesian Efficiency Tests for Commodity Futures Markets.” *American Journal of Agricultural Economics* 75 (5): 1206–10.

Dorfman, Jeffrey H., and Myung D. Park. 2011. “Estimating the Risk-Return Tradeoff in Agribusiness Stocks: Linkages with the Broader Stock Market.” *American Journal of Agricultural Economics* 93 (2): 426–33. doi:10.1093/ajae/aaq158.

Dufour, Jean-Marie, and Maxwell L. King. 1991. “Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary AR (1) Errors.” *Journal of Econometrics* 47 (1): 115–43.

Eicher, Theo S., Christian Henn, and Chris Papageorgiou. 2012. “Trade Creation and Diversion Revisited: Accounting for Model Uncertainty and Natural Trading Partner Effects.” *Journal of Applied Econometrics* 27 (2): 296–321. doi:10.1002/jae.1198.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. 1992. *Efficient Tests for an Autoregressive Unit Root*. National Bureau of Economic Research Cambridge, Mass., USA. http://www.nber.org/papers/t0130.

Engle, Robert F. 1982. “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.” *Econometrica* 50 (4): 987–1007. doi:10.2307/1912773.

Engle, Robert F., and Victor K. Ng. 1993. “Measuring and Testing the Impact of News on Volatility.” *The Journal of Finance* 48 (5): 1749–78. doi:10.1111/j.1540-6261.1993.tb05127.x.

Fan, Jianqing, and Runze Li. 2001. “Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties.” *Journal of the American Statistical Association* 96 (456): 1348–60.

Fernandez, Carmen, Eduardo Ley, and Mark FJ Steel. 2001. “Model Uncertainty in Cross-Country Growth Regressions.” *Journal of Applied Econometrics* 16 (5): 563–76.

Fernández-Villaverde, Jesús, and Juan Francisco Rubio-Ramírez. 2004. “Comparing Dynamic Equilibrium Models to Data: A Bayesian Approach.” *Journal of Econometrics* 123 (1): 153–87. doi:10.1016/j.jeconom.2003.10.031.

Gelfand, Alan E., and Adrian FM Smith. 1990. “Sampling-Based Approaches to Calculating Marginal Densities.” *Journal of the American Statistical Association* 85 (410): 398–409.

Geman, Stuart, and Donald Geman. 1984. “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images.” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6: 721–41.

Geweke, John. 1992. “Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments.” In *Bayesian Statistics 4*, edited by Jose M.

Bernardo, James O. Berger, A.P. dawid, and J.F.M Smith, 169–93. Oxford University Press.

Griffiths, Charley Xia. 2012. *Bayesian Unit Root Testing: The Effect of Choice of Prior On Test Outcomes*. The University of Melbourne.

Hamilton, James Douglas. 1994. *Time Series Analysis*. Vol. 2. Princeton university press Princeton.

Hansen, Peter R., and Asger Lunde. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics* 20 (7): 873–89. doi:10.1002/jae.800.

Harris, R. I. D. 1992. "Testing for Unit Roots Using the Augmented Dickey-Fuller Test: Some Issues Relating to the Size, Power and the Lag Structure of the Test." *Economics Letters* 38 (4): 381–86. doi:10.1016/0165-1765(92)90022-Q.

Harvey, Andrew C, and AC Harvey. 1993. *Time Series Models*. Vol. 2. Harvester Wheatsheaf New York.

Hastings, W. K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57 (1): 97–109. doi:10.2307/2334940.

Hoeting, Jennifer A., David Madigan, Adrian E. Raftery, and Chris T. Volinsky. 1999. "Bayesian Model Averaging: A Tutorial." *Statistical Science* 14 (4): 382–401.

Jing, Liang. 2010. "Hastings-within-Gibbs Algorithm: Introduction and Application on Hierarchical Model." University of Texas at San Antonio. http://georglsm.r-forge.r-project.org/site-projects/pdf/Hastings_within_Gibbs.pdf.

Koop, Gary, Eduardo Ley, Jacek Osiewalski, and Mark FJ Steel. 1997. “Bayesian Analysis of Long Memory and Persistence Using ARFIMA Models.” *Journal of Econometrics* 76 (1): 149–69.

Koop, Gary, and Simon Potter. 2004a. “Forecasting in Dynamic Factor Models Using Bayesian Model Averaging.” *The Econometrics Journal* 7 (2): 550–65.

———. 2004b. “Forecasting in Dynamic Factor Models Using Bayesian Model Averaging.” *The Econometrics Journal* 7 (2): 550–65.

Koske, I., and I. Wanner. 2013. “The Drivers of Labour Income Inequality – an Analysis Based on Bayesian Model Averaging.” *Applied Economics Letters* 20 (2): 123–26. doi:10.1080/13504851.2012.683164.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin. 1992. “Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” *Journal of Econometrics* 54 (1): 159–78.

Leamer, Edward E. 1983. “Let’s Take the Con Out of Econometrics.” *The American Economic Review* 73 (1): 31–43.

Lee, Peter M. 2012. *Bayesian Statistics: An Introduction*. John Wiley & Sons.

Ling, Shiqing, and WK Li. 2003. “Asymptotic Inference for Unit Root Processes with GARCH (1, 1) Errors.” *Econometric Theory* 19 (04): 541–64.

Lopez, J. Humberto. 1997. “The Power of the ADF Test.” *Economics Letters* 57 (1): 5–10. doi:10.1016/S0165-1765(97)81872-1.

Lubrano, Michel. 1995. “Testing for Unit Roots in a Bayesian Framework.” *Journal of Econometrics* 69 (1): 81–109.

Lucas, André. 1995. “Unit Root Tests Based on M Estimators.” *Econometric Theory* 11 (02): 331–46.

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. “Equation of State Calculations by Fast Computing Machines.” *The Journal of Chemical Physics* 21 (6): 1087–92.

Min, Chung-ki, and Arnold Zellner. 1993. “Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates.” *Journal of Econometrics* 56 (1–2): 89–118. doi:10.1016/0304-4076(93)90102-B.

Mitchell, James, Nigel Pain, and Rebecca Riley. 2011. “The Drivers of International Migration to the UK: A Panel-Based Bayesian Model Averaging Approach*.” *The Economic Journal* 121 (557): 1398–1444.

Moral-Benito, Enrique. 2013. “Model Averaging in Economics: An Overview.” *Journal of Economic Surveys*.

Morales, Knashawn H., Joseph G. Ibrahim, Chien-Jen Chen, and Louise M. Ryan. 2006. “Bayesian Model Averaging with Applications to Benchmark Dose Estimation for Arsenic in Drinking Water.” *Journal of the American Statistical Association* 101 (473): 9–17.

Nelson, Charles R, and Charles R Plosser. 1982. “Trends and Random Walks in Macroeconmic Time Series: Some Evidence and Implications.” *Journal of Monetary Economics* 10 (2): 139–62.

Nelson, Daniel B. 1991. “Conditional Heteroskedasticity in Asset Returns: A New Approach.” *Econometrica: Journal of the Econometric Society*, 347–70.

Ng, Serena, and Pierre Perron. 1995. “Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag.” *Journal of the American Statistical Association* 90 (429): 268–81. doi:10.2307/2291151.

———. 2001. “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power.” *Econometrica* 69 (6): 1519–54.

Pagano, Marcello. 1974. “Estimation of Models of Autoregressive Signal plus White Noise.” *The Annals of Statistics*, 99–108.

Perron, Pierre, and Serena Ng. 1996. “Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties.” *The Review of Economic Studies* 63 (3): 435–63.

Phillips, Peter CB. 1991. “To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends.” *Journal of Applied Econometrics* 6 (4): 333–64.

Phillips, Peter C. B., and Pierre Perron. 1988. “Testing for a Unit Root in Time Series Regression.” *Biometrika* 75 (2): 335–46. doi:10.2307/2336182.

Phillips, Peter C. B., and Peter Schmidt. 1989. *Testing for a Unit Root in the Presence of Deterministic Trends*. Cowles Foundation Discussion Paper 933. Cowles Foundation for Research in Economics, Yale University. https://ideas.repec.org/p/cwl/cwldpp/933.html.

Phillips, Peter CB, and Peter Schmidt. 1989. *Testing for a Unit Root in the Presence of Deterministic Trends*. Cowles Foundation for Research in Economics, Yale University. http://ideas.repec.org/p/cwl/cwldpp/933.html.

Poirier, Dale J. 1991. “A Bayesian View of Nominal Money and Real Output through a New Classical Macroeconomic Window.” *Journal of Business & Economic Statistics* 9 (2): 125–48. doi:10.2307/1391778.

Raftery, Adrian E. 1995. “Bayesian Model Selection in Social Research.” *Sociological Methodology* 25: 111–64.

Roberts, G. O., and A. F. M. Smith. 1994. “Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms.” *Stochastic Processes and Their Applications* 49 (2): 207–16. doi:10.1016/0304-4149(94)90134-1.

Rothenberg, Thomas J, and James H Stock. 1997. “Inference in a Nearly Integrated Autoregressive Model with Nonnormal Innovations.” *Journal of Econometrics* 80 (2): 269–86.

Said, Said E., and David A. Dickey. 1984. “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order.” *Biometrika* 71 (3): 599–607. doi:10.2307/2336570.

Sala, Xavier, Gernot Doppelhofer Martin, and Ronald I Miller. 2004. “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach.” *The American Economic Review*.

Schotman, Peter C. 1994. “Priors for the AR(1) Model: Parameterization Issues and Time Series Considerations.” *Econometric Theory* 10 (3/4): 579–95.

Schotman, Peter C, and Herman K Van Dijk. 1991. “On Bayesian Routes to Unit Roots.” *Journal of Applied Econometrics* 6 (4): 387–401.

Schwert, G. William. 1989a. “Tests for Unit Roots: A Monte Carlo Investigation.” *Journal of Business & Economic Statistics* 7 (2): 147–59. doi:10.2307/1391432.

———. 1989b. “Tests for Unit Roots: A Monte Carlo Investigation.” *Journal of Business & Economic Statistics* 7 (2): 147–59. doi:10.2307/1391432.

Seo, Byeongseon. 1999. “Distribution Theory for Unit Root Tests with Conditional Heteroskedasticity.” *Journal of Econometrics* 91 (1): 113–44.

Shiller, Robert J, and Pierre Perron. 1985. “Testing the Random Walk Hypothesis: Power versus Frequency of Observation.” *Economics Letters* 18 (4): 381–86.

Sims, Christopher A. 1988. “Bayesian Skepticism on Unit Root Econometrics.” *Journal of Economic Dynamics and Control* 12 (2): 463–74.

Sims, Christopher A, and Harald Uhlig. 1991. “Understanding Unit Rooters: A Helicopter Tour.” *Econometrica: Journal of the Econometric Society*, 1591–99.

Stock, James H. 1994. “Unit Roots, Structural Breaks and Trends.” *Handbook of Econometrics* 4: 2739–2841.

Tanner, Martin A. 1998. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. 3rd edition. New York: Springer.

Tibshirani, Robert. 1996. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)* 58 (1): 267–88.

Wright, Jonathan H. 2008. “Bayesian Model Averaging and Exchange Rate Forecasts.” *Journal of Econometrics* 146 (2): 329–41.

Zellner, Arnold. 1996. *An Introduction to Bayesian Inference in Econometrics*. 1 edition. New York; Chichester: Wiley-Interscience.

Zivot, Eric. 1994. “A Bayesian Analysis of the Unit Root Hypothesis within an Unobserved Components Model.” *Econometric Theory* 10 (3-4): 552–78.

Zivot, Eric, and Jiahui Wang. 2007. *Modeling Financial Time Series with S-PLUS®*. Vol. 191. Springer. http://books.google.com/books?hl=en&lr=&id=Mby9u8_WXXgC&oi=fnd&pg=PA1&dq=Modeling+Financial+Time+Series+with+S-PLUS&ots=ppvXOD137o&sig=MznbxY3r1ntQm8-obUvNSfui9CA.

Appendix A.

A. 1. Simulation Data Generating Process

The formula of data generating process (DGP) is introduced in Section 3, specifically (37) and (39). For every data, a random variable valued 1 to 6 is generated to decide the number of lags will be included in the process; the drift and deterministic trend are also randomly specified by a binomial distributed variable with probability 0.5. The dominant roots will range from 0.8 to 1, with 200 replications for small dominant root and 500 replications for larger one in order to better demonstrate test power. And for each dominant root various data sets are generated with different signal to noise ratio and sample size from 50 to 300. The detailed parameter setting is listed in Table A1.

