SPURIOUS REGRESSION PROBLEM IN KALMAN FILTER ESTIMATION OF TIME VARYING PARAMETER MODELS

A Master's Thesis

by
BURAK ALPARSLAN EROĞLU

Department of Economics
Bilkent University
Ankara
September 2010

SPURIOUS REGRESSION PROBLEM IN KALMAN FILTER ESTIMATION OF TIME VARYING PARAMETER MODELS

The Institute of Economics and Social Sciences
of
Bilkent University

by

BURAK ALPARSLAN EROĞLU

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS

in

THE DEPARTMENT OF
ECONOMICS
BILKENT UNIVERSITY
ANKARA

September 2010

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Asst. Prof. Dr. Taner Yiğit
Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Refet Gürkaynak
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Prof. Dr. Orhan Arıkan
Examining Committee Member

Approval of the Institute of Economics and Social Sciences

Prof. Dr. Erdal Erel
Director

To Day

ABSTRACT

SPURIOUS REGRESSION PROBLEM IN KALMAN FILTER ESTIMATION OF TIME VARYING PARAMETER MODELS

Eroğlu, Burak Alparslan

M.A., Department of Economics

Supervisor: Asst. Prof. Dr. Taner Yiğit

September 2010

This thesis provides a simulation based study on Kalman Filter estimation of time varying parameter models when nonstationary series are included in regression equation. In this study, we have performed several simulations in order to present the outcomes and ramifications of Kalman Filter estimation applied to time varying regression models in the presence of random walk series. As a consequence of these simulations, we demonstrate that Kalman Filter estimation cannot prevent the emergence of spurious regression in time varying parameter models. Furthermore, so as to detect the presence of spurious regression, we also propose a new method, which suggests penalizing Kalman Filter recursions with endogenously generated series. These series, which are created endogenously by utilizing Cochrane's variance ratio statistic, are replaced by state evolution parameter T_t in transition equation of time varying parameter model. Consequently, Penalized Kalman Filter performs well in distinguishing nonsense relation from a true cointegrating regression.

Keywords: Spurious regression, Cointegration, Time varying parameter models, Kalman Filter

ÖZET

ZAMANLA DEĞİŞEN PARAMETRE MODELLERİNİN KALMAN FİLTRESİYLE TAHMİNİNDE SAHTE İLİŞKİ PROBLEMİ

Eroğlu, Burak Alparslan
Yüksek Lisans, İktisat Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Taner Yiğit

Eylül 2010

Bu tez, durağan olmayan serilerin zamanla değişen parametre modellerine dahil edildiğinde Kalman Filtresi yöntemiyle tahmin edilmesi üzerine simulasyonlara dayanan bir çalışma sunmaktadır. Bu çalışmada, tümleşik serilerinin varlığında zamanla değişen regresyon modellerine uygulanan Kalman filtresi yönteminin sonuçlarını ve çıkarımlarını incelemek için çok sayıda similasyona baş vurulmuştur. Bu similasyonların sonucunda, Kalman filtresinin zamanla değişen parametre modellerinde Sahte ilişkinin ortaya çıkışını engelleyemediği gösterilmiştir. Ayrıca, bu sahte ilişkiyi tespit edebilmek için Kalman Filtresi yinelemelerini içsel olarak oluşturulmuş seriler yardımıyla cezalandırmayı öngören yeni yöntem önerilmiştir. İçsel olarak, Cochrane' in varyans oran istatistiği yardımıyla oluşturulmuş bu seriler, zamanla değişen parametre modelinin geçiş denklemindeki durum değişim parmetresi T_t yerine kullanılmıştır. Sonuç olarak, Cezalandırılmış Kalman Filtresi sahte ilişkinin gerçek bir eşgüdüm ilişkisinden ayrılması hususunda iyi bir performans göstermiştir.

Anahtar Kelimeler: Sahte ilişki, eşgüdüm, Zamanla değişen parametre modelleri, Kalman Filtresi

ACKNOWLEDGEMENTS

I would like to express gratitude to my supervisor Professor Taner Yiğit for his excellent guidance, encouragement and support. Without his assistance and support, I believe that I could not have been able to achieve so much.

I thank Professor Refet Gürkaynak for valuable suggestions and criticism. I also thank Professor Orhan Arıkan for his kindness in commenting on my thesis and for his excellent remarks.

I am very grateful to Şeyma Gün Eroğlu for her friendship. Her endless support and encouragement have been profound throughout the difficult times of M.A course.

I am also grateful to Koray Birdal, Yasin Dalgıç and Müge Diler for their friendship, assistance and support.

I would like to convey thanks to TUBİTAK-BİDEB for providing financial support during my studies in M.A. program.

I am also intdepted to Meltem Sağtürk for her invaluable assistance during M.A. program.

I am thankful to my beloved family, Mustafa Eroğlu, Meryem Eroğlu, İbrahim Alperen Eroğlu and Fatma Selenge Eroğlu for their support and trust.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In time series models, econometricians ineluctably encounter non-stationary series as dependent and independent variables. The presence of non-stationary series in regression models, however, may engender certain problems by vitiating the resulting estimation. Spurious regression, which can be defined as a nonsense relation between two unrelated unit root processes, is considered as one of these problems. Granger and Newbold (1974), in a precursory attempt, conducted several simulations to reveal the mechanics behind the emergence of Spurious regression. With the help of these simulations, the authors unveiled that OLS estimation produces a specious relationship between independently generated variables each following driftless random walk processes (Granger and Newbold, 1974). Moreover, they showed that the regression of two unrelated unit root processes will yield serially correlated or random walk residual term accompanied with high R^2. As a result, Granger and Newbold (1974) pointed out that the presence of serially correlated residuals may invalidate the usual inference procedures in estimation since this is a potential indicator of spurious regression. Subsequent to the cutting-edge study of Granger and Newbold, many researchers sought to explain spurious regression for different types of time series processes. These studies indicate that spurious regression is not peculiar to the regression of two independent driftless random walk processes. In chapter 2, we will briefly review different types of spurious regression.

On the other hand, as spurious regression literature was expanding, Granger and Newbold's pioneering work (1974) also inspired the emergence of the cointegration literature. Cointegration, which is first examined by Granger (1981), can simply be defined as a long term relationship between dependent variable and regressors. In his paper, Granger (1981) brings forth formal descriptions for the cointegration concept and cointegrated variables on nonstationary time series models. He states that if linear combination of nonstationary series, which corresponds to residual term in regression equation, is stationary, then these series are called cointegrated variables (Granger, 1981). In other words, the existence of the cointegrating relationship between variables depends on the presence of stationary residuals obtained from regression models. Moreover, the absence of cointegration results in nonstationary residuals and, inevitably, a very high R^2. This will lead to spurious relation between time series processes: a strong but fake relation between variables appears in estimation. Hence, there is a very adamant link between Spurious Regression and Cointegration. This link provides that Spurious regression problem is examined along with cointegration framework. Furthermore, Spurious regression cannot be directly realized because standard estimation methods (for instance OLS) fail to circumvent nonsense relation. Therefore, researchers seek to formulate formal tests to detect cointegration against spurious regression. Nevertheless, we will not go over these testing procedures in this paper since these tests are beyond the scope of our thesis.

Spurious regression and cointegration researchers mainly focused on regression models with time invariant coefficients though recently there are few attempts to model time varying cointegration. Hence, almost all inquiries on nonsense regression mentioned above, rule out time evolving parameter specification. However, fixed coefficient specification in econometrics is quite restrictive if structural break literature is considered (Bierens and Martins, 2010). The presence of structural breaks in true data generation process may weaken the inference with time invariant parameter models. This drives the scholars to practice upon time varying parameter models. What is more, the absence of cointegration in time invariant coefficient models also forces scholars to use time varying coefficient models. An early example of this situation is the paper of Canarella, Pollard and Lai (1990). In their article, they claim that the nonexistence of cointegration in Purchasing power parity models is due to time invariant parameter specification (Canarella, Pollard and Lai, 1990). Canarella, Pollard and Lai (1990) believe that PPP models can be captured better by time varying parameter models. However, using time varying parameter specification does not guarantee the circumvention of spurious regression problem. Consequently, a detailed analysis regarding the inspection of this problem in time varying parameter models is required. There have been some attempts to identify cointegration in time varying parameter models. However, current time varying cointegration models do not clearly explain how spurious regression emerges. In Chapter 2, we will discuss these cointegration models and their shortcomings in identifying spurious regression. Additionally, previous studies on time varying cointegration prefer to work on econometric methods other than Kalman Filter. Nevertheless, Kalman Filter is frequently used and a very powerful tool in time series econometrics (Harvey, 1989). We find it very curious that scholars have neglected Kalman Filter estimation in time varying cointegration framework. We believe that Kalman Filter should take the place it deserves in cointegration literature. In this study, we seek to present the outcome and ramifications of time varying parameter models estimation with Kalman Filter, in which non-stationary variables are included.

In order to scrutinize whether spurious regression emerges in Kalman Filter estimation of time varying parameter models, we will conduct series of simulations. In these simulations, we will adopt state space form (SSF). Once we put the time series model into SSF, we can apply Kalman Filter for estimating the time variant parameters (Harvey, 1989). On the other hand, these simulations will contain Kalman Filter estimation of two sets of random walk series generated with different structures. Basic difference in data generation structure stems from the existence of a cointegration relation between variables. In other words, our simulations will include both noncointegrated and cointegrated series in Kalman Filter estimation. We will also show that Kalman Filter cannot prevent the appearance of spurious regression in the estimation of noncointegrated series. After showing the emergence of the Nonsense regression in TVP models estimated by Kalman Filter, we seek to develop a new method which aims to detect Spurious Regression in Kalman filter. This new method will suggest penalizing Kalman filter equations so as to prevent Kalman Filter to create nonsense time varying relation for noncointegrated series. We will apply this new method to both cointegrated and noncointegrated series. Thus, we can explore the performance of our modified Kalman Filter under different data generation processes.

The plan of the paper is as follows. Chapter 2 provides literature review of spurious regression, cointegration models, time varying parameter models in time series econometrics and time varying cointegration respectively. This section is needed in order to understand the nature of the spurious relation in time series. In chapter 3, we review State-space form and Kalman Filter. After describing these concepts, we will discuss how one can utilize them in time varying parameter models in econometrics. So as to designate spurious regression some simulations are conducted with different data generation processes in Chapter 4. After showing spurious regression cannot be avoided by Kalman Filter algorithm itself, Chapter 5 introduce a new method to detect and prevent nonsense regression in time varying parameter models and some conclusions are drawn in Chapter 6.

CHAPTER 2

LITERATURE SURVEY

2.1. Spurious Regression and Cointegration Models

One of the core topics of this thesis is Spurious regression, which has been examined by surprisingly few researchers. Among these researchers, Yule (1926) is the first who professes the significance of nonsense correlation. He reports a correlation of 0.95 between the proportion of church marriage to all marriages and the mortality rate over 1866-1911 in England. He believes that both variables share a common factor, which influences them at the same time. He concludes that the correlation between these variables is nonsense. Furthermore, it was 1971 when spurious regression was in consideration among the scholars again. In their article, Box and Newbold indicates that some nonsense relation may emerge if sufficient care is not taken (1971). Even though Box and Newbold do not clearly identify the problem, Granger and Newbold take the first step to analyze spurious regression in 1974. They perform Monte Carlo simulations in order to illustrate nonsense regression in time series models. In these simulations, they consider two uncorrelated random walks without drift processes generated according to following equation:

$$\begin{aligned} y_t &= y_{t-1} + u_t, \quad u_t \sim iin\left(0, \sigma_u^2\right) \\ x_t &= x_{t-1} + v_t, \quad v_t \sim iin\left(0, \sigma_v^2\right) \end{aligned} \tag{1}$$

where u_t and v_t are both serially and mutually uncorrelated. After Granger and Newbold (1974) generates independent random walk series, they run the following regression equation:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \tag{1}$$

In the estimation of above regression equation, R^2 is expected to tend to zero, since y_t and x_t are uncorrelated random walk processes (Maddala and Kim, 1998). However, Granger and Newbold (1974) argued that this may not be the case. They report a very high R^2 accompanied with autocorrelated residual terms (Granger and Newbold, 1974). Furthermore, they calculate the test statistics for the significance of $\hat{\beta}_1$ by:

$$S = \frac{\left|\hat{\beta}_1\right|}{S.E(\hat{\beta}_1)} \tag{2}$$

After 100 simulations, Granger and Newbold observe that $\hat{\beta}_1$ is significant in 77 simulations in which variables are independent random walks (1974). From their simulations, they conclude that "if one's variables are random walks, or near random walks, and one includes in regression equations variables which should in fact not be included, then it will be the rule rather than the exception to find spurious relationships" (Granger and Newbold, 1974). On the other hand, the findings of Granger and Newbold inspire Philips (1986). Philips develops an analytical framework for Spurious regression. In his paper, Philips supports Granger and Newbold's empirical findings by using an asymptotic theory for regression of nonstationary series. Among the cardinal results of this paper are; limiting distribution of OLS estimator does not converge to a constant as sample size goes to infinity, and conventional test statistics for significance of estimates diverges in such regressions of driftless random walks (Philips, 1986). Therefore, Philips's asymptotic theory suggests that customary tests for coefficient significance are not valid under limiting distribution (1986). Consequently, Philips and Granger-Newbold's findings can be summarized as follows: if one include independent random walk variables in their estimation, it is inevitable that parameters in the model appear to be significant and there is high degree of fit.

Both studies of Granger and Newbold, and Philips focused on random walks without drift variables. However, econometricians deal with random walk with drift processes in some time series models. Entorf (1992) considers these processes and examine them in spurious regression framework. He claims that the results of spurious regression will differ if the independent random walks with drift processes are included in regression model (Entorf, 1992). In order to prove his claims, he considers the following data generation process:

$$\begin{aligned} y_t &= \alpha_y + y_{t-1} + u_t \\ x_t &= \alpha_x + x_{t-1} + v_t \end{aligned} \qquad (3)$$

Using this data generation process, Entorf points out that in regression of y_t on x_t, $\hat{\beta}_1$ converges to α_y / α_x in probability, instead of a random variable as in random walk without drift scenario (1992). As a consequence, the results of spurious regression will be different in accordance with the existence of drift in data generation processes (Maddala and Kim, 1998).

However, the studies on spurious regression do not remain limited to regression of two independent random walk processes. In their article, Nelson and Kang (1984) take the nonsense regression framework out of the conventional discussion. They argue that the regression of a random walk without drift process against a time trend will also result in the inappropriate inference about significance of trend coefficient (Nelson and Kang, 1984). With this study, they also show spurious regression is not peculiar to independent random walk processes. Furthermore, Philips and Darlauf, in 1988, examine asymptotical features of Nelson and Kang's findings. They provide an asymptotic theory for the regression coefficients in models which attempt to regress I(1) process on time trend.

Another attempt to delineate spurious regression for different types of time series variables again belongs to Philips. It was 1988 when Philips was the first who introduced near integrated processes and included these processes in Spurious regression framework. In his article, he mentioned that regressions with near integrated processes have similar properties as the regression with I(1) processes (Philips, 1988). On the other hand, in 1999, Tsay and Chung extended spurious regression literature by analyzing fractionally integrated processes which are the generalization of the I(1) process that exhibit a broader long-run characteristics (Tsay and Chung, 1999). Their findings indicated that when a long memory fractionally integrated process is regressed on another independent long memory fractionally integrated process, no matter whether these processes are stationary, spurious regression occurs (Tsay and Chung, 1999). Moreover, Spurious regression with stationary series was also examined by Granger, Hyung and Jeon in 2001. They found that nonsense regression emerges when positively autocorrelated autoregressive series or long moving average processes are included in estimation equation (Granger, Hyung and Jeon, 2001). Additionally, another interesting study comes from Kim, Lee and Newbold, who showed that when Independent I(0) series with linear trends are included in OLS estimation, spurious relation occurs inevitably (2003).

On the other hand, as spurious regression literature was expanding, studies on nonsense regression also inspired the development of the cointegration literature. This literature took its root in the pioneering work of Granger in 1981. In his study, Granger (1981) propounded the concept of cointegrated variables: If linear combination of two $I(1)$ variables is $I(0)$, then these two variables are cointegrated. This concept has been adopted by many researchers for a long time. However, the recent cointegration models offer a different perspective to the cointegration framework. These models seem to depart from conventional cointegration studies by means of the assumptions made for residual terms, which are no longer supposed to be stationary. Some important examples of these models are discussed below:

Hansen (1991) made a well-founded attempt to identify different types of cointegration structure between econometric time series. In his paper, Hansen (1991) examines the estimation and inference procedure of cointegrated regression models in which error terms are displaying nonstationary variance. This feature is absolutely different from the properties of classical cointegration models. Moreover, Hansen (1991) points out that the cointegration identification of Granger (1981) does not fully cover all nonstationary models in economic theory. For empirical evidence, Hansen (1991) explores some famous cointegrating regressions by using sample split variance test. From these tests, the author concludes that the variance of cointegrating regressions may exhibit time varying structure, which violates the construction of classical cointegration relation. He attempts to reconstruct new cointegration framework by relaxing the usual variance properties, which can be listed as having constant mean and bounded variance. Furthermore, Hansen (1991) defines Bi-integrated (BI) process which can be described as follows: w_t is called bi-integrated if it is the multiplication of two series namely σ_t and u_t, where

$u_t \sim I(0)$ and $\sigma_t \sim I(1)$. After defining BI process, Hansen (1991) also propounds the notion of heteroscedastically cointegration. In this study, the pair (y_t, x_t) is defined to be heteroscedastically cointegrated if x_t has a unit root and w_t follows BI process in the model: $y_t = \beta' x_t + w_t$. Hansen (1991) also considers the time varying parameter framework with $y_t = \beta_t x_t + u_t$, where $x_t \sim I(1)$ and $\beta_t \sim I(0)$. He argues that this regression model also becomes HCI model with extra error term u_t (Hansen, 1991). We will use this structure in data generation processes of our simulations.