Table A1. Simulation Design

DOMINANT ROOT	LAGS	COEFFICIENT	ROOTS
0.8	1	0.8	0.8
(REP 200)	2	0.68,0.1	0.8,-0.12
	3	0.6,-0.15,0.25	0.8,-0.1+0.55i,-0.1-0.55i
	4	0.52,-0.25,0.2,0.15	0.8,0.04+0.71i,0.04-0.71i,-0.36
	5	0.45,-0.25,0.15,0.1,0.1	0.8,0.17+0.7i,0.17-0.7i,0.34+0.34i,
	6	0.39,-0.15,0.15,0.05,0.05,0.05	-0.34-0.34i 0.8,0.26+0.57i,0.26-0.57i, -0.46,-0.23+0.54i,-0.23-0.54i
0.85	1	0.85	0.85
(REP 200)	2	0.733,0.1	0.85,-0.12
	3	0.682,-0.15,0.25	0.85,-0.08+0.54i,-0.08-0.54i
	4	0.625,-0.25,0.2,0.15	0.85,0.07+0.69i,0.07-0.69i,-0.36
	5	0.583,-0.25,0.15,0.1,0.1	0.85,0.2+0.69i,0.2-0.69i,
	6	0.57,-0.25,0.15,0.1,0.05,0.05	-0.33+0.34i,-0.33-0.34i 0.85,0.25+0.62i,0.25-0.62i, -0.16+0.5i,-0.16-0.5i,-0.46
0.9	1	0.9	0.9
(REP 200)	2	0.8,0.1	0.9,-0.1

	3	0.77,-0.15,0.25	0.9,-0.07+0.52i,-0.07-0.52i
	4	0.74,-0.25,0.2,0.15	0.9,0.09+0.68i,0.09-0.68i,-0.35
	5	0.72,-0.25,0.15,0.1,0.1	0.9,0.23+0.67i,0.23-0.67i, -0.32+0.33i,-0.32-0.33i
	6	0.66,-0.25,-0.1,0.2,0.15,0.1	0.9,0.39+0.79i,0.39-0.79i, -0.57,-0.23+0.44i,-0.23-0.44i
0.95 (REP 200)	1	0.95	0.95
	2	0.845,0.1	0.95,-0.11
	3	0.832,-0.15,0.25	0.95,-0.06+0.5i,-0.06-0.5i
	4	0.817,-0.25,0.2,0.15	0.95,0.1+0.7i,0.1-0.7i,-0.35
	5	0.81,-0.25,0.15,0.1,0.1	0.95,0.25+0.66i,0.25-0.66i, -0.32+0.33i,-0.32-0.33i
	6	0.78,-0.25,-0.1,0.2,0.15,0.1	0.95,0.42+0.77i,0.42-0.77i, -0.56,-0.23+0.44i,-0.23-0.44i
0.96 (REP 500)	1	0.96	0.96
	2	0.856,0.1	0.96,-0.1
	3	0.845,-0.15,0.25	0.96,-0.05+0.5i,-0.05-0.5i
	4	0.835,-0.25,0.2,0.15	0.96,0.11+0.66i,0.11-0.66i,-0.34
	5	0.828,-0.25,0.15,0.1,0.1	0.96,0.26+0.66i,0.26-0.66i, -0.32+0.33i,-0.32-0.33i
	6	0.805,-0.25,-0.1,0.2,0.15,0.1	0.96,0.43+0.76i,0.43-0.76i, -0.56,-0.23+0.44i,-0.23-0.44i
0.97 (REP 500)	1	0.97	0.97
	2	0.868,0.1	0.97,-0.1
	3	0.86,-0.15,0.25	0.97,-0.06+0.5i,-0.06-0.5i
	4	0.852,-0.25,0.2,0.15	0.97,0.11+0.66i,0.11-0.66i,-0.34
	5	0.846,-0.25,0.15,0.1,0.1	0.97,0.26+0.65i,0.26-0.65i, -0.32+0.32i,0.32-0.32i
	6	0.83,-0.25,-0.1,0.2,0.15,0.1	0.97,0.43+0.76i,0.43-0.76i, -0.55,-0.23+0.44i,-0.23-0.44i
0.98 (REP 500)	1	0.98	0.98
	2	0.878,0.1	0.98,-0.1
	3	0.873,-0.15,0.25	0.98,-0.05+0.5i,-0.05-0.5i
	4	0.868,-0.25,0.2,0.15	0.98,0.12+0.65i,0.12-0.65i,-0.34
	5	0.865,-0.25,0.15,0.1,0.1	0.98,0.26+0.65i,0.26-0.65i, -0.32+0.32i,-0.32-0.32i
	6	0.855,-0.25,-0.1,0.2,0.15,0.1	0.98,0.44+0.75i,0.44-0.75i, -0.55,-0.23+0.44i,-0.23-0.44i
0.99 (REP 500)	1	0.99	0.99
	2	0.889,0.1	0.99,-0.1
	3	0.887,-0.15,0.25	0.99,-0.05+0.5i,-0.05-0.5i
	4	0.885,-0.25,0.2,0.15	0.99,0.12+0.65i,0.12-0.65i,-0.35
	5	0.883,-0.25,0.15,0.1,0.1	

	6	0.878,-0.25,-0.1,0.2,0.15,0.1	0.99,0.27+0.65i,0.27-0.65i,-0.32+0.32i,-0.32-0.32i 0.99,0.44+0.75i,0.44-0.75i,-0.55,-0.22+0.44i,-0.22-0.44i
1 (REP 500)	1	1	1
	2	0.9,0.1	1,-0.1
	3	0.9,-0.15,0.25	1,-0.05+0.5i,-0.05-0.5i
	4	0.9,-0.25,0.2,0.15	1,0.12+0.65i,0.12-0.65i,-0.35
	5	0.9,-0.25,0.15,0.1,0.1	1,0.27+0.64i,0.27-0.64i,-0.32+0.32i,-0.32-0.32i
	6	0.9,-0.25,-0.1,0.2,0.15,0.1	1,0.45+0.75i,0.45-0.75i,-0.55,-0.22+0.44i,-0.22-0.44i

CHAPTER 3

DATA FREQUENCY ON THE STATIONARY TEST OF COMMODITY FUTURES PRICE

3.1 Introduction

Unit root testing is one of the most important procedures when performing time series analysis, since all the estimation and tests should be performed on a stationary series otherwise certain preprocess should be taken (like first difference) before adopting any analysis. Thus it is crucial to test the stationarity of the time series in hand accurately and efficiently. Two paths can lead researchers to achieve this goal: obtaining more data, or improving the unit root test, and these are also related to the two tasks in present paper which will be elaborated soon.

For collecting more data, one can try to consider a longer time span of data which will have more information related to the stationary property and can lead to a more reliable testing result. An alternative would be using higher frequency data while keeping the same time span. This is generally believed to not provide much information since intuitively, stationarity requires a series to pass its mean regularly at least within the test sample, and increasing the frequency while keeping the time span does not change this mean reversion within the sample (Boswijk and Klaassen 2012). However this is not the case if the low frequency data is constructed by systematic sampling, i.e., skipping certain intermediate observations from high frequency sample, and this type of sampling

is usually seen in stock market or asset market variables. For example, researchers sometimes pick the price of one day each week to construct weekly data from daily data. Choi (1992) demonstrated by simulation that this kind of data aggregation will lower the power of augmented Dickey-Fuller tests and Phillps-Perron tests, although Chambers (2004) showed that this is a finite sample effect and asymptotically it is still possible to consistently test for a unit root when sampling frequency varies.

Recently, Boswijk and Klaassen (2012) proved that the effects of systematic sampling on unit root testing is not negligible when a high-frequency sample has volatility clustering with fat-tailed innovations, which are the typical characteristics of financial market data. They simulated data sets and using likelihood ratio-based tests and conclude that these tests can have more power than the traditional ADF test on data processes holding the aforementioned behavior characteristics. This leads to the second way to test for unit roots more accurate: improve the testing method. As the two authors claimed in their paper, traditional asymptotic based test methods like the augmented Dickey-Fuller test and Phillips-Perron test will have low power when being applied on volatility clustering and heavy-tailed samples. Therefore many researches are devoted into efficiently testing these kind of data, which mainly concentrate on modeling the heavy-tail by some non-Gaussian likelihood function, or describing the high volatility behavior by specifying the conditional variance structure. For example, unit root tests which focus on a non-Gaussian likelihood are considered by Lucas (1995), and Rothenberg and Stock (1997). Seo (1999), Boswijk (2001) and Ling and Li (2003) studied tests based on a Gaussian GARCH likelihood. Although these tests increased the power when testing the financial data, one of the common issue for the existing testing

methods is that they all require some specific model specification assumption, no matter on the functional form (like the ADF test requires the number of lags specified in the model) or the error term distribution (Gaussian distribution, etc). Even for the GARCH likelihood mentioned above one can only assume the innovation conditional variance has GARCH structure not others, and meanwhile they still need to specify the GARCH level before actually performing the test. However, due to various circumstances such as data incompleteness, variable selection, distribution misspecification, etc., this procedure may lead to inappropriate model assumptions which will produce erroneous conclusion since estimates may well depend on the particular model considered (Moral-Benito 2013).

In this paper, we will devote efforts into the two aforementioned directions in hopes of improving unit root test results. Using 5 commodity futures price data (corn, soybean, cotton, live cattle and lean hog), which all display typical financial series characteristics, we first show that systematic sampling does have effects on the results of unit root testing by testing three different frequency samples: daily, weekly, and monthly. Then, more importantly, we will test the stationarity of these series by averaging 24 models using a Bayesian Model Averaging unit root test method derived in the previous chapter to confront the model specification uncertainty issue, and compare results with traditional unit root tests to show the performance of the BMA methods, as well as its ability to handle the model specification issue.

The rest of the paper is structured as follows. Section 2 introduces the robust numerical Bayesian model averaging unit root test method which will be adopted in the later analysis, including mathematical derivation of the method, specific models that will be averaged as well as what behaviors they are modeling for respectively. After the

comprehensive introduction, data sets used in the analysis are introduced, followed by the priors, posterior distributions and the sampling methods. Section 3 shows and analyzes the results of the test, while section 4 concludes the whole paper.

3.2 Robust Numerical Unit Root Test for Model Uncertainty

3.2.1 Model Parameterization

Although the variance function specification is the key description for high frequency data considered in this paper, it is also of importance to specify the mean function accurately. Besides, model uncertainty issue also arises here which can be handled by the approach adopted in the analysis. In this paper we will adopt a common autoregressive model with lag p as mean function which has the following form:

$$A(L)x_t = \varepsilon_t \tag{41}$$

where

$$A(L) = 1 - \rho_1 L - \rho_2 L^2 - ... - \rho_p L^p \tag{42}$$

The "L" is a lag operator which makes $Lx_t = x_{t-1}$ and in general $L^k x_t = x_{t-k}$. The stochastic term $\varepsilon_t \sim f(0, \sigma^2)$ has some distribution with mean 0 and variance σ^2. This can be any distribution as long as it satisfies the white noise assumption. Notice that the deterministic trend is eliminated here by theory since any deterministic trend would suggest market inefficiency (Dorfman 1993). Equation (41) can be re-written to a form that is more familiar (and denote $\mathbf{x} = (x_{t-1}, x_{t-2}, ..., x_{t-p})'$, $\boldsymbol{\rho} = (\rho_1, \rho_2 ..., \rho_p)$):

$$\begin{aligned} & x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + ... + \rho_p x_{t-p} + \varepsilon_t \\ & \Rightarrow x_t = \mathbf{x}'\boldsymbol{\rho} + \varepsilon_t, \varepsilon_t \sim f(0, \sigma^2) \end{aligned} \tag{43}$$

Given the mean function specification as above, we will consider models with autocorrelation lags from 1 to 6 to incorporate the uncertainty issue in the mean function, which will be discussed in detail later.