Harris, McCabe and Leybourne (2002) reconsider Hansen's Heteroskedastic cointegration model which allows only Heteroskedastically integrated dependent variable and conventionally integrated regressors. The authors point out that this asymmetry may cause some problems in estimation. One of these problems emerges in the form of an inconsistency of the OLS estimators in Hansen's framework unless above asymmetry condition on the integration structure of variables is imposed (Harris, McCabe and Leybourne, 2002). In order to remove these problems, they simply relax the restrictions put in Hansen's paper and bring forth stochastic cointegration model.

Moreover, Harris, McCabe and Leybourne (2006) modify their stochastic cointegration framework by including a time trend into data generation process. However, the main contribution of the paper is a test for stochastic cointegration against the alternative of no cointegration. In addition they propose a secondary test for stationary cointegration against the Heteroskedastic alternative. Thus, they conduct two tests for determining whether the model is cointegrated. Moreover, even though there is no direct application of these tests to time varying parameter models, papers of HML brought a different perspective to cointegration literature.

2.2. Time Varying Parameter Models in Econometrics and Time varying Cointegration

Another focal point of our study is the time varying parameter models in time series econometrics. The time varying parameter models became popular among econometricians as an alternative of fixed coefficient models. Especially, some particular situations led theses scholars to avoid using constant parameter models. These situations, which are reported to cause parameter variation, are listed by Tanizaki (1999) as structural changes, specification errors, nonlinearities, proxy variables and aggregation. Another reason for parameter variation was implied by Lucas (1976). Lucas (1976) argues that policy changes will systematically alter the structure of the econometric models, which makes econometric models inappropriate tools for long term economic evaluation. Accordingly, under these conditions, scholars consider the parameters of the regression model as a function of time. Cooley and Prescott (1976) are among those scholars who attempted to build an econometric model with stochastic parameter variation. They addressed the source of parameter variation to both misspecification and underlying economic theory (Cooley and Prescott, 1976).

After the study of Cooley and Prescott, some researchers combine time varying parameter models with nonstationarity. Kitawa is one of those authors, who use nonstationary series in time varying parameter models. In his article, Kitawa (1987) suggests a non-Gaussian state space approach for modeling non-stationary time series. He contended that it is not adequate to use a Gaussian State space form for the time series that has a mean value function with abrupt and gradual change (Kitawa, 1987). Concomitantly, he proposed a model, in which disturbance terms of measurement and transition equations are not Gaussian (Kitawa, 1987). Finally, he concluded that a non-Gaussian model might help dealing with variety of time series (non-stationary, non-linear…etc.) (Kitawa, 1987). However, there is nothing in Kitawa's study about spurious regression which is the result of inclusion of independent non-stationary series in the model. On the other hand, WanChun and Lun, in 2009, theorized a time varying parameter framework for autocovariance non-stationary time series. They also neglect potential spurious regression problem.

Another important study, which unites nonstationarity and TVP framework, belongs to Canarella, Pollard and Lai (1990). They analyze Purchasing Power Parity model, which contains nonstationary series such as exchange rates and relative prices. In their work, they examine the cointegration property of these series by using TVP approach (Canarella, Pollard and Lai, 1990). They conclude that a TVP approach supports PPP better in the presence of structural instability in the long run equilibrium relationship between exchange rates and prices. Furthermore, Ramajo

(2001) used time varying parameter specification in context of error correction models for estimating money demand relation in Venezuela. He allows variation in parameters of error correction model (ECM) different from classical ECM (Ramajo, 2001). However, both Canarella et al. and Ramajo ignore risk of Spurious regression in their models.

Recently there have been some attempts to analyze time varying cointegration after the analysis of nonstationary series in time varying coefficient models. Park and Hahn's study (1999) appears to be one of these attempts which focuses on time varying cointegration. In their article, the authors criticize the conclusion of cointegration papers in favor of the absence of cointegration. Park and Hahn (1999) claim that the absence of cointegration may be due to misspecification of cointegration models. They believe that this misspecification can be removed by adopting time varying cointegration model in which parameters are smooth functions evolving across time (Park and Hahn, 1999). Therefore, they proposed a cointegration model, which can be characterized as the following equations:

$$y_t = \alpha_t x_t + u_t \tag{5}$$

$$x_t \sim I(1) \text{ and } \alpha_t = \alpha\left(\frac{t}{n}\right) \tag{6}$$

where $\alpha(\cdot)$ is a smooth function defined on [0,1]. In order to estimate time varying coefficient α_t, Park and Hahn (1999) adopts a nonparametric (series estimator) method. The resulting estimator is consistent. Moreover, the estimator is efficient if the regressors are exogenous. However, this estimator has slower convergence rate than the OLS estimator in classical cointegration regressions (Park and Hahn, 1999). On the other hand, in the last part of the paper, Park and Hahn (1999) conduct a test of time varying cointegration against time invariant cointegration by using variable addition approach introduced by Park (1990). They also test their specification against no cointegration with nonstationary unit roots.

Another time varying cointegration paper is from Bierens and Martins. Bierens and Martins (2010) exposit another way to handle time varying cointegration. In their paper, time varying cointegrating vector error correction model (TV VECM) is analyzed under the assumption that cointegrating vector varies smoothly. This smooth cointegrating vector is approximated by the linear combinations of Chebyshev Time Polynomials. Furthermore, Bierens and Martins (2010) consider two different tests for cointegration. The first one tests the null of TV cointegration against time invariant cointegration. The second test contains the null hypothesis of TV cointegration against no cointegration. Hence, these testing procedures carry similar patterns as in Park and Hahn's (1999) paper.

To sum up, Econometricians have dealt with various types of spurious regression in linear fixed-parameter models. Some researchers suggest time varying cointegration models, in which cointegrating parameter vector varies smoothly over time. They also used residual based time varying cointegration tests. In these tests, they assume stationary residuals in time evolving cointegration models. However, it is seen that the time varying cointegration papers neglect the criticism of Hansen (1999) about possibility of having I(1) residuals in cointegration regression. On the other hand, even though Granger (1991) did not fully explain the issue, he tried to attract attention to uninteresting I(0) case, in which noncointegrated systems may have also I(0) error terms if time varying estimation methods are adopted. Finally, none of the studies on time varying cointegration use Kalman Filter as estimation algorithm. Thus, we will try to analyze whether spurious regression is also a problem in time variant coefficient estimation of Kalman Filter.

CHAPTER 3

REVIEW OF KALMAN FILTER ESTIMATION IN TIME SERIES

This chapter will present Kalman Filter algorithm and its application to time series models. First, we will go over state-space representation and Kalman Filter equations, then describe how Kalman filter can be applied to time series models. There will be also a section for the initialization of Kalman Filter algorithm.

3.1. State-space Form and Kalman Filter

State space form (SSF) is a powerful tool to handle a wide range of time series models (Harvey, 1989). Once we represent the econometric model in state-space form, Kalman Filter may be applied in order to obtain algorithms for prediction and smoothing (Harvey 1989). Additionally, there are various applications of Kalman Filter in Gaussian models. Kalman Filter provides means of constructing likelihood function by prediction error decomposition for these models (Harvey 1993). Moreover, time varying coefficient models can be represented in State space form, thus be estimated by Kalman Filter.

Before describing the state space representation of time varying parameter models, we will discuss the general features of SSF. General SSF applies to a multivariate time series y_t, containing N elements (Harvey, 1989). However, in this study, we concern with univariate model instead of multivariate series. Hence, we reduce the dimension of y_t to one. Moreover, Harvey (1989) notes that SSF consists of two basic components namely measurement and transition equations. Measurement equation relates observable y_t to $m \times 1$ state vector β_t:

$$y_t = Z_t \beta_t + d_t + \varepsilon_t \quad t = 1, \dots.., n \tag{7}$$

where Z_t is $1\times m$ vector of observable variables, d_t is deterministic component and ε_t is serially uncorrelated disturbance term with zero mean and variance H_t. Following equation summarizes the properties of disturbance term of measurement equation:

$$E\left(\varepsilon_t\right)=0 \text{ and } \operatorname{var}\left(\varepsilon_t\right)=H_t \text{ t=1,......,n} \tag{8}$$

Furthermore, the transition equation determines the evolution of state vector β_t via:

$$\beta_t = T_t\beta_{t-1} + c_t + R_t\eta_t \text{ t=1,....,n} \tag{9}$$

T_t is $m\times m$ matrix, c_t is $m\times 1$ vector of deterministic components in state, R_t is a $m\times g$ matrix and η_t is a $g\times 1$ vector of serially uncorrelated disturbance terms with mean zero and covariance matrix Q_t i.e.

$$E\left(\eta_t\right)=0 \text{ and } \operatorname{var}\left(\eta_t\right)=Q_t \tag{10}$$

This specification is complete after we impose two more assumptions:

1.) Initial state vector has mean b_0 and covariance matrix P_0, that is:

$$E\left(\beta_0\right)=b_0 \text{ and } \operatorname{var}\left(\beta_0\right)=P_0 \tag{11}$$

2.) The disturbance terms η_t and ε_t are uncorrelated with each other and initial state vector in all time periods:

$$E\left(\varepsilon_t\eta_s'\right)=0 \ \forall \ s,t=1,.....,n$$

and (12)

$$E(\varepsilon_t\beta_0')=0 \text{ and } E(\eta_t\beta_0')=0 \ \forall \ t=1,......,n$$

After we put the model in SSF, we can estimate the unknown state vector (time varying coefficients in regression analysis) with Kalman Filter algorithm. Kalman Filter is a recursive procedure that computes the optimal estimator of state vector at time t, based on information available at time t (Harvey, 1989). On the other hand, there are important assumptions of Kalman Filter algorithm. These assumptions can be listed as following:

1.) The system matrices H_t, Q_t, R_t, c_t, d_t and T_t together with a_0 and P_0 are assumed to be known in all time periods. (System matrices can also be estimated)

2.) The disturbance terms in measurement and transition equations are normally distributed

Under these assumptions, the mean of the conditional distribution of α_t is an optimal estimator of α_t in sense of that it minimizes the mean square error (MSE) (Harvey, 1989). Following section illustrates the essence of Kalman Filter algorithm.

3.1.1. General Form of Kalman Filter

Consider the SSF identified by equations (7) and (9). First, we can define a_{t-1} as optimal estimator of β_{t-1} based on information up to and including y_{t-1} (Harvey, 1989). In addition, P_{t-1} denotes covariance matrix of estimation error (Harvey, 1989), i.e.

$$P_{t-1} = E\left[(\beta_{t-1} - a_{t-1})(\beta_{t-1} - a_{t-1})'\right] \tag{13}$$

Given a_{t-1} and P_{t-1}, optimal estimator of β_t is given by *prediction equations* (14):

$$a_{t|t-1} = T_t a_{t-1} + c_t \tag{14}$$

where the covariance matrix of estimation error is as following:

$$P_{t|t-1} = T_t P_{t-1} T_t' + R_t Q_t R_t \quad \forall\ t = 1, \ldots.., n \tag{15}$$

These equations can be interpreted as the prior estimator and estimate error covariance of α_t based given knowledge of the process prior to step t (Welch and Bishop, 2001). On the other hand, these prior estimates of α_t will be updated by the *updating equations* of Kalman filter, once the new observation y_t is available. These *updating equations* are described by Harvey (1999) as follows:

$$a_t = a_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1} (y_t - Z_t a_{t|t-1} - d_t) \tag{16}$$

and

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1} \tag{17}$$

where,

$$F_t = Z_t P_{t|t-1} Z_t' + H_t, \quad t = 1, \ldots\ldots, n \tag{18}$$

(16) is posteriori estimate of α_t and (17) is estimate error covariance matrix of a_t. Finally, prediction and updating equations make up the Kalman Filter (Harvey, 1989). Furthermore, given the initial conditions a_0 and P_0, Kalman Filter algorithm described above gives optimal estimator of state vector as each new observation becomes available (Harvey, 1989).

3.1.2. Initialization of Kalman Filter

In order to start Kalman Filter recursions, one requires initial values for state vector and variance-covariance matrix of estimation error. Hamilton (1994b) suggest using unconditional mean and associated MSE for initial conditions a_0 and P_0 respectively. These values can be obtained by equations 19 and 20.

$$a_0 = E[\beta_1] = 0 \tag{19}$$

$$P_0 = E\left\{[\beta_1 - E(\beta_1)][\beta_1 - E(\beta_1)]'\right\} = [I - T_1 \otimes T_1]^{-1} vec(Q) \tag{20}$$

where I is identity matrix with dimension $m \times m$, $\otimes$ is Kronecker product and $vec(.)$ is vectorization operator. However, this method is only valid for stationary state vector. Note that in equation (19) the term $[I - T_1 \otimes T_1]^{-1}$ will be divergent if state is nonstationary. Under this condition, equations (17) and (18) are not appropriate initial conditions for nonstationary state vector. Thus, initial conditions should be redefined. At this point, Harvey claims that a diffuse prior should be replaced by initial values calculated by unconditional mean and MSE. In this thesis, we adopt the diffuse prior algorithm devised by de Jong (1988, 1989). Details of calculations of diffuse prior are beyond the scope of this study, so it is better to skip it.

3.2. Application of Kalman Filter to Time Varying Parameter Models

In previous sections, we describe very general properties of Kalman Filtering algorithm. However, in this study we are interested in time varying parameter models. In order to apply Kalman filter to time varying parameter regression models, we should first put the model into state-space form. The state space form of time varying parameter model with dependent variable y_t and single regressor x_t is given by equations 21 and 22.

$$y_t = \beta_t x_t + d_t + \varepsilon_t, \quad \varepsilon \sim iin(0, H_t) \tag{21}$$

$$\beta_t = T_t \beta_{t-1} + c_t + \eta_t, \quad \eta \sim iin(0, Q_t) \tag{22}$$

In above equations, we impose some specifications on system matrices H_t, Q_t, R_t, c_t, d_t, Z_t and T_t. First of all, we take the system matrices H_t, Q_t, R_t, c_t, d_t and T_t as time invariant. Moreover, we will consider time varying parameter (TVP) models with single regressor. Therefore, we replace Z_t with single dimensional observable vector x_t. Since dimension of x_t is equal to one, the state β_t is also single dimensional. Consequently, the time invariant system matrices H_t, Q_t, R_t, c_t, d_t and T_t become single valued in time varying parameter model we proposed. On the other hand, System matrices H_t, Q_t, R_t, c_t, d_t and T_t may depend on unknown parameters. In this case, we will estimate these unknown parameters via Maximum likelihood estimation. Otherwise, we assign values for these matrices. This assignment will be based on prior information about the model and economic intuition.

CHAPTER 4

SPURIOUS REGRESSION IN KALMAN FILTER

Spurious regression has been considered as a problem emerging in estimation of the models with time invariant parameters. However, it is still ambiguous whether we can encounter spurious regression in the time varying parameter models when Kalman Filter is preferred as the estimation method. At this point, we guess that if we regress two independent random walk processes, Kalman Filter will produce a specious state vector. We presume this nonsense state vector appears because Kalman Filter recursions always try to extract information about y_t with knowledge of x_t whatever the relation between y_t and x_t is. Consequently, Kalman Filter will produce Spurious regression in regression of unrelated series with time varying parameter models. In order to demonstrate the occurrence of this nonsense relation in Kalman Filter we will perform series of simulations. In these simulations, we run Kalman Filter to estimate the models containing independent and related random walk variables. Additionally, all of these simulations will be conducted in Gauss programming language.

Moreover, this chapter, which is devoted to investigate the occurrence of spurious regression, will contain two major sections. In the first section, we illustrate the estimation results of Kalman filter with known system matrices. In this estimation method, we require to assign values for the system matrices before we initiate Kalman Filter recursions. As we will exhibit, the assignment of these system matrices, especially state evolution parameter T_t, play very important role in Kalman Filter estimation. In the second part, we assume that the system matrices of the model are unknown parameters. We estimate theses parameters via Maximum likelihood estimation. On the other hand, 4 different types of data generation processes are utilized in this thesis. Next section will describe these Data generation procedures.

4.1. Data Generation Processes

In our simulations, 4 types of data generation processes (DGP) are concerned. In all of these types, we assume that the model contains one independent variable, and one dependent variable denoted as x_t and y_t respectively. These DGPs can be defined as follows:

1) We intend to demonstrate emergence of spurious regression by independently generated y_t and x_t series. General form of this type of data is given as:

$y_t = \alpha_y + y_{t-1} + \varepsilon_t^y$ and $x_t = \alpha_x + x_{t-1} + \varepsilon_t^x$ where ε_t^x and ε_t^y are independent random drawings from standard normal distribution. α_x and α_y drift terms of x_t and y_t respectively.

As it can be seen above, there is no valid relation between these series. Thus, it is expected that the estimated state vector in Kalman filter estimation of these series should be zero for all periods. However, we will show this is not the case.

2). In the second DGP, we will consider typical linear cointegration model. With this structure, we will explore implications of Kalman Filter estimation of linearly cointegrating variables. These variables can be described as follows:

The regressor x_t is generated as a random walk process: $x_t = \alpha_x + x_{t-1} + \varepsilon_t^x$. Further, y_t is generated as a linear function of x_t: $y_t = \alpha_y + \beta x_t + \varepsilon_t^y$. Consequently, y_t also follows random walk process. Additionally, ε_t^x and ε_t^y are independent drawings from standard normal distribution.

3) In the third data generation process, we consider another type of related series. However, this time, dependence between y_t and x_t is provided by time varying parameter vector in which the time variation is governed by AR(1) process.

Furthermore, The regressor x_t and the parameter vector β_t can be described as follows:

$$\begin{aligned} &x_t = \alpha_x + x_{t-1} + \varepsilon_t^x \\ &and \\ &\beta_t = \beta_x + \kappa\beta_{t-1} + \varepsilon_t^\beta \end{aligned} \tag{23}$$

In equation (22), We can observe that x_t follows I(1) process and β_t follows I(0) process. Thus, the multiplication $x_t\beta_t$ follows BI(1) process defined by Hansen (1991). Finally, equation (23) determines the relation between y_t and x_t:

$$y_t = \alpha_y + \beta_t x_t + \varepsilon_t^y \tag{24}$$

where, ε_t^x, ε_t^y and ε_t^β are random drawings from independent normal distributions each with mean 0. Moreover, equation (23) exhibits that y_t is also a BI(1) process. This yields that y_t and x_t are heteroscedastically cointegrated. By using these series, we will examine the Heteroscedastic cointegration models estimated by Kalman Filter.

4) In the last DGP, the dependence between y_t and x_t is again time varying, but in this case, the parameters vary smoothly as in Park and Hahn's article (1999). The parameter variation is defined by the following equation:

$$\beta_t = \beta_x + (t/N)^\gamma \tag{24}$$

where N is sample size and γ is the curvature parameter. In this specification the time varying term $(t/N)^\gamma$ is defined on the interval $[0,1]$. Moreover, the value of γ determines the shape of the graph of β_t. For instance, If $\gamma \le 1$ the graph of β_t will be convex, and otherwise β_t will be concave. We will consider both convex and concave parameter processes in our simulations.

Further, we have single regressor which can be defined as unit root process: $x_t = \alpha_x + x_{t-1} + \varepsilon_t^x$. Finally, the same structure in equation (23) governs the dependence between y_t and x_t.

4.2. Identification of Spurious Regression in Kalman Filter

In this section, we will concentrate on spurious regression problem in Kalman Filter estimation of time series models. In order to exhibit how spurious regression appears in Kalman Filter estimation, we conduct simulations with the data sets described in the previous section. In these simulations, we will use 3000 different y_t and x_t pairs for each data generation processes with the sample size 100 and 400. Furthermore, in order to run Kalman filter with these series, we first put our time series model into State-space form. This SSF containing the measurement and transition equations is given by equations (21) and (22) respectively. Additionally, in this SSF, d_t and c_t represent constant terms in measurement and transition equations respectively.

After we put the time series model into SSF, we run Kalman filter recursions using prediction (equation 13 and 14) and updating (equations 15 and 16) equations. From these Kalman Filter recursions, we obtain predicted value of y_t denoted as $y_{t|t-1}$, prediction of state vector $a_{t|t-1}$ and a_t namely updated estimate of β_t. With the help of these variables, we will show that Kalman Filter bear the risk of spurious regression. On the other hand, we will use RMSE so as to compare degree of the fit we obtain in the estimation with different specifications. Formula for RMSE is given below:

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(y_t - y_{t|t-1})^2} \tag{25}$$

where, we can express predicted value of y_t as in equation 26.

$$y_{t|t-1} = x_t a_{t|t-1} + d_t \tag{26}$$

Furthermore, the simulation outputs will also contains average values of the related statistics about Kalman Filter recursions over 3000 simulations. We can list related statistics as mean and standard deviation of a_t and $a_{t|t-1}$, coefficient of variation in a_t and correlation between a_t and y_t/x_t. These statistics reveal the patterns of the Kalman filter estimation under different data generation processes.

On the other hand, we will focus on two types of estimation method namely Kalman Filter estimation with known system matrices and unknown system matrices. These two estimation methods are essential in order to fully capture the patterns of the Kalman Filter recursions under each data generation process. The next sections describe these methods and exhibits simulation results for each estimation technique.

4.2.1. Kalman Filter with Known System Matrices

We first consider Kalman Filter estimation with known system matrices. In this case, the system matrices of the Kalman Filter recursions are assigned to particular numbers before the estimation. The assignment of these matrices may depend on some prior information or econometric intuition. We try to explain how we choose the values of the system matrices below.

In our model, we only allow slope parameter to vary in time. As a result, the intercept term of the model is represented by a fixed coefficient. This fixed coefficient coincides to the deterministic component of measurement equation (d_t) in state-space form. Furthermore, in Kalman Filter literature, there are different views about estimating this fixed intercept term. Harvey (1989) suggests using GLS algorithm so as to estimate the constant term of the measurement equation. Moreover, Hamilton (1994) claims that OLS can be used to estimate this parameter. We will follow the Hamilton's procedure in our simulations as Harvey's method is more complicated. Additionally, we ignore the constant term in transition equation, thus we choose c_t as 0. If any value other than 0 is chosen for c_t without any prior information, this may invalidate the Kalman Filter recursions because there is high risk of choosing wrong value. This critic may also be directed to choice of intercept term. However, we will show that the choice of d_t serves well in our simulations.

Another important system matrix of the model is Q_t, which is the variance of error term in transition equation. This matrix determines the volatility of the state vector by adjusting dispersion of the error term. For instance, high values of Q_t cause state vector to fluctuate more. Whereas, in the time series models, we do not confront with highly volatile time varying parameters. Further, many econometricians concentrate on smooth parameter specification as Bierens and Martins (2010) and Park and Hahn

(1999). Nonetheless, we should also consider the sets of time varying parameters which can not be defined as smooth transitions. Therefore, we determine the value of Q_t as 0.2, which will not cause too much volatility in state vector and also suits well for smooth transition models. On the other hand, we assume the error term in measurement equations follows standard normal distribution, so we set H_t as 1.

The most important part in the assignment of the system matrix is the choice of T_t. As we will show, T_t has very crucial role in Kalman Filter recursions. The importance of T_t stems from its power to determine the fluctuation level of the Kalman Filter prediction and updating equations. Furthermore, the previous studies in time series econometrics also attract attention to the choice of T_t. For instance, Prescott and Cooley adopts nonstationary state vector in their model (1976). Moreover, Evans suggests that in existence of nonstationary variables, one should use random walk parameters in the estimation (1991). We can conclude that these two papers suggest assigning T_t as 1, which provides nonstationary state vector. We will follow the suggestions of these scholars and use random walk parameter specification. Additionally, we will also investigate how estimation results change according to the process that parameter space follows. Following specifications for state vector will be analyzed in our study:

1.) $\beta_t = \beta_{t-1} + \eta_t$

2.) $\beta_t = 0.7\beta_{t-1} + \eta_t$

3.) $\beta_t = 0.3\beta_{t-1} + \eta_t$

After determining the specifications of T_t, the assignment of values of system matrices is complete. Moreover, following table summarizes the values of system matrices, which will be used in estimation:

Table 1. System Matrices Used in Estimations

H_t	Q_t	d_t	c_t	T_t	R_t
1	0.2	α_0	0	1, 0.7 and 0.3	1

Since we determine the values of system matrices, we only require to set the initial values a_0 and P_0 in order to start Kalman Filter recursions. These initial values are adjusted as described in the section named "Initialization of Kalman Filter". In that section, we emphasized that we will use diffuse prior for nonstationary state specification. Moreover, unconditional mean and MSE can be used as initial values of a_0 and P_0 respectively, if state vector is stationary. On the other hand, we only use Independent random walk and linearly related random walk variables in estimations since using only these data generation processes is sufficient to show existence of spurious regression in Kalman Filter with known system matrices.

4.2.1.1. Independent Random Walk Processes

We first consider estimation of the independently generated random walk series so as to explore the occurrence of spurious regression in Kalman Filter with known system matrices. In the estimation of these series, we expect that Kalman Filter will create a specious state vector. This state vector will deceive the researchers as if there is a cointegrating vector between these variables. Moreover, we think spuriousness is related with the choice of T_t in the Kalman filter estimation. Thus, if T_t is set as 1, appearance of Spurious regression will be clearer than in other cases. Additionally, we anticipate that Nonsense regression will occur in the estimation of both independent unit root processes which contain a drift term and which are generated as driftless series. Thus, we divide this section into two subsections which will check whether a significant difference appears in the estimation results of independent random walks with drift and without drift.

4.2.1.1.1. Independent Random Walk without Drift

We first focus on random walk without drift series. The data generation process of two series is given below:

$$\begin{aligned} y_t &= y_{t-1} + \upsilon_t, \quad \upsilon \sim iin(0,1) \\ x_t &= x_{t-1} + \omega_t, \quad \omega \sim iin(0,1) \end{aligned} \tag{27}$$

where ω_t and υ_t are serially and mutually uncorrelated white noise processes. These series are generated with the sample size 100 and 400. On the other hand, as mentioned previously, State-space form of the estimation is given by equations (21) and (22). After we put the model in SSF, we compute initial values of a_0 and P_0 as described in section 1.2 of chapter 3. Given these initial values and the system matrices, Kalman Filter recursions, which are defined by equations 13, 14, 15 and 16, are put in progress. We repeat this procedure for 3000 different sample. After the execution of 3000 simulations, we calculated average values of RMSE and statistics related with a_t and $a_{t|t-1}$.

Table 3 and 4 show the results of Kalman Filter estimation with the sample size 100 and 400 respectively. The results in these tables are important since they demonstrate how RMSE and other statistical values of estimates of β_t change according to specifications of T_t. Furthermore, notation used in these tables can be summarized in Table 2.

Table 2. Notations Used in Simulations

at	at_1	stdc	coeffvar	Corr(y/x~at)
Updated estimate of β_t	Predicted estimate of β_t	Standard deviation	Stdc(at)/Mean(at)	Correlation between y/x and at

Table 3. Identification of the Problem: Independent Random Walks without Drift Case N=100 with Known System Matrices

N=100		Q=0.2	H=1	x~I(1) without drift y~I(1) without drift			
iters=3000							
T	**mean(at)**	**stdc(at)**	**mean(at1)**	**stdc(at1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
1	-0.006	0.569	0.002	1.040	0.817	0.019	0.260
0.7	-0.005	0.406	-0.003	0.287	5.106	0.019	0.309
0.3	-0.004	0.345	-0.001	0.105	0.539	0.019	0.613

Table 4. Identification of The Problem: Independent Random Walks without Drift Case N=400 Kalman Filter with Known System Matrices

N=400		Q=0.2	H=1	x~I(1) without drift y~I(1) without drift			
iters=3000							
T	**mean(at)**	**stdc(at)**	**mean(at1)**	**stdc(at1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
1	-0.015	1.053	-0.020	1.417	1.443	0.016	0.121
0.7	-0.015	0.790	-0.010	0.552	1.905	0.016	0.206
0.3	-0.015	0.648	-0.005	0.194	3.686	0.016	0.423

One of the main purposes of this exercise is to see the impact of T_t on estimation results. From these tables, we can see that there is an inverse relationship between RMSE and T_t. As T_t falls, RMSE tends to increase. However, the change in RMSE due to T_t has different patterns for each sample size. In the estimation with sample size 100, when we alter T_t to 0.7 from 1, the observed change in RMSE is around 20 per cent whereas this difference is more than 100 per cent if we change T_t from 0.7 to 0.3. In spite of the fact that change increments in T_t is very close to each other, RMSE tends to increase more when the value of T_t is near 0. On the other hand, for sample size 400, this pattern is not as clear as in previous case. The increase in RMSE demonstrates similar patterns in both changes from 1 to 0.7 and from 0.7 to 0.3. Nevertheless, this difference does not change the fact that RMSE reaches its minimum when T_t approaches to zero. Thus, if one keeps T_t near zero in his estimation, he will relinquish the fit of the data. Additionally, there is another interesting feature of RMSE. For all values of T_t, RMSE is lower in sample size 400 case. This shows that in the large samples Kalman Filter tends to generate spurious regression with better fit than in small samples. This may be supported by the idea that Kalman Filter extracts better information from variables as the number of the recursions increases. Therefore, spurious regression problem in Kalman filter become severe in estimation of the large samples.

On the other hand, we mentioned that the time varying parameter should be zero for all periods, as there is no cointegrating vector between the unrelated random walk series. Furthermore, In both tables, mean value of estimated (a_t) and predicted ($a_{t|t-1}$) time varying parameter is near 0. This situation may preoccupy the readers about the behavior of a_t and $a_{t|t-1}$. However, we can also observe very high standard deviation in both a_t and $a_{t|t-1}$, which prevent estimated and predicted value of state to converge to zero. Thus we can conclude that Kalman Filter produce estimates which fluctuate around 0 with high volatility. In addition, we emphasized that we encounter lower RMSE in large sample. This may be again related to volatility rate of a_t and $a_{t|t-1}$. Standard deviations of a_t and $a_{t|t-1}$ are higher in large samples.

Thereby, there might be relation between the volatility of time varying parameter estimates and the degree of fit. Moreover, another relevant statistic in our simulations is the coefficient of variation, which measures the dispersion within a single variable after scaling with its mean. We will compare the dispersion in a_t by using this statistic in next sections.

The tables 3 and 4 give some idea about the patterns of the Kalman Filter estimation of the independent random walks without drift. However, we cannot fully observe the existence of spurious regression from these tables. Thus, we will refer to some figures which are picked from the simulations that we conducted. We will only use the figures with sample size 100, since graphs of simulations with sample size 400 have very similar patterns. Further, all figures are placed in Appendix A.

First of all, our claim about the emergence of spurious regression in Kalman Filter is verified by figure 1. In this figure, $y_{t|t-1}$, which is composed of d_t and $a_{t|t-1}x_t$, mimics the actual value of y_t when T_t is equal to 1. In other words, y_t can be predicted with the knowledge of x_t despite of the absence of the relation between these series. On the other hand, in order to show spurious regression, we should also check the patterns of predicted and estimated state vector. Although, there is no satisfying theoretical explanation about the patterns of the state vector, it is obvious that Kalman filter yields a prediction of the state, which mimics the ratio between dependent variable and regressor after a certain adjustment. We call this series as z_t which is defined below:

$$z_t = (y_t - d_t) / x_t \tag{28}$$

Figure 2 illustrates a_t, $a_{t|t-1}$ and z_t sequences. It is clear that these series are strongly linked to each other. They seem to follow the same path with some minor distortions. Further, a_t and $a_{t|t-1}$ is not a sequence, which fluctuates very close to 0 as expected. On the contrary, it seems there is a cointegrating vector between y_t and x_t which is exactly different than the zero vector. Thus, when T_t is equal to 1, we can observe Kalman Filter generates a nonsense state estimate, which cause spurious regression problem. Furthermore, A similar argument was proposed by Hanohan (1993).

Hanohan (1993) comments on behavior of time varying intercept term instead of time varying slope parameter.

Now, does spurious regression, shown above, persist with other choices of T_t? In order to see this, we first take a look at the figures exhibiting the estimation results when $T_t = 0.7$. Figure 3 shows that the link between $y_{t|t-1}$ and y_t gets worsened when we reduce the value of T_t. Hence, decreasing T_t has negative impact on degree of the fit. This negative impact can be explained by the loss of flexibility of Kalman Filter in updating the state vector. As seen in equation (13), lower values of T_t will draw the predicted state away from the value which provides a spurious fit in estimation. Furthermore, via equation (14) T_t is also directly related with the prediction equation of the variance covariance matrix of estimation error, which is denoted as $P_{t|t-1}$. This direct relation also has implicit impact on state prediction. First, we can see that $P_{t-1|t-2}$ involves in a_{t-1}. Moreover, $a_{t|t-1}$ is a function of a_{t-1}. As a result, T_t influences $a_{t|t-1}$ via the prediction equation of the variance covariance matrix of the estimation error of the previous period. In addition, figure 4 illustrates the departure of $a_{t|t-1}$ and a_t from z_t which is spurious state vector when T_t is equal to 1. However, a_t seems to depart from z_t with slower rate than $a_{t|t-1}$. This is again a natural result, since Kalman Filter creates a_t after y_t is included in estimation.