As mentioned before, financial series usually contains high-frequency innovations with fat-tailed distribution and displays volatility clustering (Boswijk and Klaassen 2012), which may have a effect on the results of unit root tests (Choi and Chung 1995). So it is necessary to incorporate modeling of these characteristics when performing the test in order to achieve a more reliable conclusion. This behavior usually requires extra modeling on the variance term and many possible models exist in the literature. Famous and widely used models include the Autoregressive Conditional Heteroscedasticity model (ARCH) developed by Engle (1982) which includes a nonlinear function in variance, and the Generalized Autoregressive Conditional Heteroskedasticity model (GARCH) proposed by Bollerslev (1986). After these two seminal papers many similar derivations of these models have been developed to handle different situations, like NGARCH (Engle and Ng 1993) and EGARCH (D. B. Nelson 1991). And for each model specification different assumptions about the distribution of stochastic term could be made to consider various data behaviors. In this paper, keeping the linear mean function to be autoregressive with lag 1 to 6, we will consider averaging four types of different variance function specifications to reflect the uncertainty on the variance structure:

(1). GARCH (1, 1) with Student's t distribution:

The Student's t distribution with small degrees of freedom will capture the heavy-tail characteristic of the analyzed data while the high volatile can be modeled by a GARCH specification, which has the following form:

$$
\begin{aligned}
x_t &= E(x_t \mid x_{t-1},...) + \varepsilon_t \\
\varepsilon_t &= z_t \sqrt{h_t} \\
h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}
\end{aligned} \tag{44}
$$

Model parsimony is one concern here since the purpose is investigating frequency effect on unit root tests rather than estimation or forecasting; besides, Hansen and Lunde (2005) compared 330 ARCH-type models in terms of their ability to describe the conditional variance and concluded that GARCH (1, 1) outperformed other higher level GARCH models so it seems enough to consider only lag 1 here. In this case, the conditional variance is described by h_t while the unconditional variance is $E\varepsilon_t^2 = \frac{\alpha_0}{1-\alpha_1-\beta_1}$. And the mean function $E(x_t \mid x_{t-1},...)$ comes from (43) with lag number from 1 to 6, respectively. To consider the possible fat-tail density, Student's t distribution with small degrees of freedom is specified to z_t which has the following distribution form:

$$
z_t \sim t(\nu) \sim \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}(1+\frac{x^2}{\nu})^{-\frac{\nu+1}{2}} \tag{45}
$$

with small value of ν satisfies $\nu > 2$. We will also set a prior to ν which will be discussed in detail later.

(2). GARCH (1, 1) with standard Normal distribution:

In this case, we consider the standard GARCH model, i.e. $z_t \sim N(0,1)$ with the familiar density function:

$$
z_t \sim N(0,1) \sim \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} z_t^2) \tag{46}
$$

All the other settings are same as (44).

(3). ARCH (1) with Student's t distribution:

The ARCH model is another common used model to describe highly volatile error behavior so we also consider it here, and it has the following specification:

$$\begin{aligned} x_t &= E(x_t \mid x_{t-1}, \ldots) + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t} \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned} \tag{47}$$

Meanwhile z_t is assumed to have Student's t distribution with form (45).

(4). Autoregressive in mean with Student's t distribution:

In this case we do not add any conditional heterskedasticity effect in the variance formula and focus only on modeling the heavy-tail behavior. The mean function is (43), and the stochastic term has t distribution with form (45).

The aforementioned four different models mainly focus on describing the variance behavior, together with mean specification of autoregressive model with lag 1 to 6 for each of the above four models, 24 models in total will be averaged to analyze one data series to get a comprehensive conclusion. In the next subsection the framework of the Bayesian model averaging unit root test will be introduced.

3.2.2 Bayesian Unit Root Test under Model Uncertainty

The main focus of unit root test will be on the coefficients of the lags, and the dynamic properties of the series can be investigated by examination of the following matrix:

$$A=\begin{bmatrix} \rho_1 & 1 & 0 & 0 & \cdots & 0 \\ \rho_2 & 0 & 1 & 0 & \cdots & 0 \\ \rho_3 & 0 & 0 & \ddots & & \vdots \\ \vdots & 0 & 0 & \cdots & 0 & 1 \\ \rho_p & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{48}$$

The first column is composed of coefficients of lags in the AR model. The roots, or eigenvalues of this matrix, determine the behavior of the modeled process. If any eigenvalues of $\mathbf{A}$ have moduli greater or equal than one, the time series defined by these coefficients has a unit root and is nonstationary. So denote $\phi_i = \|\lambda_i\|$ which is the moduli of the eigenvalues of $\mathbf{A}$ sorted by its magnitude where $\|a\|$ stands for the Euclidian distance of a. Thus the statistical hypothesis of unit root can be expressed as:

$$H_0 : \phi_1 \geq 1 \text{ vs. } H_1 : \phi_1 < 1 \tag{49}$$

Unlike the classical unit root test method which focuses on deriving the asymptotical distribution of the coefficient on which the inference is to be made, in the Bayesian framework there is really no "test" concept similar to the classical world since the Bayesian method requires a posterior distribution which combines the likelihood function with prior distributions and usually does not have standard form to perform a regular "test". One alternative is the Bayes factor approach, which defines the Bayes factor B in favor of H_0 against H_1 is defined as (Lee 2012):

$$B = \frac{(p_0 / p_1)}{(\pi_0 / \pi_1)} = \frac{p_0 \pi_1}{p_1 \pi_0} \tag{50}$$

And the elements in (50) are defined as:

$$
\begin{aligned}
p_0 &= P(\theta \in \Theta_0 \mid \mathbf{x}) \\
p_1 &= P(\theta \in \Theta_1 \mid \mathbf{x}) \\
\pi_0 &= P(\theta \in \Theta_0) \\
\pi_1 &= P(\theta \in \Theta_1)
\end{aligned} \tag{51}
$$

in which θ is the unknown parameter from a set Θ ($\Theta_0 \bigcup \Theta_1 = \Theta, \Theta_0 \bigcap \Theta_1 = \varnothing$) we want to test, and $\mathbf{x}$ is the data available whose density $f(\mathbf{x} \mid \theta)$ depends on θ. Here p_0 and p_1 are posterior probabilities of parameter supports of null (Θ_0) and alternative (Θ_1) hypotheses, while π_0 and π_1 are prior probabilities. We will use this approach to perform our test since it is also convenient to numerical Bayesian technique. Let $\mathbf{\Phi} = (\phi_1, \phi_2, ..., \phi_p)'$ be the vector of the moduli of the eigenvalues of the matrix $\mathbf{A}$, and $\boldsymbol{\eta}$ stands for all the other parameters which varies with different variance specification (for example, in GARCH (1,1) with student t distribution, $\boldsymbol{\eta}=(\alpha_0, \alpha_1, \beta_1, \nu)'$; in GARCH (1, 1) with normal distribution, $\boldsymbol{\eta}=(\alpha_0, \alpha_1, \beta_1, \sigma^2)'$), and the prior distributions are denoted as $\pi(\mathbf{\Phi}) = \pi(\phi_1, \phi_2, ..., \phi_p)$ and $h(\boldsymbol{\eta})$ respectively (the detailed prior distribution discussion will be given in next part). With likelihood function denoted as $f(x; \boldsymbol{\eta}, \mathbf{\Phi})$, the marginal posterior of the dominant root would be:

$$
p(\phi_1 \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; x) = \int_{\phi_2} \cdots \int_{\phi_p} \int_{(\mathbf{\Phi}, \boldsymbol{\eta})} f(x; \boldsymbol{\eta}, \mathbf{\Phi}) \pi(\mathbf{\Phi}) h(\boldsymbol{\eta}) d\boldsymbol{\eta} d\phi_2 \cdots d\phi_p \tag{52}
$$

Using numerical methods (to be discussed later) samples can be drawn from this marginal posterior distribution. Define an indicator function:

$$
D(\phi_{(1)}) = \begin{cases} 1 & \phi_{(1)} < 1 \text{ in the sample (support } H_1) \\ 0 & \text{otherwise (support } H_0) \end{cases} \tag{53}
$$

The posterior probability in support H_1 is given by:

$$K_1 = p(\mathrm{H}_1 \mid \mathbf{x}) = \frac{\sum_{i=1}^{B} D(\phi_{(1)})\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{x})}{\sum_{i=1}^{B} p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{x})} \tag{54}$$

where B is the number of valid iterations used in the sampler. Defining K_0 in a analogous way to (54), we have:

$$K_0 = P(\mathrm{H}_0 \mid \mathbf{x}) = \frac{\sum_{i=1}^{B} [1 - D(\phi_{(1)})]\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{x})}{\sum_{i=1}^{B} p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathbf{x})} \tag{55}$$

and the posterior odds ratio in favor of nonstationarity of the tested series is:

$$K_{01} = \frac{\sum_{i=1}^{B} [1 - D(\phi_{(1)})]\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathrm{y})}{\sum_{i=1}^{B} D(\phi_{(1)})\, p(\phi_{(1)} \mid \boldsymbol{\eta}, \phi_2, ..., \phi_p; \mathrm{y})} = \frac{K_0}{K_1} \tag{56}$$

This statistic is the center of the numerical Bayesian unit test we derived in our method. And one can declare the series to have a nonstationary root if $K_{01} > 1$ (or depending on your loss function, at a different threshold).

The above discussion showed how to infer the stationarity properties for a series using numerical Bayesian unit root test based on single model. For our purpose to incorporate model specification uncertainty, the next step would be averaging all results concluded from possible candidate models and come to a final comprehensive conclusion. Suppose there are k possible candidate models which can hold differences in various aspects, for instance in our cases, different number of time lags can be included

in the model, and the specification of the variance terms could also be different. We assign a model prior probability distribution across the possible model space, so each potential model has probability $pr(\mathrm{M}_k)$, so the posterior probability for model M_k is:

$$pr(\mathrm{M}_k \mid \mathrm{x}) = \frac{pr(\mathrm{x} \mid \mathrm{M}_k)\, pr(\mathrm{M}_k)}{\sum_{l=1}^{k} pr(\mathrm{x} \mid \mathrm{M}_l)\, pr(\mathrm{M}_l)} \tag{57}$$

and the marginal likelihood function under M_k is

$$pr(\mathrm{x} \mid \mathrm{M}_k) = \int_S pr(\mathrm{x} \mid \boldsymbol{\Omega}_{\mathbf{k}}, \mathrm{M}_k)\, pr(\boldsymbol{\Omega}_{\mathbf{k}} \mid \mathrm{M}_k)\, \mathrm{d}\boldsymbol{\Omega}_{\mathbf{k}} \tag{58}$$

$\boldsymbol{\Omega}_{\mathbf{k}}$ here is all the parameters in the model M_k (in our case, $\boldsymbol{\Omega}_k = (\boldsymbol{\Phi}_k, \boldsymbol{\eta}_k, \ \upsilon_k)'$, and υ_k is another parameter which will be introduced later) and $pr(\boldsymbol{\Omega}_{\mathbf{k}} \mid \mathrm{M}_k)$ is the prior distribution of parameters $\boldsymbol{\Omega}$ under model M_k, and S is the support of $\boldsymbol{\Omega}$. Then the final comprehensive probability of a possible unit root across the whole model space is:

$$pr(\phi_1 \geq 1 \mid \mathrm{x}) = \sum_{i=1}^{k} pr(\phi_1 \geq 1 \mid \mathrm{M}_k, \mathrm{x})\, pr(\mathrm{M}_k \mid \mathrm{x}) \tag{59}$$

in which $pr(\phi_1 \geq 1 \mid \mathrm{M}_k, \mathrm{x})$ is the posterior probability of a unit root given model M_k derived in(52). Finally we can compute the probability of the dominant root larger than 1 considering all these circumstances and draw a conclusion according to the decision rule defined in (56).

3.3 Testing the Unit Root in Commodities Futures Data with Different Frequency

3.3.1 Data

In this paper, 5 commodity futures prices series are used to test and compare the unit root results: corn, soybean, cotton, live cattle and lean hog. To evaluate the effect of data frequency on the testing result, 3 different frequencies are used for each series: daily, weekly and monthly. The high-frequency sample is the real daily settled price of each commodity from Chicago Board of Trade (corn, soybean, live cattle and lean hog) and Intercontinental Exchange (cotton)[1]. Each daily data sample size is 2,000 which is from March, 2007 to March, 2015. The low-frequency sample is constructed from the daily data by what is usually referred to as systematic sampling. Assume the daily sample is $\mathbf{Y}_t$, we skip certain observations to achieve the low-frequency data (Boswijk and Klaassen 2012), i.e.,

$$\mathbf{Y}_j^* = \mathbf{Y}_{mj},\ j = 0,...,n^* = \frac{n}{m} \tag{60}$$

where for weekly data we take m=5 and m=20 for monthly data, which can be treated as end-of-week and end-of-month price given 5 trading days in a week and 20 days in a month. Since the daily data sample size is 2000, the constructed weekly sample size is 400 and the monthly data size is 100. The first row of each subgraph of Figure 1 shows the time series plot of each commodity with different frequency samples which gives a direct view of the data.

[1] The continuous futures data are available for downloading from the Open Financial Data Project in Quandl: https://www.quandl.com. The front month contact price is used to build the continuous data.

3.3.2 Priors, Posterior Distribution and Sampling Methods

The choice of priors for the parameters is always a crucial issue in Bayesian econometric analysis. Numerous literature has been devoted into the area of trying to find properties and suitability for different priors. Berger and Yang (1994) compared various approaches to the development of a noninformative prior for the AR(1) model, and recently Griffiths (2012) performed a Monte Carlo experiment as well as an application to real world data to examine the effects of choices of different priors on Bayesian unit root test outcomes. In light of this continuing debates, in this research we do not plan to spend much time on judging the priors since this is not the main point of this research and it is still hard or impossible to find a perfect prior that works well under all circumstances due to the fact that it may well depend on the data and topic (preknowledge). Instead we will follow the approach similar to Dorfman (1993), assigning independent Beta priors to the moduli of roots of (48), while maintaining an uninformative priors on other parameters. Meanwhile, a slightly explosive situation is allowed for the value of the dominant root in order to take the possible upward bias due to sampling error from estimation of the root into account (Dorfman 1993). Specifically, the informative priors of all the roots are:

$$\begin{aligned} &\pi(\mathbf{\Phi}) = \pi(\phi_1, \phi_2, ..., \phi_p) = \pi_1(\phi_1)\pi_2(\phi_2)...\pi_p(\phi_p) \\ &\pi_1(\phi_1 - \tau) \sim \text{Beta}(30, 2); \pi_2(\phi_2) = .\pi_3(\phi_3) = ... = \pi_p(\phi_p) \sim \text{Beta}(1.1,1.1) \end{aligned} \tag{61}$$

These priors are informative on the moduli of the dominant root since it severely skewed to the right, and ranging from τ to $(1+\tau)$. The priors on all the other roots are quite weakly informative with very flat, round curves ranging over [0,1]. Parameter υ controls the degree of allowed explosive from unity, and it is also assigned a prior to fully involve

in the Bayesian framework. Griffiths (2012) suggested a hierarchical approach which we follow here, assigning an exponential distribution with mean $E(\tau) = \kappa^{-1}$ as a prior to τ, i.e.,

$$f(\tau \mid \kappa) = \kappa e^{-\kappa\tau} \tag{62}$$

where κ is chosen such that $P(\tau > 0.15) = 0.05$.

For the parameters in the conditional variance specification, the prior can be uninformative as long as it satisfies positivity in all circumstances. The common way in past practices was to strictly restrict all the parameters to be non-negative and the constant to be positive. Instead doing so, we impose positivity directly on volatility which is inspired by Dorfman and Park (2011). By doing so we allow negative coefficients in the conditional variance which widened the parameter space and possibility of getting better estimates if the negative coefficients are truly more appropriate than their positive counterpart. Thus all priors are set to have a normal distribution with mean zero and are mutually independent. So the prior of the volatility coefficient can be represented as:

GARCH (1, 1): $p(\alpha_0, \alpha_1, \beta_1) = I(h_t) \bullet N_{\alpha_0}(0,3) \bullet N_{\alpha_1}(0,3) \bullet N_{\beta_1}(0,3)$

ARCH (1): $p(\alpha_0, \alpha_1) = I(h_t) \bullet N_{\alpha_0}(0,3) \bullet N_{\alpha_1}(0,3)$

The $I(h_t)$ here is an indicator function used to control the positivity of volatility which equals one if a sampled parameter vector provides for a positive conditional variance in all time periods and equals zero otherwise.

Another parameter to consider is the degrees of freedom ν of the Student's t distribution. For modeling the fat-tail behavior well, the large part of the density should

be kept around a small value, yet we also need to have some degree of spreading to incorporate more possible candidates into consideration. In this case we will assign a truncated exponential distribution as a prior for ν which has the following density form:

$$p(\nu \mid \lambda, \delta) = \lambda \exp(-\lambda(\nu - \delta)) \bullet I(\nu > \delta) \tag{63}$$

The values of λ and δ can be set to control the density shape of the exponential distribution. The indicator function and parameter δ are used to make sure the generated t distribution is well defined (so $\delta \geq 2$); it can also be used to approximate a Normal distribution if the value of δ is set to be relatively large.

Lastly, for the models assuming the stochastic term follows Normal distribution, an inverse-gamma distribution is assigned as the prior on the variance σ^2 which has the following density form:

$$\sigma^2 \sim IG(\alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} [\sigma^2]^{-\alpha-1} \exp(-\frac{\beta}{\sigma^2}) \tag{64}$$

This is commonly used in practice since it is the conjugate prior of the Normal likelihood function. Now we can write the full form of the joint posterior distribution for each model as follows:

(1). GARCH (1, 1) with Student's t distribution:

$$\begin{aligned} p(\mathbf{\Phi}, \alpha_0, \alpha_1, \beta_1, \nu, \upsilon \mid \mathbf{x}) \propto\ & L(\mathbf{x} \mid \boldsymbol{\rho}, \alpha_0, \alpha_1, \beta_1, \nu) \times \pi(\mathbf{\Phi}) \\ & \times p(\alpha_0, \alpha_1, \beta_1) \times p(\nu \mid \lambda, \delta) \times f(\upsilon \mid \kappa) \end{aligned} \tag{65}$$

The term $L(\mathbf{x} \mid \boldsymbol{\rho}, \alpha_0, \alpha_1, \beta_1, \nu)$ is the likelihood function of a Student's t distribution which has the following form (assume sample size is T, ignoring the effects of number of lags included in the model to the sample size):

$$L(\mathbf{x}\,|\,\boldsymbol{\rho},\alpha_0,\alpha_1,\beta_1,\nu)=\left[\frac{\Gamma[(\nu+1)/2]}{\pi^{1/2}\Gamma(\nu/2)}(\nu-2)^{-1/2}\right]^T \exp\left\{-\frac{1}{2}\prod_{t=1}^{T}h_t-\prod_{t=1}^{T}\left[1+\frac{(x_t-\mathbf{x}'\boldsymbol{\rho})^2}{h_t(\nu-2)}\right]^{[(\nu+1)/2]}\right\} \quad (66)$$

$$h_t=\alpha_0+\alpha_1\varepsilon_{t-1}^2+\beta_1 h_{t-1}$$

And all the other terms in (65) are defined before in the previous part of this paper.

(2). GARCH (1, 1) with Normal distribution:

$$\begin{aligned} p(\boldsymbol{\Phi},\alpha_0,\alpha_1,\beta_1,\sigma^2,\upsilon\,|\,\mathbf{x}) \propto\; & L(\mathbf{x}\,|\,\boldsymbol{\rho},\alpha_0,\alpha_1,\beta_1,\sigma^2)\times\pi(\boldsymbol{\Phi}) \\ & \times p(\alpha_0,\alpha_1,\beta_1)\times p(\sigma^2\,|\,\alpha,\beta)\times f(\upsilon\,|\,\kappa) \end{aligned} \quad (67)$$

Again the term $L(\mathbf{x}\,|\,\boldsymbol{\rho},\alpha_0,\alpha_1,\beta_1,\sigma^2)$ is the likelihood function of Normal distribution which has the following:

$$L(\mathbf{x}\,|\,\boldsymbol{\rho},\alpha_0,\alpha_1,\beta_1,\sigma^2)=\left(2\pi\right)^{-T/2}\prod_{t=1}^{T}h_t^{-1/2}\exp\left[-\frac{1}{2h_t}\left(\prod_{t=1}^{T}(x_t-\mathbf{x}'\boldsymbol{\rho})^2\right)\right] \quad (68)$$

$$h_t=\alpha_0+\alpha_1\varepsilon_{t-1}^2+\beta_1 h_{t-1}$$

The posterior densities of the other two models (ARCH (1) with Student's t distribution and regular autoregressive model with Student's t distribution) will be omitted here since they have similar formula as the above with minor parameter changes.

With all the discussion and assembling above, to find a final decision of stationarity of the analyzed series all we need to do is that find the marginal posterior of the dominant root using equation (52) by integrating out all the nuisance parameter for every model, then averaging the results of all 24 models by (59). However, the first task is quite difficult, sometimes impossible to do analytically, since the joint posterior density is usually too complicated to have a closed form. In our case, the combination of different prior densities makes the posterior not have a standard distribution form; furthermore, we

specify the Beta prior to the moduli of the eigenvalues of matrix **A** in (48) for the purpose of being logically straightforward, while the calculation of likelihood function actually needs the value of the coefficient $\boldsymbol{\rho}$. One needs to explicitly formulate the mapping process in order to achieve the corresponding likelihood function. Without solving these arduous issues analytically, numerical method will be adopted to achieve an empirical sample from the marginal posterior distribution of interested parameters, and Metropolis–Hastings (MH) algorithm will be adopted in our computation. It is one of the Markov Chain Monte Carlo (MCMC) methods that samples from a conditional distributions of subsets of parameters in order to constructing a Markov chain whose convergence distribution is the desired distribution. It was first suggested by Metropolis et al. (1953) and later generalized by Hastings (1970) and now is widely applied in all areas which need numerical sampling technique including econometrics, statistics, physics and engineering. The basic idea to the Metropolis-Hastings algorithm is setting up an "accept/reject" rule to keep revising the Markov chain so that it can reach the aimed convergence distribution (Tanner 1998) . Assume that random variables $\mathbf{X}$ have joint distribution $\psi(\mathbf{X})$ from which the sample need to be drawn. A MH algorithm requires choosing a proposal distribution $q(\cdot \mid \mathbf{X})$ which generates a candidate chain based on the current values of random variables, then decides to accept or reject according to the detailed balance rule. To be more specific, for this paper, the algorithm can be described as below:

(1). Assign initial values to parameter space $\mathbf{X}$: $\mathbf{X}^{(0)} = (\boldsymbol{\rho}^{(0)}, \boldsymbol{\eta}^{(0)})$, here $\boldsymbol{\rho}$ is coefficients of different lags whose length may depend on models, and $\boldsymbol{\eta}$ is all the other parameters which varies according to different volatility specification;

(2). For iteration m from 1 to N, repeat the following procedures:

a. generate a candidate value from proposal density, $\mathbf{X}^{*} \sim q(\cdot \mid \mathbf{X}^{(\mathrm{m})})$, in this paper the proposal density is multivariate Normal distribution;

b. generate a random variable $u \sim Uniform(0,1)$;

c. decide the next sample unit using the following rule:

$$\mathbf{X}^{(\mathrm{m}+1)} = \begin{cases} \mathbf{X}^{*} & \text{if } u \le \alpha(\mathbf{X}^{(m)}, \mathbf{X}^{*}) \\ \mathbf{X}^{(\mathrm{m})} & \text{otherwise} \end{cases}, \text{ in which}$$

$$\alpha(\mathbf{X}^{(\mathrm{m})}, \mathbf{X}^{*}) = \begin{cases} \min\{\dfrac{\psi(\mathbf{X}^{*})\, \mathrm{q}(\mathbf{X}^{(\mathrm{m})} \mid \mathbf{X}^{*})}{\psi(\mathbf{X}^{(m)})\, \mathrm{q}(\mathbf{X}^{*} \mid \mathbf{X}^{(\mathrm{m})})}, 1\} & \text{if } \psi(\mathbf{X}^{(\mathrm{m})})\, \mathrm{q}(\mathbf{X}^{*} \mid \mathbf{X}^{(\mathrm{m})}) > 0 \\ 1 & \text{otherwise} \end{cases}$$

(3). Return the final values $(\mathbf{X}^{(0)}, \mathbf{X}^{(1)}, ..., \mathbf{X}^{(N)})$ which is the sample from the posterior density.

The MH algorithm is flexible since the proposal density can be arbitrary, but there is a price for this convenience (Jing 2010). If the proposal density is not chosen appropriately, the acceptance rate is low which means that the sample we create may include lots of noise compared to the convergence distribution we actually want. Also the Markov Chain may not cover the whole support (or at least a large part) of the convergence distribution; sometimes it even can stick to a small area which leads to very poor sampling efficiency. We will adjust the proposal density repeatedly by tuning the mean and variance, as well as trying different sets of starting values, to achieve certain rate of acceptance to guarantee the good mixture of posterior density.

3.3.3 Results

We tested for unit roots in all 3 different sample frequencies of 5 commodities futures price by averaging 4 types of conditional variance structures introduced in Section 3.2: GARCH (1, 1) with error term modeled as Student's t distribution; GARCH (1, 1) with error term assumed as Normal distribution, ARCH (1) with Student's t distribution and regular AR models with Student's t distribution. For each type of variance structure, the mean functions are assumed to be autoregressive model with maximum lag length between 1 and 6 to incorporate model specification uncertainty in the mean process. So for each data series the final decision is made by averaging 24 models. For each model, we set 51,000 Monte Carlo iterations for the MH algorithm to achieve better convergence with elimination of first 21,000 draws for better posterior sample mixture, and the Geweke test (Geweke 1992) is adopted to examine the convergence of each posterior sample.