On the other hand, for the case $T_t = 0.3$, both the departure of $y_{t|t-1}$ from y_t and the departure of $a_{t|t-1}$ from z_t can be seen clearly. Figure 5 shows the graph of $y_{t|t-1}$ and y_t, and figure 6 shows the a_t, $a_{t|t-1}$ and z_t sequences. In figure 5, we can observe that Kalman filter cannot predict y_t as good as previous case. Moreover, $y_{t|t-1}$ and y_t seem to be irrelevant series depicted in the same picture. Thereby, this specification of T_t almost limits Kalman filter to create spurious relation between variables as T_t approaches to 0. Moreover, Another interesting result can be observed in figure 6. $a_{t|t-1}$ deviates from z_t and approaches to zero line. This result explains the reason of disappearance of spurious regression. Since $a_{t|t-1}$ takes values close to 0, $y_{t|t-1}$ carries very limited information extracted from x_t. Therefore, spurious regression tends to vanish with lower values of T_t.

To sum up, we have shown that in Kalman filter estimation of independently generated random walk series Spurious regression inevitably emerge when T_t is defined as unity. Moreover, The emergence of the Spurious regression has some prognostications. First of all, Spurious regression is accompanied with high volatility in state vector. However, highly volatile state vectors may not be valid indicator of Nonsense regression, since there may exist other data generation process, which yields similar results. Furthermore, mean value of a_t and $a_{t|t-1}$ is very close to zero. Nevertheless, these sequences actually do not converge to zero vector, but mimic another variable, which is the ratio of $y_t - d_t$ and x_t. As a consequence, the spurious regression occurs if one unfortunately uses independent random walk without drift series in his estimation. On the other hand, T_t is a very important factor in occurrence of nonsense regression. In our simulations, we observe that T_t directly influence the degree of fit. For instance, when T_t is equal to 1, the best fit is achieved. On contrary, lower values of T_t cause the link between actual and predicted values of y_t to vanish.

4.2.1.1.2. Independent Random Walks With Drift

It is also an interesting exercise to examine the random walk with drift variables in Kalman Filter. This exercise may help us see whether Entorf's findings can be observed in time varying parameter framework. However, we expect that independent random walk with drift case will not give very different results from independent random walk without drift case. Since, Entorf's (1997) findings give importance on intercept term which we estimate by OLS (not in Kalman Filter), his results are irrelevant with our case. Nevertheless, it may be helpful to analyze this case in order to understand nature of spurious regression.

We apply exactly the same method described in the previous section. The only difference between two sets of simulations occurs in the data generation processes. In this section, we use the independent random walks defined as in the equation (29):

$$y_t = \alpha_y + y_{t-1} + \varepsilon_t^y,\ \varepsilon_t^y \sim iin(0,1)$$
$$x_t = \alpha_x + x_{t-1} + \varepsilon_t^x,\ \varepsilon_t^x \sim iin(0,1) \tag{29}$$

where, α_y and α_x are constants, ε_t^y and ε_t^x are serially and mutually uncorrelated white noise processes. In our simulations, we set $\alpha_x = 0.02$ and $\alpha_y = 0.04$. Moreover, in the evaluation of the simulation results, we utilize the same statistics used for independent random walks without drift case. The tables 5 and 6 exhibit these statistics for the sample size 100 and 400 respectively.

Table 5. Identification of the problem: Independent Random walks with drift case Sample size=100 with Known System Matrices

N=100		Q=0.2	H=1	x~I(1) with drift y~I(1) with drift			
iters=3000							
T	**mean(at)**	**stdc(at)**	**mean(at_1)**	**stdc(at_1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
1	0.023	0.570	0.009	1.076	1.584	0.205	0.298
0.7	0.019	0.398	0.013	0.282	1.372	0.205	0.339
0.3	0.019	0.340	0.006	0.103	0.451	0.205	0.685

Table 6. Identification of the problem: Independent Random walks with drift case sample size=400 with Known System Matrices

N=400		Q=0.2	H=1	x~I(1) with drift y~I(1) with drift			
iters=3000							
T	**mean(at)**	**stdc(at)**	**mean(at_1)**	**stdc(at_1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
1	0.102	1.043	0.097	1.268	-0.187	0.185	0.125
0.7	0.102	0.795	0.071	0.557	-0.452	0.185	0.252
0.3	0.101	0.661	0.030	0.198	-0.113	0.185	0.535

Tables 5 and 6 demonstrate very similar results with tables 3 and 4. First, the reaction of state vector to the changes in T_t is the same as the random walk without drift case. As T_t decrease, we can see that RMSE, thus, degree of fit between actual and predicted y_t declines. Moreover, when we compare the magnitudes of RMSE in each case, we do not confront with significant differences. Average RMSE seems to be slightly higher than without drift case. However, the major difference between these cases emerges in simulations with sample size 400. In this case, the mean value of the estimated state vector is near 0.1 which is different from the without drift case. This difference also divulges that the estimated and predicted state vector may not have certain pattern, which provides the same mean in all cases. On the other hand, we also encounter with high standard deviations of the predicted and estimated state vector in the estimation of the unrelated random walks with drift variables. Additionally, an interesting difference from without drift case appears in coefficient of variation calculated for sample size 400. We can observe negative coefficient of variation in this case. This should be the result of the negative mean of a_t. However, this result does not have much impact on our analysis. In conclusion, similar comments made for estimation results of driftless random walks are valid for this case.

In addition, for sample size 100 figures 7, 9 and 11 illustrate relation between the actual and the predicted values of y_t with different selections of T_t. In the first instance, Figures 7 shows that Kalman Filter creates the prediction of y_t very close to the actual value of the variable when T_t is equal to 1. These predictions worsen when T_t is reduced as seen in figures 9 for $T_t = 0.7$. Finally, when T_t is set to 0.3, y_t and $y_{t|t-1}$ sequences seem to be very different processes from each other. These results are very similar to the findings of the previous section. Furthermore, the reason in these results is not different than what we report in the estimation of the independent random walk series without drift.

On the other hand, we obtain resembling results about patterns drawn by a_t and $a_{t|t-1}$ as in without drift case. Figure 8 for the sample size 100 exposes the relation between a_t, $a_{t|t-1}$ and z_t as T_t is equal to 1. From this figures, it seems that a_t, $a_{t|t-1}$ follow almost the same process as z_t. As a result, there is a strong relation between a_t, $a_{t|t-1}$ and z_t, which creates spurious regression in the estimation. Furthermore, fall in T_t causes the strength of this relation to decrease, thus the spurious regression vanishes as a consequence. We can see these results in figure 10 for sample size 100. In figure 10, when T_t is equal to 0.7, a_t and $a_{t|t-1}$ tend to move away from z_t. Additionally, setting T_t as 0.3, increase the gap between a_t, $a_{t|t-1}$ and z_t as seen in the figure 12. Actually an expected result emerges from these figures. When T_t is 0.3 or in other words near zero, the sequence $a_{t|t-1}$ tends to move close to zero for both sample sizes. Moreover, from these results, we can also deduct that Kalman Filter tends to estimate the state vector as a ratio of dependent and independent variables in regression given above if T_t is chosen as 1. So, one may suspect that this tendency of Kalman Filter causes spurious regression also in random walk with drift data.

Consequently, in both with drift and without drift series, spurious regression is inevitable if $T_t = 1$. Furthermore, the estimation results of these two different types of independent random walk series have many patterns in common. Both estimations contain highly volatile state estimates, which mimics the variable z_t. In addition, our simulations reveal that Spurious regression in both cases substantially depends on the values of T_t. In these simulations, we observe that there is a tendency of spurious regression to disappear when T_t is decreased. This situation make us think that Nonsense regression is all about the mechanism of Kalman Filter, which seeks to create best prediction of dependent variable with the knowledge of regressor whether there is a relation between these variables. However, this prediction is based on a delusive state vector if the series are unrelated. Thus, researchers should be aware of the risk of spurious results, which is likely to appear in Kalman filter estimation with known system matrices.

4.2.1.2. Linearly Related Random Walks

In the previous section, we analyze the Kalman filter estimation of the independently generated random unit root series and exhibit spurious regression. However, this analysis is not sufficient for understanding the mechanics of Kalman Filtering in estimation of random walk series. Therefore, we apply the same procedure to related random walk series. However, since we do not have much space, we deal with the simplest case. First, we generate x_t series as a unit root process with drift. After x_t is obtained, y_t is formed as linear function of x_t. This procedure can be observed below:

$$\begin{aligned} x_t &= \alpha_x + x_{t-1} + \varepsilon_t^x, \qquad && \varepsilon_t^x \sim iin(0,1) \\ y_t &= \alpha_y + \delta x_t + \varepsilon_t^y, \qquad && \varepsilon_t^y \sim iin(0,1) \end{aligned} \tag{30}$$

Thereby, y_t and x_t are related via equation (30), where α_x, α_y and δ are some known constants, and ε_t^y and ε_t^x are serially and mutually uncorrelated white noise processes. Furthermore, δ can also be interpreted as cointegrating vector that relates x_t and y_t. Accordingly, If classical linear regression models are applied to these series, a cointegration relation between variables can be observed. However, it is not certain that Kalman Filter can capture this linear cointegrating vector. Therefore, we will check what happens when estimation method for these series is chosen as Kalman filter. In order to see this, we again conduct series of simulations. In these simulations, we define $\alpha_x = 0.02$, $\delta = 0.9$ and $\alpha_y = 0.5$. As in previous case, we obtain relevant statistics from Kalman filter estimation. These statistics are summarized in the following tables.

Table 7. Identification of the problem: Related Random walks case Sample size=100 with Known System Matrices

N=100		Q=0.2	H=1	x~I(1) with drift and y=0.5+0.9*x+ε ε~N(0,1)			
iters=3000							
T	**mean(at)**	**stdc(at)**	**mean(at_1)**	**stdc(at_1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
1	0.892	0.164	0.900	0.138	0.184	0.020	0.136
0.7	0.848	0.124	0.587	0.112	0.158	0.020	0.495
0.3	0.805	0.136	0.239	0.051	0.202	0.020	1.062

Table 8. Identification of the problem: Related Random walks case Sample size=400 with Known System Matrices

N=400		Q=0.2	H=1	x~I(1) with drift and y=0.5+0.9*x+ε ε~N(0,1)			
iters=3000							
T	**mean(at)**	**stdc(at)**	**mean(at_1)**	**stdc(at_1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
1	0.898	0.131	0.900	0.128	0.146	-0.034	0.069
0.7	0.858	0.123	0.599	0.095	0.150	-0.034	0.298
0.3	0.822	0.145	0.246	0.047	0.192	-0.034	0.671

Above results exhibit different features from the estimation outcomes with unrelated series. First of all, for all choices of T_t, a_t and $a_{t|t-1}$ no longer fluctuate around zero mean. Moreover, when T_t is equal to 1, both a_t and $a_{t|t-1}$ are concentrated around the true value of slope coefficient with very small standard deviation. This depicts Kalman Filter generated approximately constant state estimates very close to the actual value of the slope parameter as expected. Furthermore, We also see that for all values of T_t and sample sizes, in the estimation of the linearly related series RMSE is lower than in the estimation of the unrelated series. However, we do not guess this can be used in the detection of spurious regression, since there is no threshold level of RMSE which can be used for comparison. On the other hand, there are some similar results with the independent random walk case. For instance, RMSE tends to rise in response of a reduction in T_t. In both sample sizes, the rate of rise in RMSE is increasing as T_t approaches to 0. Further, in the estimation of the linearly related random walk processes with Kalman filter, best fit of the data is obtained when $T_t = 1$ is imposed. This is the same result we observe in the estimation of the unrelated series. Another common result can be explored in the pattern of standard deviation of $a_{t|t-1}$. As T_t decreases, volatility of $a_{t|t-1}$ also fall. However, we do not think that this is the main source of the loss in degree of the fit when T_t moves down. We guess that RMSE dramatically rises because of the fall in mean value of $a_{t|t-1}$. In both tables, average $a_{t|t-1}$ tends to move down, as T_t converges to 0. Since $x_t a_{t|t-1}$ and d_t constitute $y_{t|t-1}$, the fall in average $a_{t|t-1}$ means that on average predicted value of y_t moves away from the fitted value obtained in case $T_t = 1$. As a consequence, change in T_t substantially impacts degree of the fit via $a_{t|t-1}$.

We will check the related figures in order to enhance the above comments. The figure 13 illustrates that Kalman Filter generates y_t prediction very close to actual value of y_t when T_t is equal to 1. However, T_t again plays very crucial role in degree of fit. Although, for $T_t = 0.7$, in figure 15 predicted y_t slightly moves away from its actual value, figure 17 clearly shows that degree of fit significantly dwindles when T_t is defined as 0.3. Hence, setting T_t as 0.3 considerably distorts the fit between actual and predicted values of y_t. The conclusions in these pictures are also compatible with results of table 7 and 8.

Furthermore, when T_t is equal to 1, Kalman filter produces estimates which have almost same pattern as z_t defined in previous section (Figure 14). However, decreasing T_t to 0.7 and 0.3 causes conspicuous deviation in both a_t and $a_{t|t-1}$ from z_t, which is also indicated as a fall in the mean value of a_t and $a_{t|t-1}$ in the tables 7 and 8. If we consider N=100, Figures 16 and 18, respectively for $T_t = 0.7$ and

$T_t = 0.3$ illustrates the sharp deviations of $a_{t|t-1}$ from z_t. But, the response of a_t to decrease in T_t is weaker than $a_{t|t-1}$. We have already discussed these outcomes when interpreting the table results. In addition, we show in the estimation of the unrelated series, a_t and $a_{t|t-1}$ display very similar patterns when we change T_t. However, in estimation of linearly related series, $a_{t|t-1}$ is affected more than a_t in response to change in T_t. On the other hand, for all values of T_t, the estimated and predicted state seem to fluctuate around a horizontal line. This line corresponds to the true parameter value 0.9 when T_t is imposed as 1. As a result, Kalman Filter captures the slope parameter in linear model when T_t is equal to 1. Further, decreasing T_t will not spoil this pattern, instead, the level of the line moves down with lower values of T_t. Accordingly, Kalman Filter also generates approximately constant state estimate when T_t is lower than 1. However, as T_t approaches to 0, the line, around which state vector fluctuates, moves toward the x-axis.

Consequently, above results show that when T_t is given as 1, both in estimation of related series and independently generated random walk series Kalman Filter will yield prediction as if there is a relation between these series. However, it is very hard for a researcher to determine whether these series are actually cointegrated or not, since both estimation of cointegrated and noncointegrated series have similar features. One of these similarities is observed in the response of y_t prediction to the change in T_t. In both data generation processes, predicted value of y_t moves away from actual value. On the other hand, best fit of the data, which may be spurious, is achieved when T_t is equal to 1. Thus, we can conclude that setting T_t as 1 bears the risk of spurious regression in estimation of random walk series in Kalman filter with known system matrices. Furthermore, the solution to avoid nonsense regression may be considered as using low values of T_t. However, this will distort the performance of Kalman filter in estimation of linearly related series as well. Moreover, if we glance at average standard deviation of the estimates of β_t in their estimation of both related and independent series, it is obvious that in independent random walk case standard deviations tends to be higher. This may imply that if there is no valid cointegrating vector between y_t and x_t, Kalman Filter create false parameter estimates, which fluctuates more. In our simulations, we also exhibit that these estimates mimic the ratio of y_t and x_t after some adjustment. In addition, The lack of the cointegrating relation between variables may be the reason of higher standard deviations in Kalman estimates. However, one should still be careful to detect spuriousness, since only looking at behavior of time varying parameter estimates may not be sufficient for detection of nonsense relation.