Table 3.1 lists the probability of a dominant root greater than 1 using our BMA methods according to (55), where a value greater than or equal to 0.5 is considered to support nonstationarity. To compare, the p value of two other commonly used unit root test methods, the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test are also shown in the table. One thing to remember is that Bayesian and frequentist tests are complete different logically. For instance, Bayesian tests do not conform to the conventional 0.05 size from the traditional sampling theory (Griffiths 2012). Also one fundamental difference between these two frameworks is that the traditional frequentist framework treats unknown parameters as fixed and data as random, while the Bayesian technique believes data is fixed and treats parameters as random. In this case it might be

inappropriate to evaluate a “pure” Bayesian test which makes decisions conditional on the fixed observed data using criteria constructed by sampling theory whose estimates based on data at hand plus hypothetical repeated sampling in the future with similar data, and one should be cautious when using standards in frequentist econometrics like “power” to compare tests in Bayesian framework although it might be a good measurement of the performance of a test in frequentist econometrics. Regardless of these potential debates, in this paper some of the terminologies will be used for conceptional convenience when comparing the results of BMA with ADF and PP test. One should bear in mind that no matter how the comparison results turn out, the BMA method proposed in this paper provided a solution to performing unit root tests under model specification uncertainty, which cannot be achieved by other traditional test methods since they all require some kind of model restrictions.

Generally speaking the results various among different commodities as well as different data frequencies, which is the focus of this paper. First, notice that for the BMA result, although the probability of having a unit root varies between different frequency samples, the conclusion is basically consistent except for cotton. It is believed that the futures prices are stationary for the commodities soybean, live cattle and lean hogs, but nonstationary for corn. For cotton, the BMA method indicates a unit root exists for weekly and monthly data but not in daily data, although the probability is 0.498 which is quite close to 0.5, which might be thought as “marginally stationary” and could be caused by sampling error.

Another result to notice is that for each commodity, the probability of a dominant root greater than 1, $pr(\phi_1 \geq 1 | \mathrm{x})$, computed by averaging 24 models using BMA method

is increasing as the frequency of tested data decreasing. This indicates that more mean-reversion information is provided by using the high frequency data which is consistent with the conclusion in Boswijk and Klaassen (2012). To be more specific, high frequency samples carry more information through high volatile and fat-tail behavior which can be captured by GARCH and ARCH models with Student's t distribution in the BMA method. This information will be lost when constructing low frequency sample through system sampling (constructing low frequency sample by skipping intermediate high-frequency observations).

Table 3.2 shows the residual kurtosis of the AR (1) mean function. Here the mean function will not have significant effect on the variability of the residuals so only AR (1) residuals are shown here for simplicity. Kurtosis is a statistic which primarily describes the peakedness (width of peak), tail weight, and lack of shoulders of probability distribution; and the larger the kurtosis is, the higher and sharper the central peak is, and the longer and fatter the tails will be. It is clear that the high frequency daily data has leptokurtic property, while for low frequency weekly and monthly data the kurtosis value significantly dropped to small values. So this extra information contained in daily sample can be well captured by GARCH and ARCH models in the BMA method but will be ignored by traditional methods like ADF or PP test. In Table 1, one can find that the conclusion based on ADF or PP test does not vary across different sample frequencies, which indicates the power of the ADF and PP test is hardly effected by the sampling frequency (Boswijk and Klaassen 2012). As mentioned before which is the most desirable property of BMA method is it can handle model specification uncertainty in a unit root test. This issue exists in all empirical modeling work since we all need some sort

Table 3.1. Unit Root Test Results of Five Commodity Futures Prices Data

		BMA	DF 1	DF 2	DF 3	DF 4	DF 5	DF 6	PP
CORN	Day	0.503	0.576	0.538	0.475	0.339	0.390	0.315	0.542
	Week	0.717	0.530	0.535	0.497	0.577	0.648	0.622	0.552
	Month	0.784	0.576	0.538	0.475	0.338	0.390	0.315	0.519
SOYBEAN	Day	0.372	0.271	0.312	0.282	0.332	0.359	0.357	0.298
	Week	0.406	0.346	0.381	0.319	0.296	0.277	0.256	0.317
	Month	0.481	0.227	0.098	0.163	0.113	0.0431	0.0264	0.263
COTTON	Day	0.498	0.622	0.627	0.628	0.624	0.579	0.647	0.621
	Week	0.535	0.593	0.651	0.662	0.641	0.550	0.499	0.616
	Month	0.685	0.489	0.320	0.216	0.468	0.463	0.402	0.519
LIVE CATTLE	Day	0.222	0.537	0.570	0.545	0.559	0.546	0.553	0.542
	Week	0.254	0.530	0.535	0.497	0.577	0.648	0.622	0.552
	Month	0.408	0.576	0.538	0.475	0.338	0.390	0.315	0.519
LEAN HOG	Day	0.141	0.084	0.060	0.010	0.010	0.010	0.010	0.010
	Week	0.188	0.010	0.010	0.010	0.010	0.010	0.0188	0.010
	Month	0.396	0.010	0.010	0.010	0.010	0.010	0.0262	0.010

Notes:
1. First column shows the probability of the dominant root equals to or greater than 1 ($pr(\phi_1 \geq 1 | \mathrm{x})$), and the value equals to or greater than 0.5 indicates evidence of nonstationarity.
2. Column 2-8 lists the p-values from Phillips-Perron (PP) test and augmented Dickey-Fuller tests with lag length 1 to 6 (DF1-DF6). Usually p-value larger than 0.1 is considered as evidence of nonstationarity.

of assumptions in the analysis, but the importance of this issue may depend on the goal of the research or researcher's belief. Sometimes the model uncertainty may cause contradictory results which may bring troubles to the succeeding analysis. In this paper one such example is the soybean monthly sample. Using the ADF test on the monthly data and under a commonly used 10% significance level, it is confirmed nonstationary if the model specification is AR (1), AR (3) or AR (4), while for AR (2), AR (5) and AR (6) the test indicates stationarity of the data which is opposite to the result using other lags as well as daily and weekly data. So the model specification uncertainty problem is important here since improper specification of the lag will lead to completely different results which will affect the following analysis. The BMA method confronts this problem by averaging all 6 possible lag specification (or more if the researcher needed) and reaching a final, more robust conclusion. Table 3.3 lists the number of unit root detected by each variance structure specification in BMA methods. Notice that although in many situations all AR (1) to AR (6) have detected the existence of unit root, under several cases only part of the specifications did, which proved the issue of model uncertainty has an effect on the testing result, as well as the ability of our BMA approach to handle this issue. Also in this paper, four different treatments of the conditional variance are averaged to consider the uncertainty in the specification of the variance structure. This is also an advantage compared to the aforementioned papers which also investigated the effects of sample frequency on the unit root test result but still based their conclusions on one preassumed model functional form. As can be seen from Table 3.3, most findings of nonstationarity are concentrated on GARCH model with Normal errors and the regular AR model with Students' t errors, which capture the heteroscedasticity and heavy tailed

Table 3.2. AR (1) Residual Kurtosis of Different Samples

COMMODITY	FREQUENCY	AR (1) RESIDUAL KURTOSIS
CORN	daily	23.97
	weekly	5.32
	monthly	3.53
SOYBEAN	daily	26.40
	weekly	1.03
	monthly	0.70
COTTON	daily	60.22
	weekly	17.80
	monthly	6.31
LIVE	daily	23.97
CATTLE	weekly	5.32
	monthly	3.53
LEAN HOG	daily	23.97
	weekly	5.32
	monthly	3.53

behaviors respectively, so any single model specification may not fully describe the behavior of the data and information will be lost which may lead to an inappropriate conclusion. Table 3.4 shows the posterior ratio of each type of conditional variance specification (aggregating the mean function AR (1)-AR (6)) in the BMA method for 3 frequency samples of each commodity. It shows that for the daily sample, GARCH and ARCH models takes a almost majority of the posterior weights which captures the high volatile and heavy tail behavior. For low frequency data like monthly, the weight of AR-t is increasing comparing to other frequencies since the kurtosis is low so the high volatility property is fading away, which reduced the weights of ARCH and GARCH models; meanwhile, they still contain heavy-tail behavior which can be captured by the Student's t distribution used in the model. The heavy-tail character can be seen from the second row of Figure 3.1, which plots the densities of residuals of AR (1) model (for simplicity purpose we skip the other mean function form, but the fat tail behavior exists in models with all AR lags).

The results shows that BMA method can handle model specification uncertainty well. Meanwhile, it can also self-adjust the posterior weight to the "right" model which can describe the data behavior most accurately. Fernández-Villaverde and Francisco Rubio-Ramírez (2004) actually proved that asymptotically, Bayesian methods will give the highest posterior probability to the best model under the Kullback-Leibler distance.

Table 3.3 Number of Unit Roots Detected by Each Model Specification in BMA Methods

FREQUENCY	MODEL	CORN	SOYBEAN	COTTON	LIVE CATTLE	LEAN HOG
DAILY	GARCH-T	0	0	0	0	0
	GARCH-N	6	6	6	6	0
	ARCH-T	0	0	0	0	0
	AR-T	6	0	6	5	0
	Total	12	6	12	11	0
WEEKLY	GARCH-T	0	0	0	0	0
	GARCH-N	6	4	6	3	0
	ARCH-T	0	0	0	0	0
	AR-T	6	6	6	6	0
	Total	12	10	12	9	0
MONTHLY	GARCH-T	0	0	2	0	3
	GARCH-N	6	6	6	6	0
	ARCH-T	0	0	0	0	0
	AR-T	6	1	6	6	0
	Total	12	7	14	12	3

Note:
Numbers indicate that how many unit roots detected under each variance structure specification, and the "Total" shows the summation of all 4 structures which gives the overall number of unit roots in 24 models.

Table 3.4 Posterior Weights of Each Type of Model for Different Frequency

COMMODITY	MODEL	SAMPLE FREQUENCY		
		Daily	Weekly	Monthly
CORN	GARCH (1, 1)-t	0.244	0.0307	0.189
	GARCH (1, 1)-Normal	0.277	0.521	0.361
	ARCH (1)-t	0.231	0.0193	0.182
	AR-t	0.248	0.429	0.268
SOYBEAN	GARCH (1, 1)-t	0.0892	0.169	0.192
	GARCH (1, 1)-Normal	0.462	0.172	0.158
	ARCH (1)-t	0.0758	0.176	0.181
	AR-t	0.373	0.483	0.469
COTTON	GARCH (1, 1)-t	0.173	0.305	0.263
	GARCH (1, 1)-Normal	0.368	0.417	0.249
	ARCH (1)-t	0.152	0.135	0.261
	AR-t	0.307	0.143	0.227
LIVE	GARCH (1, 1)-t	0.388	0.351	0.308
CATTLE	GARCH (1, 1)-Normal	0.115	0.158	0.196
	ARCH (1)-t	0.385	0.347	0.312
	AR-t	0.112	0.144	0.184
LEAN HOG	GARCH (1, 1)-t	0.305	0.218	0.223
	GARCH (1, 1)-Normal	0.321	0.433	0.322
	ARCH (1)-t	0.205	0.215	0.221
	AR-t	0.169	0.134	0.234

3.4 Conclusion

Financial data series usually display high volatility clustering with heavy tail distributions, which may affect the results of traditional asymptotic theory-based unit root tests. Some likelihood ratio-based tests may mitigate this problem by considering special conditional variance structures like GARCH, but such approaches depend on some preassumed functional form, which is sometimes difficult to specify and may lead to incorrect conclusions if the model structure is not appropriately specified. This paper investigated these issues by testing 5 commodity futures data: corn, soybean, cotton, live cattle and lean hog futures prices. The tests are performed on three different sample frequencies: daily settlement price (high frequency), weekly and monthly price (constructed by skip intermediate observations on the daily data, low frequency). By applying a new developed Bayesian Model Averaging unit root test, we first showed that sample frequency has effects on the unit root test results. High frequency data contains more mean-reversion information which will be ignored by traditional frequentist tests like the Augmented Dickey-Fuller test.

Results from the empirical application offer several useful insights. First, specifying the conditional variance structure in the test improved the test results. Second, for unit root tests to rely on a single model specification can sometimes produce results that are not robust. This could be seen from the ADF test on soybean monthly data, by specifying different lags included in the test, completely different decisions are made on the stationarity of the series which could bring troubles to the succeeding analysis steps. Using BMA methods, however, the result is consistency across different sampling frequency, and it can average all the possible model candidates the researcher believes is

possible and then reached a comprehensive decision which solves the model specification problem in the unit root test. By examining the posterior weight of BMA methods, we find that the model weights will “self-adjust” to the models that may best describe the behavior of certain sample frequency to make sure they have higher posterior probability.