4.2.2. Kalman Filter with Unknown System Matrix

In this section, we are going to focus on Kalman Filter estimation with unknown system matrices, which should also be estimated. Thus, Hamilton (1994) suggests that these unknown system parameters could be estimated by maximum likelihood estimation. This means that these parameters can be estimated via a composite estimation method, which consists of Kalman Filter recursions and ML procedure. In order to see the implication of this estimation method on time series models, we will perform simulations as in previous sections. By this exercise, we aim to investigate whether a ML estimation with Kalman filter recursions can help us to detect spurious regression. We expect that ML estimation via Kalman filter cannot avoid spurious regression. On contrary, it will worsen the problem, since ML estimation will seek to enhance the prediction performance of Kalman Filter. Furthermore, so as to apply ML estimation, we will redefine the measurement and transition equations for unknown parameters of the model.

$$\begin{aligned} y_t &= \alpha + \beta_t x_t + \varepsilon_t \qquad && \varepsilon_t \sim N(0,\sigma_h^2) \\ \beta_t &= \mu + T\beta_{t-1} + \eta_t && \eta_t \sim N(0,\sigma_q^2) \end{aligned} \tag{31}$$

In above system we replaced d_t, c_t, T_t, H_t and Q_t with unknown parameters $\alpha, \mu, T, \sigma_h^2, \sigma_h^2$ respectively. Here, we have 5 parameters to estimate in Kalman Filter recursions. Then, we can write $\Phi = (\alpha, \mu, T, \sigma_h^2, \sigma_h^2)$ as the parameter space. Furthermore, we require a likelihood function so as to estimate these unknown parameters. The equation (32) gives this sample log-likelihood function for Kalman Filter.

$$\sum_{t=2}^{N} \log f(y_t \mid x_t, \Upsilon_{t-1}) = -(N/2)\log(2\pi) - (1/2)\sum_{t=2}^{N} \log(x_t P_{t|t-1} x_t + \sigma_h^2)$$
$$- (1/2)\sum_{t=2}^{N} (y_t - x_t a_{t|t-1})^2 / (x_t P_{t|t-1} x_t + \sigma_h^2) \tag{32}$$

where, $\Upsilon_{t-1} = (y_{t-1}, y_{t-2}, ..., y_1)$, $P_{t|t-1}$ and $a_{t|t-1}$ are defined in chapter 3. After, we obtain the sample likelihood function, we can apply a general nonlinear optimization procedure (called OPTMUM library) created in Gauss programming language in order to maximize the sample log likelihood function with respect to parameter space Φ. On the other hand, we obtain the covariance matrix of ML estimators by taking inverse of OPG estimator.

Different from previous section, we will use 4 types of data generation process. These can be listed as: 1) unrelated random walks, 2) linearly related random walks, 3) Time varying parameter model with AR(1) specification and 4) Time varying parameter model with smooth transitions. All of these data generation processes are described at the beginning of this chapter.

4.2.2.1. Independent Random Walk Processes

We first analyze estimation of unrelated series with Kalman Filter via ML. We use exactly the same set of variables generated in 2.1.1. Accordingly, we consider 4 cases, which can be listed as random walks with drift and without drift consisting of 400 and 100 observation. By the help of these simulations, we can observe whether spurious regression appears in the Kalman filter estimation with ML. What is more, we have chance to verify our choices about the system matrices in previous section. Below table summarizes the estimation results for a_t and $a_{t|t-1}$ related statistics and RMSE.

Table 9. Identification of the problem: Independent Random Walks case with Unknown System Matrices

	x~I(1) and y~I(1) independently generated						
N=100	**mean(at)**	**stdc(at)**	**mean(at_1)**	**stdc(at_1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
With Drift	-0.002	0.389	-0.002	0.372	-0.526	0.468	0.143
Without drift	-0.014	0.391	-0.013	0.373	-0.547	0.500	0.143
N=400							
With Drift	0.032	0.722	0.033	0.712	0.395	0.362	0.100
Without drift	-0.008	0.709	-0.008	0.699	0.834	0.381	0.098

We can see that table 9 have some common patterns with tables 3, 4, 5 and 6, in which results of Kalman Filter estimation with known system matrices are exhibited. First, in all cases a_t and $a_{t|t-1}$ seem to have zero mean on average. This indicates that a_t and $a_{t|t-1}$ move around zero line. Additionally, standard deviations of a_t and $a_{t|t-1}$ are still high. However, the volatility of state estimates and predictions seem to be slightly lower than the previous cases. We guess that the reason in the reduction of the volatility of a_t and $a_{t|t-1}$ is about ML estimate of σ_q^2 which is computed around 0.01 for all cases and sample sizes. Adversely, we use 0.2 for σ_q^2 in previous case. Thus, optimal value of this parameter is very low than our choice for Q_t in Kalman filter estimation with known system matrices. On the other hand, RMSE seems to improve which can be interpreted as follows: Estimating unknown parameters with MLE enhances the degree of fit. However, this enhanced fit is surely a spurious one which is outcome of the integration of two optimizing procedure namely MLE and Kalman Filter.

To be sure about the existence of spurious regression, we will check the relevant figures. For sample size 100 with drift and without drift cases, the spuriousness of estimated state vector become obvious if we examine the figures 19 and 21 respectively. These figures demonstrate the predicted value of y_t that follows a path very near to the actual value. Thus, the regressor x_t looks as if it is related to y_t via a time varying cointegrating vector. Nevertheless, we know this is not a true relation, since we generate series independently with each other. Now, let us set our eyes on the estimated and predicted state vector. As seen in the figures 20 for without drift case and 22 for with drift case, the estimated and predicted state moves along a variable $z_t^{mle} = (y_t - \alpha_{mle})/x_t$ where, α_{mle} is the ML estimate of the constant term in measurement equation. Although there are sharp deviations at some periods, a_t and $a_{t|t-1}$ appear to trace z_t^{mle} very closely in the most of the observations. Further, we guess that these sharp distortions in z_t^{mle} actually are caused by the patterns of x_t or y_t at those points. Furthermore, the relation between a_t, $a_{t|t-1}$ and z_t^{mle} seems to be stronger in Kalman Filter estimation with known system matrices (See figures 2 and 8).

On the other hand, it is also interesting to check the distribution of ML estimators. We will report the results for only with drift case and sample size 100 and 400, since there is not much difference with without drift case. Table 10 and 11 exhibits the histograms of ML estimates of α, T and μ out of 3000 simulation respectively for N=100 and N=400. Both in the simulations with sample size 100 and 400, the intercept term α do not possess a well behaved distribution around zero. In addition, its distribution is highly volatile. However, since there is no valid cointegrating vector between y_t and x_t, we expect α to be zero. Instead, the MLE of α moves to a random number. Moreover, the distribution of ML estimator of T seems to be concentrated near 1 with a low variance. It is highly probable that we have highly volatile time varying parameter estimates since T goes to a number close to 1. Furthermore, we pointed out that Kalman Filter produce best fit (spurious or true) when T is set to one in Known system matrices case. Hence, these results also validates the choice of T_t in the previous section. Additionally, MLE of μ is clustering around zero with very low variance. This indicates that state vector follows a near unit root process without drift. Finally, we can add that the tendencies of these distributions become obvious as we increase sample size to 400 as expected since Kalman filter performs better with large samples.

Table 10. Histograms of MLE Estimator of Parameters: Random Walk with Drift Sample Size=100

α			T			μ	
bin	*Frequency*		*bin*	*Frequency*		*bin*	*Frequency*
-1.2	493		-1.2	0		-1.2	6
-1.1	2		-1.1	0		-1.1	2
-1	6		-1	0		-1	5
-0.9	5		-0.9	0		-0.9	1
-0.8	6		-0.8	0		-0.8	1
-0.7	5		-0.7	2		-0.7	2
-0.6	2		-0.6	4		-0.6	3
-0.5	2		-0.5	2		-0.5	2
-0.4	3		-0.4	6		-0.4	2
-0.3	4		-0.3	4		-0.3	5
-0.2	6		-0.2	9		-0.2	5
-0.1	8		-0.1	2		-0.1	15
0	7		0	11		0	1362
0.1	1		0.1	12		0.1	1569
0.2	3		0.2	4		0.2	7
0.3	6		0.3	3		0.3	2
0.4	3		0.4	0		0.4	1
0.5	1		0.5	0		0.5	1

0.6	4		0.6	0		0.6	2
0.7	8		0.7	1		0.7	2
0.8	3		0.8	24		0.8	1
0.9	6		0.9	269		0.9	1

Table 10 continued: Histograms of MLE Estimator of Parameters: Random Walk with Drift Sample Size=100

1	4		1	2453		1	0
1.1	2		1.1	190		1.1	1
1.2	2		1.2	4		1.2	1
More	2408		More	0		More	1
Average	16.79898			0.929735			-0.00505
Variance	18.612			0.167234			0.122221
# of times the null that ML estimate is insignificant is rejected							
84			8			2522	

Table 11. Histograms of MLE Estimator of Parameters: Random Walk with Drift Sample Size=400

α			T			μ	
bin	*Frequency*		*bin*	*Frequency*		*bin*	*Frequency*
-1.2	454		-1.2	0		-1.2	1
-1.1	2		-1.1	0		-1.1	0
-1	0		-1	0		-1	0
-0.9	7		-0.9	0		-0.9	2
-0.8	7		-0.8	1		-0.8	0
-0.7	7		-0.7	1		-0.7	0
-0.6	5		-0.6	1		-0.6	0
-0.5	4		-0.5	1		-0.5	0
-0.4	4		-0.4	1		-0.4	0
-0.3	4		-0.3	0		-0.3	0
-0.2	5		-0.2	1		-0.2	0
-0.1	3		-0.1	2		-0.1	0
0	3		0	1		0	1144
0.1	5		0.1	0		0.1	1848
0.2	2		0.2	0		0.2	1
0.3	5		0.3	0		0.3	0
0.4	5		0.4	0		0.4	0
0.5	3		0.5	0		0.5	0
0.6	4		0.6	0		0.6	1
0.7	1		0.7	0		0.7	1
0.8	6		0.8	0		0.8	1
0.9	5		0.9	0		0.9	0
1	5		1	2839		1	0
1.1	1		1.1	152		1.1	0
1.2	3		1.2	0		1.2	0
More	2450		More	0		More	1
Average	20.73917			0.983391			0.001362
Variance	20.87744			0.079974			0.062749

# of times the null that ML estimate is insignificant is rejected				
197		57		2623

To sum up, In Kalman Filter estimation with unknown parameters, Spurious regression is again unavoidable. Thus, the ML estimation cannot prevent nonsense relation to appear, but conversely worsens the problem by improving degree of fit. Additionally, the ML estimators of α and T appear to be significant according to individual t-statistics in most of the simulations. However, one may expect T to be insignificant, since $T = 0$ will yield a zero vector of $a_{t|t-1}$ and prevent spurious regression. Furthermore, this result may be linked to outcomes of Spurious regression in fixed parameter models. Recall that in spurious regression with fix coefficients we obtain misleading significance test statistics. In Kalman Filter estimation, we also obtain false significance test results. What is more, from above tables and figures in appendix, we can conclude that there are two potential sources of spuriousness. First one is the incorrect ML estimation of intercept term α and second source of spurious regression is high fluctuation in time varying parameter estimates which depends on ML estimate of T. These results also indicates that T_t has very crucial role in emergence of spurious regression. Consequently, Using the ML estimation with Kalman Filter recursion may also mislead one's research by yielding spurious outcomes. Hence, researchers should be careful when using nonstationary series in Kalman Filter especially if they use ML estimation.

4.2.2.2. Linearly Related Random Walk Processes

Our analysis will continue on related series. We will first consider linearly related random walk processes. As we describe at the beginning chapter 4, the regressor x_t follows random walk and dependent variable y_t is linear function of x_t. Thus, there is fixed cointegrating vector between x_t and y_t. Further, the same data sets generated in section 2.1.2 are used so as to compare and contrast the results with conclusions of this section. We will only report driftless random walk cases with sample size 100 and 400, as there is no distinguishing feature of estimation involving random walk with drift processes. Therefore, we seek to estimate the cointegrating vector between variables and other unknown parameters of model via Kalman Filter with unknown system matrices. A ML estimation method is adopted as in the estimation of the independently generated random walk variables. Table 12 exhibits the results.

Table 12. Identification of the problem: Linearly Related Random Walks Case with Unknown System Matrices

x~I(1) and y=0.5+0.9x+ε							
iters=3000							
N=100	**mean(at)**	**stdc(at)**	**mean(at1)**	**stdc(at1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
Without drift	0.899	0.017	0.899	0.010	0.019	0.340	0.099
N=400							
Without drift	0.900	0.007	0.900	0.005	0.007	0.189	0.050

From this table, we can see MLE constitutes smaller RMSE than Kalman Filter with known system matrices. Hence, as in estimation of unrelated series, MLE enhances the performance of Kalman Filter. This result does not exhibit an unexpected picture. Optimizing nature of both MLE and Kalman Filter produces better precision in predicting y_t. Thus, we end up with lower RMSE in Kalman filter estimation with unknown system matrices. Furthermore, two features of a_t and $a_{t|t-1}$ immediately attract attention. First, in both sample size, mean value of a_t and $a_{t|t-1}$ catches the true value of slope coefficient. Secondly, the standard deviation of a_t and $a_{t|t-1}$ is very close to 0. These two patterns of estimated and predicted value of state vector imply that Kalman Filter creates a constant sequence of state estimates which move along true value of slope parameter. Thus, Kalman filter recursions accurately captures the linearity of the true data generation process. As consequence, this accuracy also reverberate through RMSE.

We can also observe the improvement in the degree of fit in the figure 23. In this graph, the predicted value of y_t seems to trace the actual value very closely. Additionally, if we compare figures 13 (estimation with known system matrices) and 23, the figure 23 exhibits stronger relation between the actual and predicted y_t which is also indicated by lower RMSE. On the other hand, in figure 24, we can explore that the graph of $a_{t|t-1}$ and a_t seems to be a vertical line passing through the true value of slope parameter even though z_t^{mle} is more volatile. There may be two reason for this result. First, ML estimate of σ_q^2 may be very close to 0. Thus, this will prevent the state vector to flexibly fluctuate. On the other hand, T may be estimated as almost 0. In this case, the predicted state vector is equal to μ^{mle} for all observations. Then, also μ^{mle} will be equal to mean value of $a_{t|t-1}$, which is found to be very close to actual slope coefficient. Hereupon, we should check the ML estimates of the parameters of the model.

Tables 20 and 21 (in appendix B) exhibit the histograms of the ML estimates of the parameters α, T, μ and σ_q for sample size 400 and 100 respectively. This set of parameters can be divided into two subgroup. The first subgroup consists of T, μ and σ_q which determine the estimate of the slope parameter in the underlying linear model. In Kalman Filter estimation. this slope parameter is assumed to be time variant. Further, it is predicted by $a_{t|t-1}$, which can be divided $a_{t|t-1}$ into two components namely time varying and time invariant parts (see equation 13). First, μ is time invariant component of state vector. In time varying part (Ta_{t-1}), variation of predicted state is mainly determined by T and σ_q (note that σ_q is involved in a_{t-1} via updating equation of Kalman filter). Since both T and σ_q is accumulated near 0, variation in $a_{t|t-1}$ will be very limited. Consequently, the predicted state vector will be almost a constant sequence, which is equal to the ML estimate of μ. Moreover, the ML estimate of μ is clustering near 0.9, which is true value of the slope parameter. On the other hand, the second subgroup contains only α which captures the intercept term of the underlying linear model. The tables illustrate that α seems to be normally distributed around 0.5, which is true value of intercept term. Moreover, except σ_q, tables demonstrate that distributions of all parameters seems to be normal though we do not have theoretical explanation for this situation. However, σ_q is stuck at 0.1. Finally, sample size has very little impact on the distributions of parameters except α, which is more volatile in sample size 100 case.

To sum up, ML estimation with Kalman filter recursions seems to be very powerful tool for estimating and predicting linear cointegration models. This estimation method successfully predicts the dependent variable and the underlying parameters of the model. Furthermore, it removes some problems encountered in Kalman Filter with known system matrices. One of these problems is the inappropriate choice of system matrices which is not sufficient to explain the data. However, the ML estimation does not require the selection of the system matrices, further, it estimates the optimal system matrices for the data set in the cointegration models. Therefore, it is natural that the ML has better performance in linearly related data generation than Kalman filter estimation with known system matrices. On the other hand, we obtain distinguishing results from the estimation outcomes of Kalman filter with known system matrices. One of these results is about the ML estimate of T. This estimate is computed as 1 for unrelated random walk case. However, T is estimated around 0 in estimation of related random walks.

4.2.2.3. Related Random Walk Series with Time Varying Cointegrating Vector

In this section, we will use two different types of data generation processes namely time varying parameter following AR(1) process, and time varying parameter with smooth transitions. TVP following AR(1) process offers more volatile parameter specification than the smooth transition. Thus, we can examine the implications of the volatility differences of two DGPs on Kalman Filter estimation. Another aim of this section is to see the performance of Kalman Filter under various time varying parameter specifications.

4.2.2.3.1. Time Varying Parameters Generated as AR(1) Process

The first time varying parameter model, we consider, is TVP generated as AR(1) processes. We will consider 4 data generation process as derivatives of the type 3 described in section 1 of chapter 4. Furthermore, we will only report results with driftless series. These data sets are generated according to the following rules:

DGP 1) Independent variable is generated as a driftless random walk: $x_t = x_{t-1} + \varepsilon_t^x$ and the parameter vector follows AR(1) process with the AR coefficient 0.7: $\beta_t = 0.5 + 0.7\beta_{t-1} + \varepsilon_t^\beta$. Finally, the relation between x_t and y_t is given by equation:

$y_t = 0.5 + \beta_t x_t + \varepsilon_t^y$ where, ε_t^x, ε_t^y and ε_t^β are random drawings from independent normal distributions each with mean 0. We also have $\varepsilon_t^\beta \sim N(0, 0.2)$.

DGP 2) $x_t = x_{t-1} + \varepsilon_t^x$ and the parameter vector follows $\beta_t = 0.5 + 0.7\beta_{t-1} + \varepsilon_t^\beta$. Finally, $y_t = 0.5 + \beta_t x_t + \varepsilon_t^y$ where, ε_t^x, ε_t^y and ε_t^β are random drawings from the independent normal distributions each with mean 0. We have $\varepsilon_t^\beta \sim N(0, 0.02)$. Thus, we consider smoother parameter process than case 1 with same AR coefficient.