Overall, this paper shows that sampling frequency did matter to the unit root result so one should be cautious when performing the systematic sampling to get low frequency sample from the high frequency one. We also provided a BMA approach to test for unit roots to accommodate the high volatility and heavy tail behavior of the financial data. Our BMA test can handle model specification uncertainty issues well which provided more robustness and flexibility compared with other similar methods, and it also suggested an alternative to the researchers who are concerned this is a significant problem in their work.

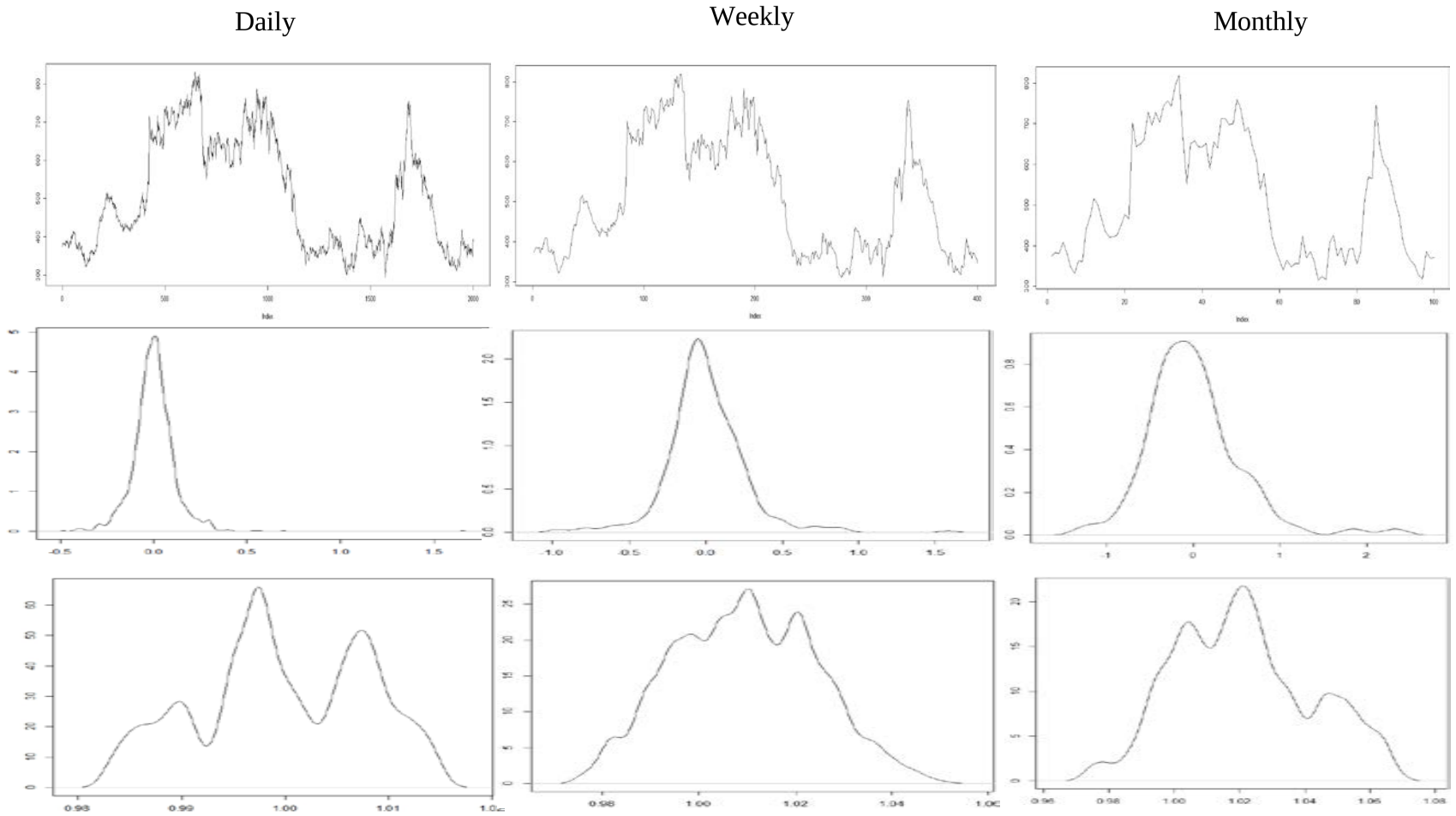


Figure 3.1 (i). Corn Futures Price. Top Row: Prices Time Series; Middle Row: Residual Density of AR (1) Model; Bottom Row: Averaged Posterior Density of the Dominant Root

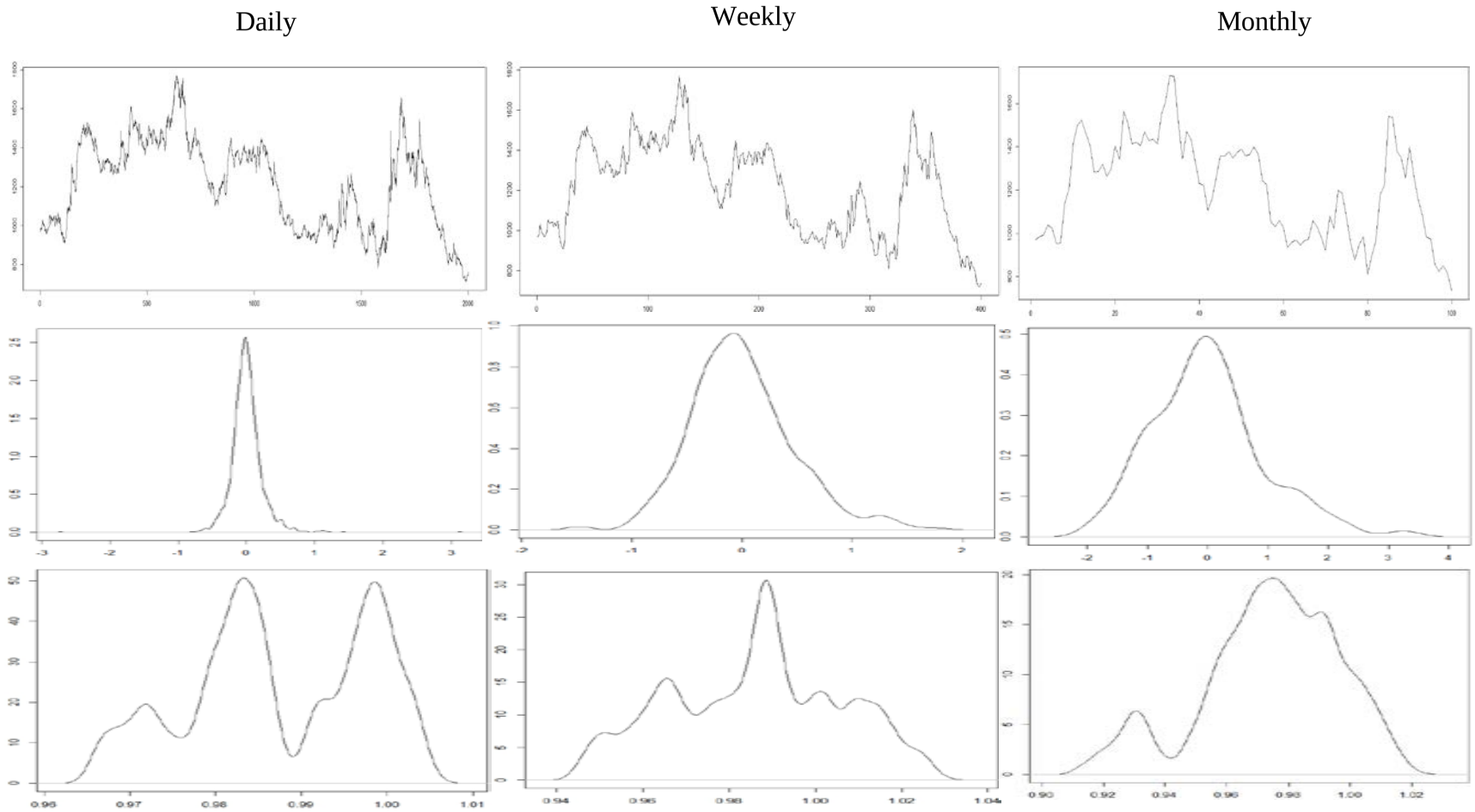


Figure 3.1 (ii). Soybean Futures Price. Top Row: Prices Time Series; Middle Row: Residual Density of AR (1) Model; Bottom Row: Averaged Posterior Density of the Dominant Root

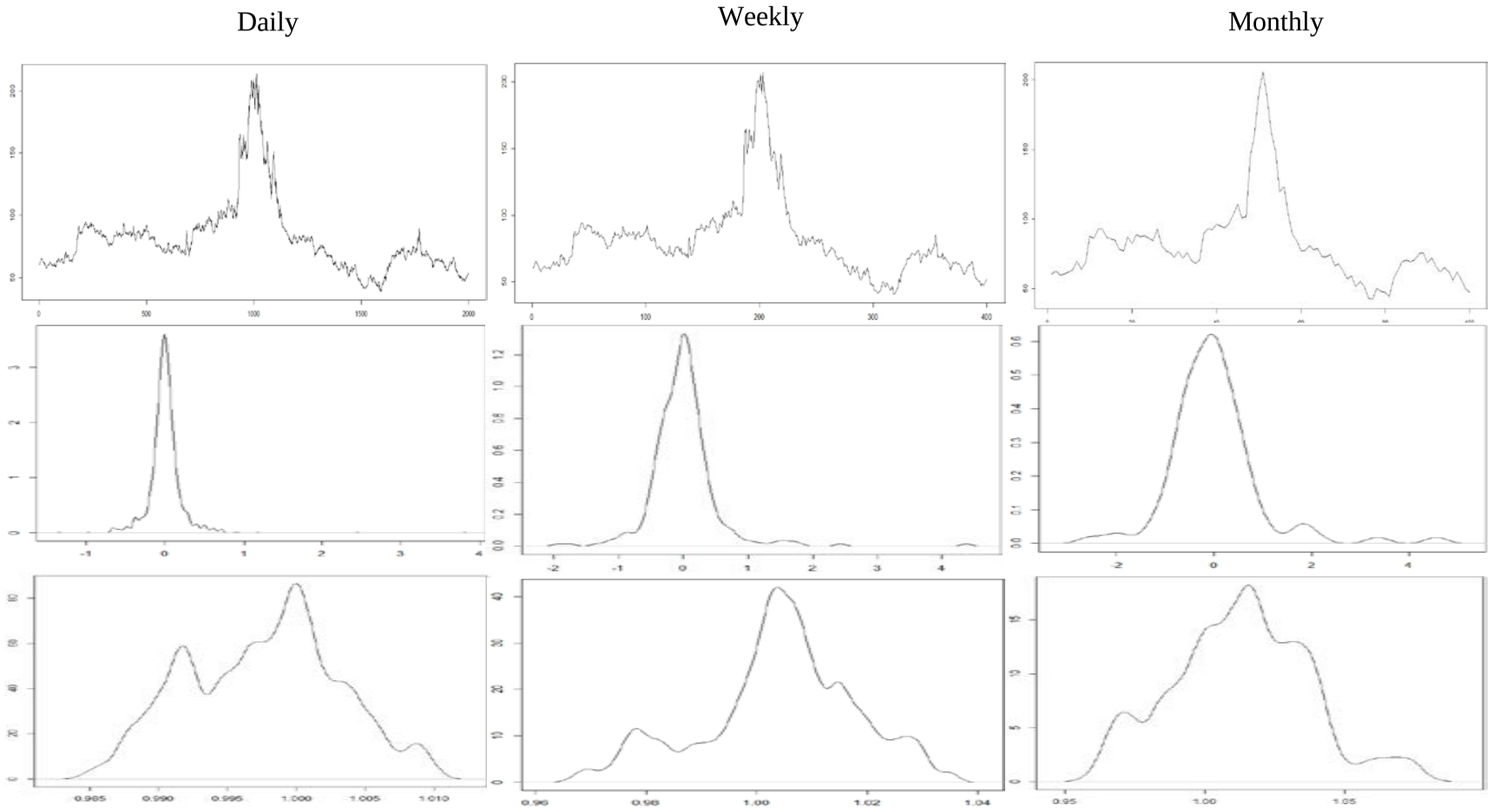


Figure 3.1 (iii). Cotton Futures Price. Top Row: Prices Time Series; Middle Row: Residual Density of AR (1) Model; Bottom Row: Averaged Posterior Density of the Dominant Root