DGP 3) $x_t = x_{t-1} + \varepsilon_t^x$ and the parameter vector follows $\beta_t = 0.5 + 0.3\beta_{t-1} + \varepsilon_t^\beta$. The x_t and y_t is related via $y_t = 0.5 + \beta_t x_t + \varepsilon_t^y$ where, ε_t^x, ε_t^y and ε_t^β are random drawings from independent normal distributions each with mean 0. Further, We have $\varepsilon_t^\beta \sim N(0, 0.2)$.

DGP 4) In this data generation process, we consider smoother version of DGP 3. the independent variable $x_t = x_{t-1} + \varepsilon_t^x$ is given and the parameter vector follows $\beta_t = 0.5 + 0.3\beta_{t-1} + \varepsilon_t^\beta$. Finally, the equation $y_t = 0.5 + \beta_t x_t + \varepsilon_t^y$ characterize the cointegrating vector between y_t and x_t, where, ε_t^x, ε_t^y and ε_t^β are random drawings from independent normal distributions each with mean 0. In addition, contrary to DGP 3, we have $\varepsilon_t^\beta \sim N(0, 0.02)$. Moreover, Following table shows mean and standard deviations of β_t for different specifications:

Table 13. Mean and Standard deviation of β_t

	mean(β_t)	std(β_t)
κ=0.7 and ε_t^β ~N(0,0.2)	1.667	0.626
κ=0.7 and ε_t^β ~N(0,0.02)	1.667	0.198
κ=0.3 and ε_t^β ~N(0,0.2)	0.714	0.469
κ=0.3 and ε_t^β ~N(0,0.02)	0.714	0.148

In the estimation of these data sets, we will focus on two important aspects of estimation. First, we aim to analyze the performance of Kalman Filter in predicting the dependent variable y_t. However, prediction of y_t is not sufficient by itself to conclude that Kalman Filter performs well in estimation of time varying cointegrated series. Thus, we will also expound how well estimated and predicted state vector grasp the true TVP data generation process. On the other hand, in the literature, there is no simulation based study, which analyze the implications of Kalman Filter estimation of cointegrated series via TVP. Therefore, we believe interesting results from simulations may emerge.

We first consider the relevant statistics of Kalman Filter recursions with ML estimation. Table 14 will display the estimation results for the data sets with sample size 100. In table 14, it is seen that Kalman Filter produces better fit by means of average RMSE when true parameter is less volatile. This indicates that volatility of true data generation process of TVP impacts the degree of fit. Another important point is that RMSE calculated in this section is never as small as in independent random walks and linearly related random walks cases with same sample size. Thus, Kalman Filter exhibits poorer performance under time varying parameter specification. However, how can we explain this situation? The potential answer for this question may be related to nature of data generation process. Volatility in the parameter vector may restrict Kalman filter to obtain better fit. On the other hand, mean and standard deviations of estimated and predicted state are downward biased according to table 13. Thus, Kalman Filter cannot catch the true data generation process as accurately as in linearly related random walks case. What is more, when $\mathrm{var}(\varepsilon_t^{\beta}) = 0.2$ volatility of estimated parameter process is as high as it is in Independent random walk case. This indicates that a high variance of a_t is not a valid implication of spuriousness in the Kalman Filter estimation.

Above we have mentioned some problems with estimation of TVP models. We will clarify these ideas with help of figures. First, the figures 25 (DGP 1), 27 (DGP 2), 29 (DGP 3) and 31 (DGP 4) depicts that the Kalman Filter predictions actually capture the actual DGP with some slight deviations. Hence, Kalman Filter cannot be considered to completely fail in fitting the data, whereas degree of fit is not as good as in graphs of previous estimations. Moreover, we will also consider the performance of Kalman Filter at predicting the true data generation process of the time varying parameter. Figures 26 (DGP 1), 28 (DGP 2), 30 (DGP 3) and 32 (DGP 4), which illustrate a_t, $a_{t|t-1}$ and actual parameter vector β_t on the same graph, will be helpful in this analysis. There is a common pattern in all these figures. a_t follows a closer path with underlying DGP of TVP even though $a_{t|t-1}$ can capture the trend of these series. Furthermore, Table 14 also shows that a_t is more volatile than $a_{t|t-1}$.

Consequently, updated state estimation has better performance at capturing true parameter process than predicted state in these cases. Furthermore, since $a_{t|t-1}$ can only explain some portion of true DGP, predicted value of y_t seems not to grasp actual value. Since Kalman Filter produce a_t after observing y_t, a_t can capture β_t better than $a_{t|t-1}$.

Table 14. Identification of the problem: TVP Related Random Walks Case with Unknown System Matrices Sample Size=100

N=100			x~I(1) , β(t)=0.5+κβ(t-1)+ξ(t)				
iters=3000			y(t)=0.5+β(t)x(t)+ε(t)				
	mean(at)	**stdc(at)**	**mean(at_1)**	**stdc(at_1)**	**coeffvar**	**corr(y/x~at)**	**RMSE**
κ=0.7 and ε_t^{β} ~N(0,0.2)	1.437	0.571	1.431	0.392	1.290	0.784	0.7073
κ=0.7and ε_t^{β} ~N(0,0.02)	1.608	0.166	1.599	0.142	0.106	0.694	0.2905
κ=0.3 and ε_t^{β} ~N(0,0.2)	0.682	0.410	0.682	0.116	0.460	0.768	0.6891
κ=0.3 and ε_t^{β} ~N(0,0.02)	0.705	0.116	0.704	0.033	0.172	0.684	0.2463

On the other hand, we also have histograms that expose the distribution of the ML estimates over 3000 iteration. These histograms are exhibited in the tables 22, 23, 24 and 25 (Appendix B) respectively for DGP 1, DGP 2, DGP 3 and DGP 4. In all of these tables, distributions of each parameter exhibit very similar patterns. First, α^{mle} seems to have a normal shaped distribution around 0.5 which is the true value of the intercept parameter. However, there are considerable accumulation in both tails of the distribution of α^{mle}. Thus, these accumulations may cause the problems mentioned above. Moreover, T^{mle} has again a normal shaped distribution around 0.7 for data generated with $\kappa = 0.7$ and around 0.3 for data generated with $\kappa = 0.3$. Hence, the ML estimation successfully captures the true value of κ. Another parameter that determines the value of the state vector is μ^{mle}. In all cases, μ^{mle} seems to have very similar distribution. Tables exhibit a normal shaped distribution for μ^{mle} that clusters around 0.5. Finally, we will examine the distribution of σ_q which determines the volatility of the state vector. There are two different results for σ_q. In the data generation processes of TVP with variance 0.2, the ML estimator of σ_q is amassed around 0.43. This value is close to square root of 0.2. Further, we can observe that the ML estimates σ_q is accumulated near 0.14 for the TVP data generation process of which variance is 0.02. Thus, σ_q is correctly estimated by the ML, and it seems to be normally distributed. Consequently, even though we do not have theoretical explanation for the distributions for the MLE estimates in Kalman Filter, we may comment all parameters are probably distributed normally around their true value. However, there are serious distortions in the histogram of α^{mle}. These distortions may cause that the mean and the standard deviation of a_t and $a_{t|t-1}$ cannot match with the mean and the standard error of β_t.

On the other hand, we also consider the case in which one have 400 observation in his sample. The summary of the result for these simulations are given in Table 15.

Table 15. Identification of the problem: TVP Related Random Walks Case with Unknown System Matrices Sample Size=400

Sample Size=400		Q=0.2	H=1	x~I(1) , β(t)=0.5+κβ(t-1)+ξ(t)			
iters=3000				y(t)=0.5+β(t)x(t)+ε(t)			
	mean(at)	stdc(at)	mean(at1)	stdc(at1)	coeffvar	corr(y/x~at)	RMSE
κ=0.7 and ε_t^{β} ~N(0,0.2)	1.623	0.595	1.621	0.415	0.375	0.682	0.436
κ=0.7and ε_t^{β} ~N(0,0.02)	1.658	0.173	1.656	0.132	0.105	0.566	0.157
κ=0.3 and ε_t^{β} ~N(0,0.2)	0.710	0.435	0.710	0.129	0.640	0.654	0.432
κ=0.3 and ε_t^{β} ~N(0,0.02)	0.713	0.122	0.713	0.036	0.171	0.549	0.148

Table 15 exhibits very similar results as the table 14. However, for each choice of κ and $\text{var}(\varepsilon_t^{\beta})$, in the simulations with sample size 400 Kalman filter exhibits better degree of fit by means of having lower RMSE. Therefore, Kalman Filter generates more accurate prediction for the large samples. However, in the large samples, the estimations of linearly related and independently generated random walks have lower RMSE. Furthermore, other statistics are not considerably different from the statistics in sample size 100 case.

The histograms for N=400 also depicts similar features. However, this time all parameters are accumulated near their true value with significantly small variance. Thus, as sample size grows, the accuracy of the ML estimation increases. Tables 26, 27, 28 and 29 in appendix B depict the histograms for the DGP 1, 2, 3 and 4 respectively. As a consequence, higher accuracy of the ML estimates in large samples provides better prediction of y_t, which is also indicated by lower RMSE.

To summarize, our simulations indicate that the ML estimation is generally successful in capturing the true parameter values of the TVP data generation processes with AR(1) parameter evolution. However, the higher RMSEs obtained in theses simulations indicate relatively low performance of Kalman filter in fitting the data. We mentioned some potential sources of this result. First, since we report the average values of the RMSE, some samples may be too far away from the general tendency. Further, in the histogram of α^{mle} we observe sharp deviations from the true value of the parameter. Thus, we suspect that the estimated value of α^{mle} causes problems in these samples. These deviations do not only influence the value of the intercept term of the model, but also impact other parameters of the model. Finally, even though there are some distortions in the estimation of the parameters, Kalman Filter still generates good prediction which can be observed from figures 21-32.

4.2.2.3.2. Time Varying Parameters Generated as Smooth Transitions

In the previous exercise, we examine the TVP model generated with high volatility. However, we will deal with more stable TVP processes in this section. Thus, we will examine performance of Kalman filter on such series. We also consider whether the problems encountered in the previous section repeats in the TVP generated as smooth transitions. We expect Kalman Filter can produce better results since these series are not as volatile as TVP following AR(1) process. We have two different sets of variables for this case. In these sets, the parameter processes differ from each other according to their curvature :

1). Regressor x_t is generated as driftless unit root as: $x_t = x_{t-1} + \varepsilon_t^x$ and the TVP process is given by $\beta_t = 0.5 + (t/N)^{0.5}$ where N is the sample size. The relation between regressor and dependent variable is governed by the equation: $y_t = 0.5 + \beta_t x_t + \varepsilon_t^y$ and ε_t^x, ε_t^y are independent standard normal variables. As we mentioned, this type of DGP offers convex functional form for the time varying parameter. Moreover, the mean of β_t is calculated as 1.43 and the variance of β_t is

0.0015 for the sample size 100. For the sample size 400, these values are given as 1.33 for the mean and 0.009 for the variance of β_t.

2). In this case, we consider concave parameter process, which is determined by: $\beta_t = \beta_x + (t/N)^{1.5}$ where N is the sample size. Moreover, the cointegration relation is given by $y_t = \beta_t x_t + \varepsilon_t^y$ and ε_t^x, ε_t^y are independent standard normal variables. In addition, for sample size 100, mean and variance of β_t are 1.32 and 0.01 respectively. finally, mean value of β_t is 1.11 and variance of β_t is 0.04 for sample size 400.

We generated above series with sample size 100 and 400 to check how Kalman Filter performs in estimation of such series. In order to run Kalman Filter recursions with the unknown system matrices, we adopt the state-space form given in equation (31). The parameters of the model are again estimated by ML. Following table exhibits the estimation results and relevant statistics for state vector and RMSE:

Table 16. Identification of the problem: TVP (Smooth Transition) Related Random Walks Case with Unknown System Matrices

Sample Size=100				x~I(1) , β(t)=0.5+(t/N)^γ			
	mean(at)	stdc(at)	mean(at1)	stdc(at1)	coeffvar	corr(y/x~at)	RMSE
γ=0.5	1.435	0.037	1.428	0.081	0.143	0.407	0.194
γ=1.5	1.317	0.098	1.308	0.147	0.075	0.568	0.195
Sample Size=400							
γ=0.5	1.339	0.103	1.336	0.122	0.077	0.454	0.068
γ=1.5	1.116	0.211	1.114	0.218	0.189	0.546	0.059

The above table exhibits the estimation results for sample size 100 and 400. From these tables we can see that Kalman Filter does not generate state estimates and predictions as volatile as in the TVP generated as the AR(1) processes. This picture possibly appears because of the absence of the noise term in the parameter process. We can also see that the parameter γ slightly affects the volatility in a_t and the mean value of a_t. On the other hand, the mean and variance of a_t and $a_{t|t-1}$ are computed very close to the mean and the variance of true parameter process β_t. Hence, Kalman filter creates the estimated and predicted state vector around the mean of β_t with almost same variance. Additionally, the histograms of all of the data generation processes (Tables 30-33 in appendix B) exhibits similar picture for ML estimates of α , T , μ and σ_q . However, the shapes of distributions are clearer for sample size 400. From these histograms, it can be observed that the estimated value of T is accumulated around 1. We also see some outliers for T around zero, whereas these outliers do not change the analysis too much. However, there is an interesting result for μ . The ML estimator μ_{MLE} is approximately zero for all sample size and γ . Thus, we can conclude Kalman Filter ignores the constant term in the data generation of β_t and include the time invariant part of the state in the time variant part. Furthermore, the table 16 reports smaller RMSE than table 14 and 15 for all cases. Thus, Kalman filter produces better prediction for y_t in the estimation of the time varying parameters with low volatility. This is an expected result since the volatility of the true parameter vector also increase the variation in the dependent variable. As a consequence, increase in volatility of y_t reduces traceability of the underlying model by Kalman Filter, since noisy data will cause Kalman Filter underperforms.

On the other hand, the figures 33 (for γ=0.5) and 35 (γ=1.5) illustrate the results of the estimation of the time varying parameters in the smooth transition model with Kalman Filter where the sample size is equal to 100. These figures verify above comments. Therefore, Kalman filter algorithm with ML estimation yields very good prediction of the data. Additionally, in all cases, Kalman Filter algorithm estimates the time varying parameter very close to the actual parameter process. Figures 34 and 26 exhibits the relation between the estimated state vector and the actual parameter process for the different choice of γ. The predicted and estimated state vectors seem to follow exactly the same path as β_t . Therefore, performance of Kalman filter is excellent in the estimation of the time varying parameter models with smooth transitions. The only problematic situation is that the ML estimate of μ is miscomputed, whereas this does not impact the estimation performance.

In conclusion, this chapter includes the simulations with 4 different types of the data generation process. In all of these simulations, Kalman Filter performs well in predicting the dependent variable y_t with the knowledge of the regressor x_t . However, a big problem emerges at the first set of the data generation process, namely the unrelated random walk series. In the unrelated random walk case, we expect that Kalman Filter detect the inexistence of relation between variables. However, this is not the case. Instead, Kalman Filter produces a nonsense state vector that links y_t and x_t. Moreover, the emergence of meaningless state vector is accompanied with highly volatile predicted and estimated state sequences. The high variations in a_t and $a_{t|t-1}$ do not help us detect spurious regression yet, since the time varying parameter models can also yield high variation in the state according to variance of the TV parameter. In addition, Kalman Filter underperforms in the estimation of TVP following AR(1) processes. We encounter with the inaccuracies in estimation of β_t. We think these imprecision potentially emerge because of high volatility of the true parameter process, as we do not face with the same problem in time fixed or smooth transition models. These models have excellent performance in capturing the true model. On the other hand, we encounter very different estimates for T in each data generation process. While ML estimate of T is computed as 0 in the linearly related random walk case, this parameter is estimated as approximately 1 in the estimation of independent random walks and TVP model with smooth transitions. Therefore, the estimated value of T does not give clue about the existence of the spurious regression.

CHAPTER 5

AN ATTEMPT TO OVERCOME SPURIOUS REGRESSION IN KALMAN FILTER

In previous sections, by the help of the simulations, we seek to demonstrate that Kalman filter is vulnerable against the Spurious regression problem. In these simulations, we try to estimate independently generated random walk series with Kalman filter. The Kalman filter estimation with these series yields specious time varying parameter estimates, which cause the Spurious regression problem. In addition, it is not possible to detect this problem by simple methods such as inspecting fluctuations in the estimated time varying parameter. Therefore, we will develop a new method so as to distinguish the nonsense regression from regression between related series. This new method requires penalizing the Kalman Filter equations in order to prevent the nonsense relations. However, we should be careful about penalizing the related series by mistake. Thus, we aim to propose a penalty term which will reduce the degree of fit in the estimation of the unrelated series, but do not affect the estimation of the related series too much.

However, how can we create such a penalty term? At this point, the results of chapter 4 will give us a lead. In chapter 4, we show that as T_t approach to zero, the degree of fit for both noncointegrated and related random walks declines. Additionally, in both type of series best fit of the data is achieved when T_t is equal to 1. In our method, we will exploit this feature of Kalman Filter recursions by imposing a penalty term on the system matrix T_t. This penalty term will take values very close to 1 in the estimation of related series, and it will move away from 1 in the existence of independent random walk series in the regression equation. On the other hand, a penalty term possessing these patterns can be obtained by the help of the Cochrane's variance ratio statistic, which is used in identifying unit roots in time series. We will first define this statistic. The following equations describe the variance ratio statistics VR(q), where q is lag difference:

$$VR(q) = \frac{\sigma_c^2(q)}{\sigma_a^2(q)} \tag{33}$$

where, $\sigma_c^2(q)$ is an unbiased estimator of 1/q of the variance of the q-th difference of the variable u_t, and $\sigma_a^2(q)$ is an unbiased estimator of the variance of the first difference of u_t. The formulas for calculating $\sigma_c^2(q)$ and $\sigma_a^2(q)$ are given below in equations (34) and (35):

$$\sigma_c^2(q) = \frac{1}{m}\sum_{t=q}^{T}(u_t - u_{t-q} - q\hat{\mu})^2 \tag{34}$$

where $m = q(T-q+1)(1-q/T)$ and

$$\sigma_a^2(q) = \frac{1}{T-1}\sum_{t=2}^{T}(u_t - u_{t-1} - \hat{\mu})^2 \tag{35}$$

where $\hat{\mu} = \frac{1}{T}(u_t - u_1)$ (Liu and He, 1991).

In our case, u_t is residuals obtained from the OLS regression of y_t on a constant and x_t. On the other hand, we calculate recursive variance ratio statistics since we seek to create an endogenous series with same length as the regression variables. After obtaining recursive variance ratio statistics we subtract whole series from one so as to create the penalty term denoted as dv_t.

$$dv_t = 1 - VR_t(q) \tag{36}$$

Before reporting estimation results of penalized Kalman Filter estimation, it is important to understand the patterns of the penalty term dv_t. Histograms in tables 34-37 show how the mean of dv_t is distributed over 3000 simulations. In these tables, we observe that in related series mean of dv_t seems to accumulate between 0.9 and 1. Furthermore, the mean value of dv_t spreads far from 1 for independent random walk processes. Therefore, the penalty term dv_t will serve our purpose to penalize nonsense regression. Moreover, the impact of this penalty term on estimation of related series will not be as much as in unrelated series case.

Table 17. Comparison of Summary Statistics of Kalman Filter Estimation when Penalty Applied and not Applied N=100

N=100								
Unrelated		**mean(at)**	**std(at)**	**mean(at_1)**	**std(at_1)**	**coeffvar**	**corr(y/x,at)**	**rmse**
x~I(1) and y~I(1) are independently generated	penalty	-0.009	0.254	-0.009	0.172	0.441	0.431	0.288
	no penalty	-0.013	0.391	-0.013	0.373	-0.547	0.500	0.143
Linearly Related								
x~I(1), y=0.5+0.9x+ε	penalty	0.534	0.131	0.528	0.142	0.419	0.248	0.147
	no penalty	0.899	0.017	0.899	0.010	0.019	0.340	0.099
TVP								
κ=0.7 and std(ξ)=0.2	penalty	0.410	0.312	0.407	0.271	0.804	0.583	0.730
	no penalty	1.437	0.571	1.431	0.392	1.290	0.783	0.707
TVP								
κ=0.7 and std(ξ)=0.02	penalty	0.408	0.140	0.403	0.137	0.277	0.428	0.272
	no penalty	1.608	0.166	1.599	0.142	0.106	0.694	0.291
TVP								
κ=0.3 and std(ξ)=0.2	penalty	0.957	0.620	0.951	0.489	0.681	0.740	0.774
	no penalty	0.682	0.409	0.682	0.116	0.459	0.768	0.689
TVP								
κ=0.3 and std(ξ)=0.02	penalty	0.972	0.290	0.964	0.281	1.070	0.513	0.335
	no penalty	0.705	0.116	0.703	0.033	0.172	0.684	0.246
TVP Smooth								
γ=0.5	penalty	0.917	0.196	0.908	0.223	0.420	0.355	0.181
	no penalty	1.435	0.037	1.428	0.081	0.026	0.407	0.194
TVP Smooth								
γ=1.5	penalty	0.885	0.206	0.879	0.225	0.192	0.492	0.171
	no penalty	1.317	0.098	1.308	0.147	0.075	0.568	0.195

Table 18. Comparison of Summary Statistics of Kalman Filter Estimation when Penalty Applied and not Applied N=400

N=400								
				Noncointegrated				
x~I(1) y~I(1) are independently generated		mean(at)	std(at)	mean(at_1)	std(at_1)	coeffvar	corr(y/x~at)	rmse
	penalty	-0.006	0.292	-0.007	0.174	2.524	0.300	0.302
	no penalty	-0.008	0.709	-0.008	0.698	0.834	0.381	0.098
				Linearly Related				
x~I(1), y=0.5+0.9x+ε	penalty	0.725	0.110	0.722	0.120	0.203	0.133	0.070
	no penalty	0.900	0.006	0.900	0.005	0.007	0.189	0.050

Table 18 Continued. Comparison of Summary Statistics of Kalman Filter Estimation when Penalty Applied and not Applied N=400

				TVP				
κ=0.7 and std(ξ)=0.2	penalty	0.573	0.369	0.571	0.331	0.971	0.548	0.476
	no penalty	1.623	0.595	1.621	0.415	0.375	0.682	0.436
				TVP				
κ=0.7 and std(ξ)=0.02	penalty	0.571	0.133	0.569	0.129	0.269	0.365	0.164
	no penalty	1.658	0.173	1.656	0.132	0.105	0.566	0.156
				TVP				
κ=0.3 and std(ξ)=0.2	penalty	1.334	0.629	1.331	0.506	0.632	0.655	0.461
	no penalty	0.71	0.435	0.71	0.129	0.64	0.654	0.432
				TVP				
κ=0.3 and std(ξ)=0.02	penalty	1.35	0.257	1.347	0.25	0.248	0.444	0.187
	no penalty	0.713	0.122	0.713	0.036	0.171	0.549	0.148
				TVP Smooth				
γ=0.5	penalty	1.134	0.185	1.131	0.2	0.18	0.407	0.077
	no penalty	1.339	0.103	1.336	0.122	0.077	0.454	0.068
				TVP Smooth				
γ=1.5	penalty	0.988	0.229	0.986	0.225	0.246	0.523	0.079
	no penalty	1.116	0.211	1.114	0.218	0.189	0.546	0.059

On the other hand, in this chapter, the system matrices d_t, c_t, H_t and Q_t are assumed to be unknown. Thus, we will adopt a ML estimation with Kalman filter. Further, in the previous chapter, T_t is also considered as constant unknown parameter, which is estimated by the MLE. However, we will replace T_t with dv_t, which is endogenously generated. After this modification, the measurement and transition equations are given by (37).

$$\begin{aligned} y_t &= \alpha + \beta_t x_t + \varepsilon_t \quad \varepsilon_t \sim N(0,\sigma_h^2) \\ \beta_t &= \mu + dv_t \beta_{t-1} + \eta_t \quad \eta_t \sim N(0,\sigma_q^2) \end{aligned} \tag{37}$$

Above state-space form will be estimated by Kalman Filter and the MLE procedures. The Kalman filter recursions will be computed by the prediction and updating equations described in chapter 3. Additionally, MLE will be used to estimate the unknown parameters of the model, which is given by vector $\Phi = (\alpha,\ \mu,\ \sigma_h,\ \sigma_q)$. Moreover, we generate 3000 different pairs of random walk without drift variables with N=100 and 400 for each data generation processes. These series are exactly the same set of variables used in chapter 4. We will apply the modified Kalman Filter to these sets. The summary statistics for these estimations can be seen in 17 and 18 table for sample size 100 and 400 respectively. These tables also includes the estimation results when the penalty term is not applied. Therefore, we can monitor the changes in these estimations after the penalty term is applied. In addition, so as to identify overall difference between the estimation with penalty and without penalty cases, we will present another table (table 19), which exhibits the percentage change in some statistics. Thus, table 19 will be very useful in interpreting the results. Moreover, Negative values in table 19 indicates a fall in the value of the statistic with respect to without penalty case.

Table 19. Percentage Change in Key Summary Statistics

		mean(at)	stdc(at)	mean(at_1)	std(at_1)	RMSE
Noncointegrated x~I(1) y~I(1) are independently generated	N=100	-29.84%	-35.03%	-30.53%	-53.88%	101.70%
	N=400	-16.36%	-58.84%	-10.98%	-75.08%	208.04%
Linearly Related x~I(1), y=0.5+0.9x+ε	N=100	-40.59%	653.55%	-41.23%	1356.95%	48.31%
	N=400	-19.49%	1595.19%	-19.76%	2349.41%	39.08%
TVP κ=0.7 and var(ξ)=0.2	N=100	-71.45%	-45.24%	-71.59%	-30.86%	3.17%
	N=400	-64.72%	-37.89%	-64.76%	-20.16%	9.03%
TVP κ=0.7 and var(ξ)=0.02	N=100	-74.63%	-15.94%	-74.77%	-3.39%	-6.53%
	N=400	-65.57%	-23.42%	-65.62%	-1.60%	4.60%
TVP κ=0.3 and var(ξ)=0.2	N=100	40.24%	51.41%	39.51%	321.18%	12.32%
	N=400	87.78%	44.64%	87.47%	292.28%	6.56%
TVP κ=0.3 and var(ξ)=0.02	N=100	37.95%	150.27%	37.03%	745.79%	36.15%
	N=400	89.39%	111.83%	89.08%	594.01%	26.45%
TVP Smooth γ=0.5	N=100	-36.11%	434.78%	-36.41%	177.18%	-6.89%
	N=400	-15.29%	79.71%	-15.33%	63.29%	13.25%
TVP Smooth γ=1.5	N=100	-32.78%	109.17%	-32.83%	53.12%	-12.06%
	N=400	-11.44%	8.56%	-11.44%	3.29%	32.12%

5.1. Estimation with Independent Random walks when Penalty Applied

As mentioned before, the main purpose of penalizing Kalman filter is to prevent spurious regression. Since spurious regression emerges in the regression of the unrelated random walk variables, we begin our analysis with the estimation of these series with penalized Kalman filter. We will apply the penalized Kalman Filter to the same samples of the independently generated random walk variables which are used in chapter 4. However, we only include driftless series in our simulations.

The results of these simulations are given in the tables 17, 18 and 19. These tables demonstrate that the major impact of the penalty term on RMSE appears in the Kalman filter estimation of the independent random walk processes. This impact appears as a significant rise in the RMSE obtained from the estimation of the unrelated series. Table 19 clarifies the idea: the percentage increase in RMSE exceeds %100 for sample size 100 and %200 for sample size 400. On the other hand, adding a penalty term to the Kalman filter equations mitigates the standard deviations of the estimated and the predicted value of the time varying parameter β_t. This implies that Kalman filter cannot produce the state estimates as flexibly as in without penalty case. In these results, the penalty term dv_t has two way impact on the Kalman Filter recursions. First, dv_t involves in the prediction equation of the state vector (see equation 14) as a multiplier of the previous periods state estimate a_{t-1}. Further, since dv_t is generated as a sequence which takes values close to 0, penalty term will pull down the value of state prediction in each period. Moreover, the second effect of dv_t is on the predicted value of the variance-covariance matrix of the estimation error. As seen in the equation 16, dv_t will decrease $P_{t|t-1}$, so $P_{t|t-1}$ will be computed with downward bias. As a consequence, Kalman Filter underestimates the estimation error. Furthermore, since $P_{t-|t-2}$ is also miscomputed, this will cause a fall in a_{t-1} via updating equation 16. Hence, the fall in a_{t-1} will provide $a_{t|t-1}$ to move away from the spurious state. On the other hand, the affects of dv_t do not cause any significant change in the mean value of the state vector. This may imply that the main reason of the fall in the RMSE is due to reduction in the volatility of the state vector.

In order to see the situation clearly, we will check the related figures. In the figure 37 for N=100, the predicted value of y_t clearly deviates from the actual value of y_t. Thus, this picture illustrates that the penalty term dv_t successfully separate $y_{t|t-1}$ from

y_t. As a result, the spurious relation between y_t and x_t is precluded. Furthermore, the figure 38 clearly exhibits that $a_{t|t-1}$ and a_t do not move in the same path with z_t as in the spurious regression case. We also view less volatile state prediction $a_{t|t-1}$ which tends to move along 0. Hence, by this method, Kalman filter no more creates a false parameter process which result in spurious fit.

Moreover, the histogram of ML estimator of the unknown parameters is exhibited in the table 38. In this table, α_{mle} seems not to have a proper distribution as in without penalty case. The estimated values of this parameter spreads on both tails of the bin. Thus, the penalty term in estimation does not have much impact on the distribution of α_{mle}. Another parameter, which the penalty term does not influence, is the constant term of the transition equation μ_{mle}. The distribution of μ_{mle} is almost same as in without penalty case. This parameter is accumulated around 0.1. Therefore, μ_{mle} has very little impact on the state evolution. What is more, the distribution of σ_q is amassed around 0. Thus, so volatility of state vector will be very low in this case.

In conclusion, the above findings indicate that the proposed penalty term is successful in removing effects of the spurious regression in Kalman Filter estimation. The main reason of this result seems to be the reduction in volatility of the estimated and predicted state vector, which is provided by both dv_t and the ML estimate of σ_q. In addition, the only problematic situation seems to be the ML estimate of α_{mle} does not possess a proper distribution. However, the distribution of this parameter does not impact the state vector, but only the level of the predicted y_t.

5.2. Estimation with Linearly Related Random Walks when Penalty Applied

In order to see the performance of the penalized Kalman Filter in the estimation of the related series, we first apply this method to the regression of the linearly related random walk variables. In order to compare the results of penalized Kalman filter estimation with the previous exercises, we generate exactly the same set of the linearly related variables used in chapter 4.

The estimation results of both with and without penalty cases can be observed in the tables 17, 18 and 19. As seen in these tables, the estimation results for this DGP strictly differ from the estimation results of the independent random walk processes. However, there are some distortions in the mean value and the standard deviation of the estimated state vector relative to without penalty case. These distortions can be classified as the increase in the volatility of the estimate of the time varying parameter and a fall in the RMSE and the mean value of the state vector. We guess the decrease in RMSE is fed by increase in volatility of state vector. Since the relation between variables is linear, fluctuations of state vector results in small bias in the estimation of the slope parameter. As a result, these small biases impact the RMSE since the predicted value of y_t deviates from its actual value a little. Moreover, the mean value of the predicted state is 3.5 point less than the actual value of the slope parameter. Nevertheless, the presence of such distortions does not prevent Kalman Filter to obtain good fit of the data. Figure 39 for sample size 100 will clarify this argument. In this figure, it seems Kalman Filter still can predict y_t as good as without penalty case. Additionally, Kalman Filter also capture the true value of the slope coefficient. The figure 40 exhibits the estimated value of the state for sample size 100. In this graph, a_t and $a_{t|t-1}$ move along the true value of the slope parameter. However, Kalman Filter requires some training period to reach the actual slope parameter. In addition, we cannot see the exaggerated volatility increase in a_t and $a_{t|t-1}$ in this graph, since the standard deviations of a_t and $a_{t|t-1}$ is initially so small that 0.1 point increase is calculated as very high percentage change. Further, the mean of a_t and $a_{t|t-1}$ is downward biased, since Kalman Filter cannot catch the true parameter in initial periods. We guess that the initial values of dv_t may cause this problem. Nevertheless, when dv_t converge to 1 after some periods, Kalman filter creates better fit of the data.

The most interesting feature of this estimation appears in the ML estimates of the unknown parameters. First of all, α_{mle} seems to cluster in two tails of the bin. Thus, we cannot say intercept term is successfully captured. Moreover, in without penalty case, μ_{mle} is accumulated near 0.9, which is true value of slope parameter. However, the distribution of μ_{mle} is very different when we apply the penalty term. μ_{mle} is amassed around 0.1 in the absence of the parameter T in the model. This result implies that the slope parameter of the model is determined by the time variant part of the state vector. Since we control T_t by the penalty term, μ_{mle} tends to zero in the presence of other factors determining the state estimation and prediction. Finally, σ_q is accumulated around 0.2 as expected.

To sum up, the penalized Kalman Filter will show poorer performance in the estimation of linearly related series relative to without penalty case. Nevertheless, the performance loss of the Kalman filter is not as bad as in the estimation independent series. In addition, the figures do not exhibit deep differences from the estimation results of without penalty case. Therefore, our method is partially successful in the estimation of the linearly related unit root processes although some problems emerge.