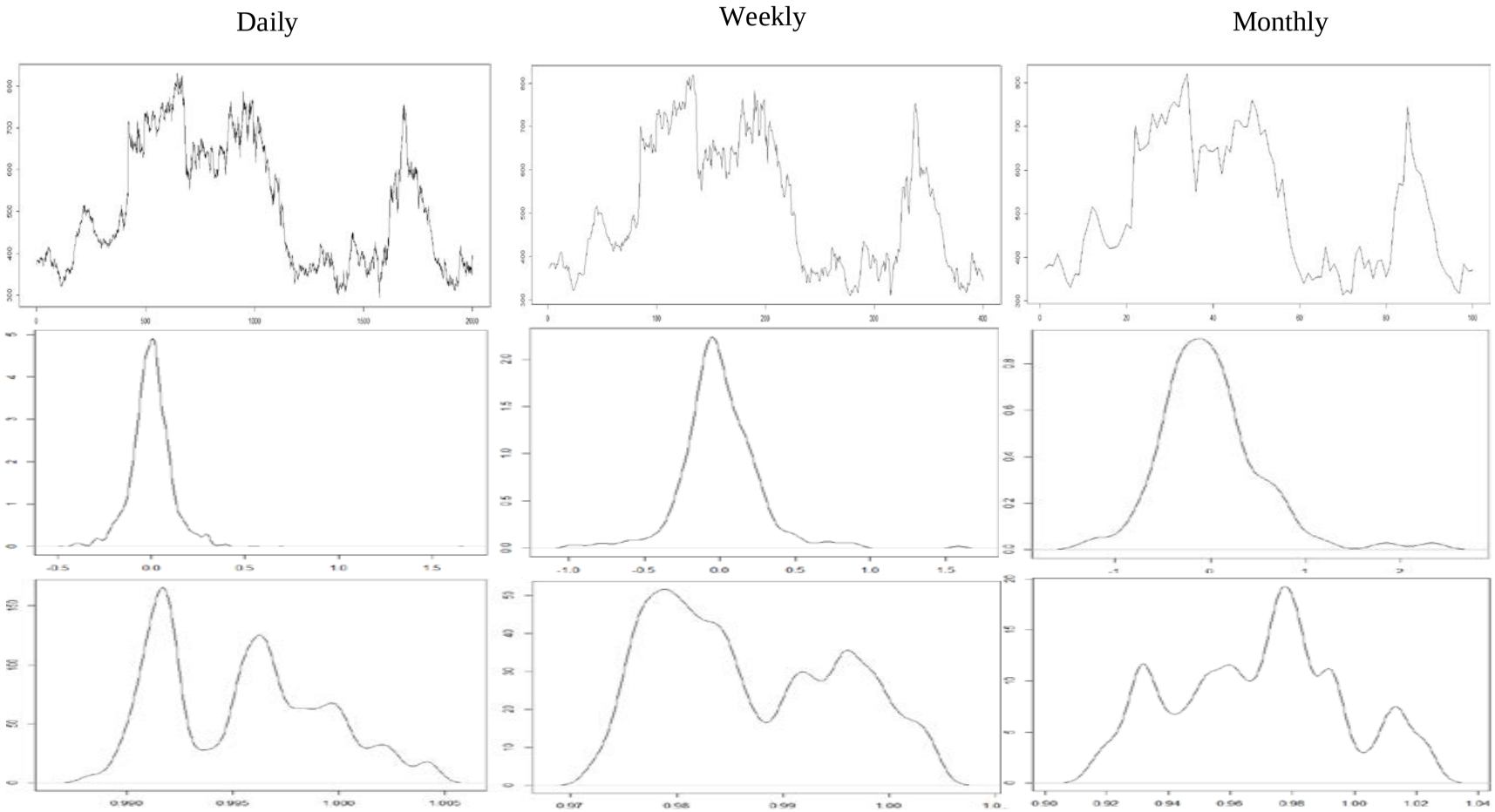


Figure 3.1 (iv). Live Cattle Futures Price. Top Row: Prices Time Series; Middle Row: Residual Density of AR (1) Model; Bottom Row: Averaged Posterior Density of the Dominant Root

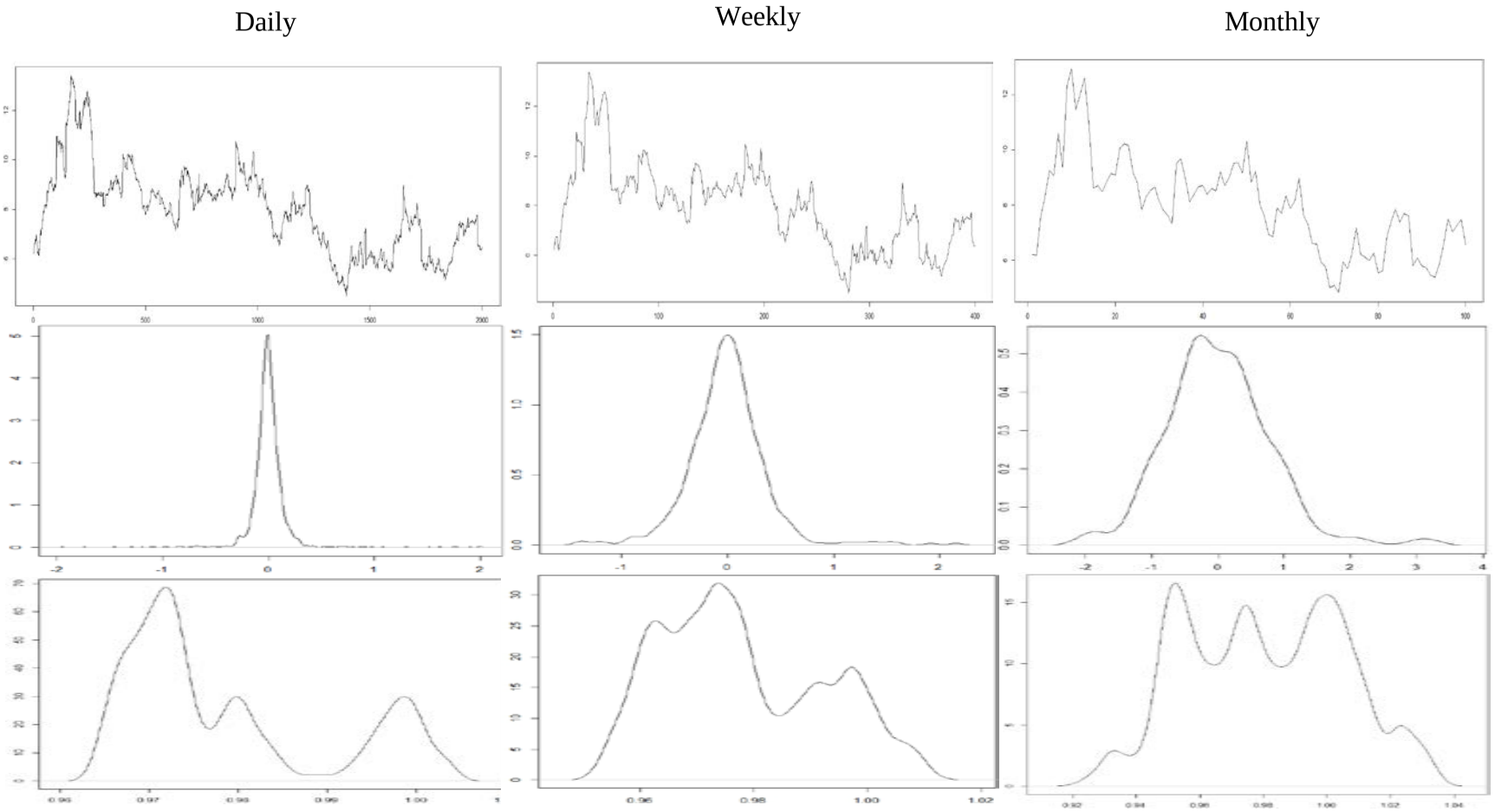


Figure 3.1 (v). Lean Hog Futures Price. Top Row: Prices Time Series; Middle Row: Residual Density of AR (1) Model; Bottom Row: Averaged Posterior Density of the Dominant Root

References

Avramov, Doron. 2002. “Stock Return Predictability and Model Uncertainty.” *Journal of Financial Economics* 64 (3): 423–58.

Barnard, G. A. 1963. “New Methods of Quality Control.” *Journal of the Royal Statistical Society. Series A (General)* 126 (2): 255–58. doi:10.2307/2982365.

Bauwens, Luc, Michel Lubrano, and Jean-Francois Richard. 2000a. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press. http://books.google.com/books?hl=en&lr=&id=hRV3XiQVWLUC&oi=fnd&pg=PA1&dq=Bayesian+inference+in+dynamic+econometric+models&ots=mJMR-mN8hn&sig=Yq1euDu-iE8DAzNeoBWaA0DIsAA.

———. 2000b. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press.

Berger, James O., and Ruo-yong Yang. 1994. “Noninformative Priors and Bayesian Testing for the AR(1) Model.” *Econometric Theory* 10 (3/4): 461–82.

Bhargava, Alok. 1986. “On the Theory of Testing for Unit Roots in Observed Time Series.” *Review of Economic Studies* 53 (3): 369.

Bollerslev, Tim. 1986. “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics* 31 (3): 307–27.

Boswijk, H. Peter. 2001. *Testing for a Unit Root with near-Integrated Volatility*. Tinbergen Institute Discussion Paper. http://www.econstor.eu/handle/10419/85890.

Boswijk, H Peter, and Franc Klaassen. 2012. “Why Frequency Matters for Unit Root Testing in Financial Time Series.” *Journal of Business & Economic Statistics* 30 (3): 351–57.

Box, George EP, and David R Cox. 1964. “An Analysis of Transformations.” *Journal of the Royal Statistical Society. Series B (Methodological)*, 211–52.

Brock, William A., Steven N. Durlauf, and Kenneth D. West. 2007. “Model Uncertainty and Policy Evaluation: Some Theory and Empirics.” *Journal of Econometrics* 136 (2): 629–64.

Campbell, John Y., and Pierre Perron. 1991. “Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots.” In *NBER Macroeconomics Annual 1991, Volume 6*, 141–220. MIT Press. http://www.nber.org/chapters/c10983.pdf.

Chambers, Marcus J. 3. “Testing for Unit Roots with Flow Data and Varying Sampling Frequency.” *Journal of Econometrics* 119 (1): 1–18. doi:10.1016/S0304-4076(03)00152-0.

Charitidou, Efstratia, Dimitris Fouskakis, and Ioannis Ntzoufras. 2013. “Bayesian Transformation Family Selection: Moving towards a Transformed Gaussian Universe.” *arXiv Preprint arXiv:1312.3482*. http://arxiv.org/abs/1312.3482.

Chatfield, Chris. 1995. “Model Uncertainty, Data Mining and Statistical Inference.” *Jorunal of the Royal Statistical Society* 158 (3): 419–66.

Choi, In. 1992. “Effects of Data Aggregation on the Power of Tests for a Unit Root: A Simulation Study.” *Economics Letters* 40 (4): 397–401.

Choi, In, and Bhum Suk Chung. 1995. “Sampling Frequency and the Power of Tests for a Unit Root: A Simulation Study.” *Economics Letters* 49 (2): 131–36.

Claeskens, Gerda, and Nils Lid Hjort. 2008. *Model Selection and Model Averaging*. 1 edition. Cambridge ; New York: Cambridge University Press.

Clemen, Robert T. 1989. “Combining Forecasts: A Review and Annotated Bibliography.” *International Journal of Forecasting* 5 (4): 559–83. doi:10.1016/0169-2070(89)90012-5.

Clyde, Merlise, and Edward I. George. 2004. “Model Uncertainty.” *Statistical Science* 19 (1): 81–94.

Cochrane, John H. 1988. “How Big Is the Random Walk in GNP?” *Journal of Political Economy* 96 (5): 893–920.

Cremers, K. J. Martijn. 2002. “Stock Return Predictability: A Bayesian Model Selection Perspective.” *Review of Financial Studies* 15 (4): 1223–49. doi:10.1093/rfs/15.4.1223.

DeJong, David N., John C. Nankervis, N. E. Savin, and Charles H. Whiteman. 1992. “The Power Problems of Unit Root Test in Time Series with Autoregressive Errors.” *Journal of Econometrics* 53 (1–3): 323–43. doi:10.1016/0304-4076(92)90090-E.

DeJong, David N., and Charles H. Whiteman. 1991. “Reconsidering ‘Trends and Random Walks in Macroeconomic Time Series.’” *Journal of Monetary Economics* 28 (2): 221–54.

Dickey, David A., and Wayne A. Fuller. 1979a. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association* 74 (366a): 427–31. doi:10.1080/01621459.1979.10482531.

———. 1979b. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association* 74 (366a): 427–31. doi:10.1080/01621459.1979.10482531.

———. 1981. “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root.” *Econometrica* 49 (4): 1057–72. doi:10.2307/1912517.

Dorfman, Jeffrey H. 1993. “Bayesian Efficiency Tests for Commodity Futures Markets.” *American Journal of Agricultural Economics* 75 (5): 1206–10.

Dorfman, Jeffrey H., and Myung D. Park. 2011. “Estimating the Risk-Return Tradeoff in Agribusiness Stocks: Linkages with the Broader Stock Market.” *American Journal of Agricultural Economics* 93 (2): 426–33. doi:10.1093/ajae/aaq158.

Dufour, Jean-Marie, and Maxwell L. King. 1991. “Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary AR (1) Errors.” *Journal of Econometrics* 47 (1): 115–43.

Eicher, Theo S., Christian Henn, and Chris Papageorgiou. 2012. “Trade Creation and Diversion Revisited: Accounting for Model Uncertainty and Natural Trading Partner Effects.” *Journal of Applied Econometrics* 27 (2): 296–321. doi:10.1002/jae.1198.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. 1992. *Efficient Tests for an Autoregressive Unit Root*. National Bureau of Economic Research Cambridge, Mass., USA. http://www.nber.org/papers/t0130.

Engle, Robert F. 1982. “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.” *Econometrica* 50 (4): 987–1007. doi:10.2307/1912773.

Engle, Robert F., and Victor K. Ng. 1993. “Measuring and Testing the Impact of News on Volatility.” *The Journal of Finance* 48 (5): 1749–78. doi:10.1111/j.1540-6261.1993.tb05127.x.

Fan, Jianqing, and Runze Li. 2001. “Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties.” *Journal of the American Statistical Association* 96 (456): 1348–60.

Fernandez, Carmen, Eduardo Ley, and Mark FJ Steel. 2001. “Model Uncertainty in Cross-Country Growth Regressions.” *Journal of Applied Econometrics* 16 (5): 563–76.

Fernández-Villaverde, Jesús, and Juan Francisco Rubio-Ramírez. 2004. “Comparing Dynamic Equilibrium Models to Data: A Bayesian Approach.” *Journal of Econometrics* 123 (1): 153–87. doi:10.1016/j.jeconom.2003.10.031.

Gelfand, Alan E., and Adrian FM Smith. 1990. “Sampling-Based Approaches to Calculating Marginal Densities.” *Journal of the American Statistical Association* 85 (410): 398–409.

Geman, Stuart, and Donald Geman. 1984. “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images.” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6: 721–41.

Geweke, John. 1992. “Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments.” In *Bayesian Statistics 4*, edited by Jose M.

Bernardo, James O. Berger, A.P. dawid, and J.F.M Smith, 169–93. Oxford University Press.

Griffiths, Charley Xia. 2012. *Bayesian Unit Root Testing: The Effect of Choice of Prior On Test Outcomes*. The University of Melbourne.

Hamilton, James Douglas. 1994. *Time Series Analysis*. Vol. 2. Princeton university press Princeton.

Hansen, Peter R., and Asger Lunde. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics* 20 (7): 873–89. doi:10.1002/jae.800.

Harris, R. I. D. 1992. "Testing for Unit Roots Using the Augmented Dickey-Fuller Test: Some Issues Relating to the Size, Power and the Lag Structure of the Test." *Economics Letters* 38 (4): 381–86. doi:10.1016/0165-1765(92)90022-Q.

Harvey, Andrew C, and AC Harvey. 1993. *Time Series Models*. Vol. 2. Harvester Wheatsheaf New York.

Hastings, W. K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57 (1): 97–109. doi:10.2307/2334940.

Hoeting, Jennifer A., David Madigan, Adrian E. Raftery, and Chris T. Volinsky. 1999. "Bayesian Model Averaging: A Tutorial." *Statistical Science* 14 (4): 382–401.

Jing, Liang. 2010. "Hastings-within-Gibbs Algorithm: Introduction and Application on Hierarchical Model." University of Texas at San Antonio. http://georglsm.r-forge.r-project.org/site-projects/pdf/Hastings_within_Gibbs.pdf.

Koop, Gary, Eduardo Ley, Jacek Osiewalski, and Mark FJ Steel. 1997. “Bayesian Analysis of Long Memory and Persistence Using ARFIMA Models.” *Journal of Econometrics* 76 (1): 149–69.

Koop, Gary, and Simon Potter. 2004a. “Forecasting in Dynamic Factor Models Using Bayesian Model Averaging.” *The Econometrics Journal* 7 (2): 550–65.

———. 2004b. “Forecasting in Dynamic Factor Models Using Bayesian Model Averaging.” *The Econometrics Journal* 7 (2): 550–65.

Koske, I., and I. Wanner. 2013. “The Drivers of Labour Income Inequality – an Analysis Based on Bayesian Model Averaging.” *Applied Economics Letters* 20 (2): 123–26. doi:10.1080/13504851.2012.683164.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin. 1992. “Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” *Journal of Econometrics* 54 (1): 159–78.

Leamer, Edward E. 1983. “Let’s Take the Con Out of Econometrics.” *The American Economic Review* 73 (1): 31–43.

Lee, Peter M. 2012. *Bayesian Statistics: An Introduction*. John Wiley & Sons.

Ling, Shiqing, and WK Li. 2003. “Asymptotic Inference for Unit Root Processes with GARCH (1, 1) Errors.” *Econometric Theory* 19 (04): 541–64.

Lopez, J. Humberto. 1997. “The Power of the ADF Test.” *Economics Letters* 57 (1): 5–10. doi:10.1016/S0165-1765(97)81872-1.

Lubrano, Michel. 1995. “Testing for Unit Roots in a Bayesian Framework.” *Journal of Econometrics* 69 (1): 81–109.

Lucas, André. 1995. “Unit Root Tests Based on M Estimators.” *Econometric Theory* 11 (02): 331–46.

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. “Equation of State Calculations by Fast Computing Machines.” *The Journal of Chemical Physics* 21 (6): 1087–92.

Min, Chung-ki, and Arnold Zellner. 1993. “Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates.” *Journal of Econometrics* 56 (1–2): 89–118. doi:10.1016/0304-4076(93)90102-B.

Mitchell, James, Nigel Pain, and Rebecca Riley. 2011. “The Drivers of International Migration to the UK: A Panel-Based Bayesian Model Averaging Approach*.” *The Economic Journal* 121 (557): 1398–1444.

Moral-Benito, Enrique. 2013. “Model Averaging in Economics: An Overview.” *Journal of Economic Surveys*.

Morales, Knashawn H., Joseph G. Ibrahim, Chien-Jen Chen, and Louise M. Ryan. 2006. “Bayesian Model Averaging with Applications to Benchmark Dose Estimation for Arsenic in Drinking Water.” *Journal of the American Statistical Association* 101 (473): 9–17.

Nelson, Charles R, and Charles R Plosser. 1982. “Trends and Random Walks in Macroeconmic Time Series: Some Evidence and Implications.” *Journal of Monetary Economics* 10 (2): 139–62.

Nelson, Daniel B. 1991. “Conditional Heteroskedasticity in Asset Returns: A New Approach.” *Econometrica: Journal of the Econometric Society*, 347–70.

Ng, Serena, and Pierre Perron. 1995. “Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag.” *Journal of the American Statistical Association* 90 (429): 268–81. doi:10.2307/2291151.

———. 2001. “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power.” *Econometrica* 69 (6): 1519–54.

Pagano, Marcello. 1974. “Estimation of Models of Autoregressive Signal plus White Noise.” *The Annals of Statistics*, 99–108.

Perron, Pierre, and Serena Ng. 1996. “Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties.” *The Review of Economic Studies* 63 (3): 435–63.

Phillips, Peter CB. 1991. “To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends.” *Journal of Applied Econometrics* 6 (4): 333–64.

Phillips, Peter C. B., and Pierre Perron. 1988. “Testing for a Unit Root in Time Series Regression.” *Biometrika* 75 (2): 335–46. doi:10.2307/2336182.

Phillips, Peter C. B., and Peter Schmidt. 1989. *Testing for a Unit Root in the Presence of Deterministic Trends*. Cowles Foundation Discussion Paper 933. Cowles Foundation for Research in Economics, Yale University. https://ideas.repec.org/p/cwl/cwldpp/933.html.

Phillips, Peter CB, and Peter Schmidt. 1989. *Testing for a Unit Root in the Presence of Deterministic Trends*. Cowles Foundation for Research in Economics, Yale University. http://ideas.repec.org/p/cwl/cwldpp/933.html.

Poirier, Dale J. 1991. “A Bayesian View of Nominal Money and Real Output through a New Classical Macroeconomic Window.” *Journal of Business & Economic Statistics* 9 (2): 125–48. doi:10.2307/1391778.

Raftery, Adrian E. 1995. “Bayesian Model Selection in Social Research.” *Sociological Methodology* 25: 111–64.

Roberts, G. O., and A. F. M. Smith. 1994. “Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms.” *Stochastic Processes and Their Applications* 49 (2): 207–16. doi:10.1016/0304-4149(94)90134-1.

Rothenberg, Thomas J, and James H Stock. 1997. “Inference in a Nearly Integrated Autoregressive Model with Nonnormal Innovations.” *Journal of Econometrics* 80 (2): 269–86.

Said, Said E., and David A. Dickey. 1984. “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order.” *Biometrika* 71 (3): 599–607. doi:10.2307/2336570.

Sala, Xavier, Gernot Doppelhofer Martin, and Ronald I Miller. 2004. “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach.” *The American Economic Review*.

Schotman, Peter C. 1994. “Priors for the AR(1) Model: Parameterization Issues and Time Series Considerations.” *Econometric Theory* 10 (3/4): 579–95.

Schotman, Peter C, and Herman K Van Dijk. 1991. “On Bayesian Routes to Unit Roots.” *Journal of Applied Econometrics* 6 (4): 387–401.

Schwert, G. William. 1989a. “Tests for Unit Roots: A Monte Carlo Investigation.” *Journal of Business & Economic Statistics* 7 (2): 147–59. doi:10.2307/1391432.

———. 1989b. “Tests for Unit Roots: A Monte Carlo Investigation.” *Journal of Business & Economic Statistics* 7 (2): 147–59. doi:10.2307/1391432.

Seo, Byeongseon. 1999. “Distribution Theory for Unit Root Tests with Conditional Heteroskedasticity.” *Journal of Econometrics* 91 (1): 113–44.

Shiller, Robert J, and Pierre Perron. 1985. “Testing the Random Walk Hypothesis: Power versus Frequency of Observation.” *Economics Letters* 18 (4): 381–86.

Sims, Christopher A. 1988. “Bayesian Skepticism on Unit Root Econometrics.” *Journal of Economic Dynamics and Control* 12 (2): 463–74.

Sims, Christopher A, and Harald Uhlig. 1991. “Understanding Unit Rooters: A Helicopter Tour.” *Econometrica: Journal of the Econometric Society*, 1591–99.

Stock, James H. 1994. “Unit Roots, Structural Breaks and Trends.” *Handbook of Econometrics* 4: 2739–2841.

Tanner, Martin A. 1998. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. 3rd edition. New York: Springer.

Tibshirani, Robert. 1996. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)* 58 (1): 267–88.

Wright, Jonathan H. 2008. “Bayesian Model Averaging and Exchange Rate Forecasts.” *Journal of Econometrics* 146 (2): 329–41.

Zellner, Arnold. 1996. *An Introduction to Bayesian Inference in Econometrics*. 1 edition. New York; Chichester: Wiley-Interscience.

Zivot, Eric. 1994. “A Bayesian Analysis of the Unit Root Hypothesis within an Unobserved Components Model.” *Econometric Theory* 10 (3-4): 552–78.

Zivot, Eric, and Jiahui Wang. 2007. *Modeling Financial Time Series with S-PLUS®*. Vol. 191. Springer. http://books.google.com/books?hl=en&lr=&id=Mby9u8_WXXgC&oi=fnd&pg=PA1&dq=Modeling+Financial+Time+Series+with+S-PLUS&ots=ppvXOD137o&sig=MznbxY3r1ntQm8-obUvNSfui9CA.