5.3. Estimation with TVP when Penalty Applied

The time varying parameter models have important place in the econometrics literature. In the previous chapter, we have shown that these models can be favorably estimated by the Kalman Filter. However, estimating these models with the penalized Kalman Filter may different implications. Thus, we will investigate the indications of the penalized Kalman filter estimation of these sets of variables On the other hand, we will consider the estimation of two types of TVP models via modified Kalman Filter. These specifications are TVP following AR(1) process and TVP generated as smooth transitions. The first type of DGP is volatile relative the smooth transition model.

5.3.1. Estimation with TVP generated as AR(1) Process When Penalty Applied

In the first type of the time varying parameter model, the parameter sequence follows AR(1) process with drift. This specification offers highly volatile parameter process which also yields fluctuant dependent variables. In addition, Kalman filter also creates very volatile state estimates and predictions under this specification. This characteristic of TVP following AR(1) process resembles independently generated random walk processes. Thus it will be interesting to inspect whether the modified Kalman Filter can distinguish these series from the unrelated random walk variables.

In order to see this, we generate the data as described in section 1 of chapter 4 and compute dv_t sequence for each sample. Tables 34-27 show the distribution of the mean value of the penalty term dv_t, which assembles near 0.9 for $\kappa = 0.3$ and 0.8 for $\kappa = 0.7$. Thus, we do not expect a picture like the independent random walk estimation.

The estimation results can be seen in tables 17 and 18. From these tables, some minor distortion in the RMSE and the standard deviation of a_t can be observed. As it can be seen from the table 19 highest distortion in the average RMSE does not exceed %37 in all cases. This indicates that penalty term does not affect the degree of fit as adversely as in the independent random walks case. On contrary, there is an improvement in the estimation of TVP with parameters $\kappa = 0.7$ and $\mathrm{var}(\varepsilon^{\beta}) = 0.02$. Additionally, some considerable change in the mean of a_t and $a_{t|t-1}$ attract attention, although these distortions do not impact the performance of the Kalman Filter in predicting y_t. With these distortions, the mean value of the estimated and predicted state vector are too far from the mean of the true parameter process. The reason is twofold. First, the constant term in the state vector μ_{mle} is computed with downward bias. Tables 40-43 shows that for $\kappa = 0.7$ μ_{mle} clusters around 0.3, which is 0.2 points less than true parameter value. However, μ_{mle} is accumulated around 0.1 when κ is equal to 0.3. Thus, μ_{mle} is downward biased for both choice of κ. Another reason of the distortion in the mean value of the state vector seems to be relevant to T_t. In estimation without penalty term, optimal T is computed around 0.7 for $\kappa = 0.7$ and 0.3 when κ is 0.3. Per contra, we use dv_t instead of a constant parameter T. Since dv_t takes values around 0.9 and 0.8 for $\kappa = 0.7$ and 0.3 respectively, in both cases Kalman Filter unnecessarily update the state and the variance covariance matrix of prediction error. In addition, the sign of the change in mean of a_t and $a_{t|t-1}$ is not same for both cases. We observe an decrease in the mean of a_t and $a_{t|t-1}$ for $\kappa = 0.7$, whereas the average values of a_t and $a_{t|t-1}$ raise when the penalty term is applied to TVP series with $\kappa = 0.3$. On the other hand, the same situation is valid for the change in the standard deviation of the state vector. Table 19 clearly demonstrates that the standard deviation of a_t and $a_{t|t-1}$ is lower than without penalty case when κ is equal to 0.7. However, when $\kappa = 0.3$, the standard deviation of a_t and $a_{t|t-1}$ rise after the application of the penalized Kalman filter

So as to see estimation outcomes better, we will also exhibit some graphs. For sample size 100, the figures 41 (DGP 1), 43 (DGP 2), 45 (DGP 3) and 47 (DGP 4) affirm that the penalized Kalman Filter performs well in fitting the data. In the figures, we can see that $y_{t|t-1}$ follows almost the same path as the actual value of y_t. In addition, an intriguing point is that when standard deviation of ε_t^{β} is high (for the case var(ε_t^{β})=0.2) Kalman recursions create a_t and $a_{t|t-1}$ very close to actual β_t process for both $\kappa = 0.7$ and $\kappa = 0.3$ as seen in figures 42 and 46. In other cases, $a_{t|t-1}$ is less volatile than β_t, whereas, a_t still follows same path as β_t (see figures 44 and 48). Additionally, difference between the mean of the actual time varying parameter and estimated state vector, which is demonstrated in table 19, seems not to be very important. This difference only causes some minor shifts in estimated and predicted state vectors. Thus, penalized Kalman Filter performs fairly well in the estimation of TVP generated as AR(1) process since it creates good prediction for the dependent variable y_t and the true parameter vector.

5.3.2. Estimation with TVP Generated as Smooth Transitions

In this section, we will be interested in the time varying parameters generated as smooth transitions. The time variation in this data generation process is very different from the TVP following AR(1) process. The basic difference between the time variations of these two data generation processes appears in the volatility of the parameter process. The volatility of the time varying coefficient in smooth transition model is very low on contrary to the TVP following AR(1). Moreover, we obtained better performance in the estimation of smooth transition model without penalty term. However, this situations may not be valid for the Kalman Filter with penalty term. In order to see this, we apply the penalized Kalman filter to the TVP model with smooth transitions. Estimation results of these estimations are given in the tables 17, 18 and 19.

The most noteworthy point in these set of estimation is that in small sample (N=100) the average RMSE is lower than without penalty case whatever the choice of γ is. Hence, the penalty term has positive impetus on degree of fit for the small samples. Conversely, we cannot explore same situation in sample size 400 case. We guess the reason of this difference is related to the change in the volatility of the estimated and predicted state vectors. In small samples, we observe very large changes in the standard deviation of a_{t} and $a_{t|t-1}$. However, we cannot detect these change in the volatility of the state vector in large samples. Thus, Kalman Filter may produce better fit of the data since excess volatility in the state vector emerge in small samples. However, the increase in fit may be partially misleading, as both mean and standard deviation of a_{t} and $a_{t|t-1}$ move far from the mean and standard error of the actual parameter process. Therefore, this issue requires further inspection.

On the other hand, graphs may be helpful to see the performance of the modified Kalman filter. Figures 49 and 51 illustrates the strong link between y_t and $y_{t|t-1}$ for $\gamma = 0.5$ and $\gamma = 1.5$ respectively. These graphs show that Kalman Filter generate very successful prediction for y_t. However, the figures 50 ($\gamma = 0.5$) and 52 ($\gamma = 1.5$) demonstrate that $a_{t|t-1}$ is not as successfully as in without penalty case in capturing the true parameter variation since there are some local upward and downward biases. In spite of the fact that Kalman Filter estimates the time varying coefficient with bias in all of the cases and sample sizes, intriguingly this bias does not reflect on the prediction of y_t. We guess there may be two reason behind this situation. First, as we adopt a MLE estimation, there may be biases in other parameters. Thus these biases may also create a distortion in the time varying parameter estimation. Furthermore, in the smooth transition model, the time varying parameter does not contain a noise term. However, we try to estimate this parameter by including error term in the state evolution. Therefore, trying to estimate a noiseless parameter with the noisy state vector may cause problems in the estimation. Moreover, the histograms of α^{MLE} and μ^{MLE} (Tables 44 and 45 in appendix) may clarify these ideas. In these tables, for all κ, α^{MLE} has no apparent tendency toward its true value 0.5. However, in estimation without penalty term, ML estimation can successfully capture the true value of the intercept term α. Thus, penalty term engenders distortions in the distribution of α^{MLE} in the smooth transition model. In addition, μ^{MLE} seems to be accumulated around 0.2 which is downward biased from true value 0.5. In presence of these biases, one may think the predicted value of y_t apart from the actual value of y_t. Nevertheless, Kalman Filter compensates these distortions by adjusting the state vector in order to predict y_t close to its actual value. Additionally, since dv_t takes values around 0.9, Kalman Filter can still manage to generate successful y_t prediction in presence of biases in the parameter estimations.

Consequently, as in all related series previously presented, penalized Kalman Filter also does a good job in predicting the dependent series y_t with the knowledge of independent variable x_t in estimation of TVP with smooth transitions. However, in presence of biases in ML estimation of parameters, we cannot confirm that we obtain expected outcomes for the time varying parameter estimation with the penalty term. Therefore, while the penalty term do not impact the prediction of y_t too much, the negative affects of the penalizing Kalman Filter arise in estimation of the time varying slope coefficient.

On the other hand these results are common in all data generation processes used in this thesis excluding independently generated random walk case. In the estimation of independently generated random walk series, predictions of both time varying parameter and y_t is exposed to significant deviations from without penalty case. These deviations clearly indicate the existence of the spurious regression in the Kalman Filter estimation. As a consequence, our method brings forth a good attempt in detection of the spurious regression, however, it is not suitable for estimating cointegrating vector between variables.

CHAPTER 6

DISCUSSION

Spurious regression has been considered as a problem emerging in OLS estimation when two unrelated random walk processes are included in regression equation. This problem engenders a nonsense relation between variables and generates false significance test statistics. The reason behind these false significance tests can be explained by the divergence of t-statistic for the OLS estimator of slope coefficient (Phillips, 1980). Phillips also argues that if this t-statistic is divided by $\sqrt{N}$, then the transformed statistic has a distribution (potentially not standard normal). This implies that conventional (non-transformed) t-statistic for $\hat{\beta}_1$ will exceed the critical values as sample size grows. Therefore, the null hypothesis of the t-statistic ($H_0 : \beta_1 = 0$) will be rejected in this case. Thus, this will yield false significance test results.

On the other hand, Spurious regression may not be peculiar to simple linear regression models and in other estimation methods spuriousness may cause some problems such as false inference about the data. One of these estimation methods, which we consider in our analysis, is time varying parameter models. In order to show that the risk of spurious regression persists also in time varying coefficient models, we use Kalman Filtering, which is accepted as very powerful tool in estimation of the various time series models (Harvey, 1989). Although Kalman filter delivers optimal estimator for the state vector in state-space representation, this does not guarantee the elimination of nonsense regression in estimation. Further, we believe that Kalman Filter worsens the spurious regression problem if two unrelated random walk variables are included in time varying regression equation. With the help of simulations, we demonstrate the emergence of spurious regression in chapter 4. In these simulations, if two unrelated random walk process are included in state space form, Kalman filter delivers a nonzero state vector. This state vector also yields a specious fit of the dependent variable y_t. However, the results of these simulations do not clearly explain the reason behind the emergences of spurious regression. At this point, one should focus on basic properties of Kalman filtering. In Kalman filter algorithm the a_t is defined as conditional mean of true parameter process β_t. Moreover, Harvey (1989) argues that the conditional mean is minimum mean square error estimator of β_t. In other words, a_t minimizes expectation of squared difference between estimated and actual value of the parameter (MSE= $E\left[(\beta_t - a_t)^2\right]$ in this case). Furthermore, the conditional mean of y_t at time $t-1$ (we define this as $y_{t|t-1}$ in equation 26) can be interpreted as minimum mean squared error estimator of y_t. This term also involves in updating equation of Kalman filter (see equation 16) as an adjustment term, which controls for the deviation from actual value of dependent variable. This implies that Kalman Filter both minimizes mean squared error in the forecast of state vector β_t and dependent variable y_t. As a consequence, the optimizing nature of the Kalman Filter will yield state estimates, which extract best prediction of dependent variable y_t with the knowledge of the regressor x_t. One can easily notice that best fit of the data can be achieved if time varying parameter is estimated as a ratio $(y_t - d_t)/x_t$.

In our simulations, we verify that Kalman Filter actually create a state vector which mimics the above ratio. However, this feature of Kalman filter is restricted by the magnitude of state evolution parameter, which actually determines how flexibly Kalman filter can update the state estimates. For instance, in estimation of independently generated random walk processes, the condition $T = 1$ produces nearly perfect fit for the data. In this case, since Kalman filter can freely update state vector, a nonzero parameter vector, which engenders spurious fit of the data, is obtained. Additionally, lowering the value of T causes sharp deviations between predicted and actual value of y_t, while estimated state vector approaches to zero line. This implies that Kalman Filter requires flexibility in updating the state vector in unrelated random walk case. This requirement can also be empirically observed in our simulations. These simulations depicts that in estimation of independent random walk process Kalman filter produces highly volatile state prediction.

Therefore, in chapter 5, we develop a new method in order to provide a solution for spurious regression in Kalman filter. This method seeks to fix Kalman filter recursion with an adjustment in the state evolution parameter. We believe that if one can reduce the value of the state evolution parameter in unrelated random walk case, spurious regression can be detected easily. However, we should be careful about penalizing estimation of related series by mistake. Thereby, we propose a time varying penalty term, which carries different patterns under cointegration and spurious regression cases. These patterns can be summarized as follows: in spurious regression, proposed penalty term moves very close to zero line and under cointegration the sequence of penalty term is computed near 1. Hence, this will yield different results under spurious regression and cointegration. In spurious regression case, Kalman filter cannot flexibly adjust the forecast error, because the presence of small state evolution parameter strictly restricts the volatility of the state estimate. Moreover, in the estimation of cointegrated variables, the proposed penalty term does not impose a severe restriction on the forecast error adjustment of Kalman filter recursions. Thus, Kalman Filter still can create good prediction of the data.

An important question is: How can we create such a penalty term? One of the potential answers to this question points Cochrane's variance ratio statistic. This statistic measures the ratio of the conditional and unconditional variance of residuals obtained in OLS estimation of the series, which are included in time varying regression equation. The conditional variance is also scaled down by the factor determined by the lag length choice in the calculation of the conditional variance. This adjustment has some consequences. First, if the series are independently generated random walks, then the conditional variance tend to rise despite of the presence of the adjustment. This will yield that the ratio is very close to 1. Secondly, under cointegration, both conditional and unconditional mean remain constant. Thus, the adjustment term will pull down the value of the numerator. As a result, the variance ratio is computed as a number close to 0. Under these conditions, we subtract calculated variance ratio from 1, so the role of the variance ratio statistic is reversed and now it is convenient for our purpose. To sum up, this penalty term does not allow Kalman Filter to flexibly adjust the forecast error in spurious regression case. However, in cointegrated variables case, the penalty term provides nearly nonstationary transition equation, which yields no restriction on predicting the data.

CHAPTER 7

CONCLUSION

Spurious regression rose to prominence in the time series models with fixed coefficients. Scholars have designed various tools to detect this problem and built cointegration framework for time invariant parameter models. However, some of these scholars are obliged to conclude that cointegration does not exist in their models. Nonetheless, they sometimes attribute the absence of the cointegration in their study to the misspecification of the fixed parameter models, when they detect spurious regression in their analysis. Furthermore, the time varying parameter models have become very popular among these researchers like Canarella, Pollard and Lai (1990). Whereas, the appearance of Nonsense regression problem in the time varying parameter models, especially when Kalman filter is preferred in the estimation, is still dangerous. Without concerning the danger of the spuriousness, one may misinterpret the data and make wrong policy recommendations. On the other hand, the conventional analysis on the detection of the spurious regression may not be valid for time varying parameter models as Granger noted (1991). This conundrum necessitates the redefinition of Spurious regression problem for the time varying parameter models especially when the Kalman Filter method is used in the estimation. In this study, we attempt to take the first step to redefine spurious regression in Kalman Filter.

In order to expound the spurious regression in the Kalman Filter, we utilize Monte-Carlo simulations. In these simulations, we adopt two different Kalman Filter estimation methods. The first estimation method is Kalman Filter with known system matrices. This method requires the assignment of the system matrices before the estimation. In this thesis, the values of these system matrices are determined by prior knowledge and economic intuition. After assigning proper values to system matrices, we run Kalman Filter recursions with both independently generated and related series. We observe that the results of these estimations are highly sensitive to the state evolution parameter T_t. When T_t is equal to 1, Kalman filter produces a time varying cointegrating vector between the variables whether the series are independent or related. However, the cointegrating vector between independently generated variables seems to be specious, since there is no valid relationship for these variables. Furthermore, decrease in T_t causes the disappearance of cointegration between variables generated as both independent and related series. Therefore, this feature of Kalman filter cannot be used in the identification of the Spurious regression. On the other hand, the second estimation method can be classified as the Kalman Filter with unknown system matrices. In this case, ML estimation method with Kalman recursions is adopted for the estimation of unknown system matrices. We apply this method to the same set of variables generated for the Kalman Filter estimation with known system matrices. The results of these simulations are similar to previous case. In these estimations, Kalman Filter also creates a cointegrating state vector in estimation of both types of data generation processes. Hence, Kalman Filter estimation with known and unknown system matrices cannot prevent spurious regression.

Additionally, in these simulations we find out that the nonsense state vector exhibits a certain pattern. The spurious state vector seems to follow the same path as the ratio of the dependent variable minus the intercept term and the independent variables, which we denote as z_t. This result has a very strong impact on the cointegration tests in the time varying parameter models. As it can be seen, when the state vector follows the ratio z_t, residuals will be stationary instead of following I(1) process. As a consequence, the conventional residual based tests will fail to detect the spurious regression. Since residual based tests are no longer valid, researchers should be cautious about the possibility of the occurrence of the spurious regression problem in the estimation of the time varying parameter models with Kalman Filter. Otherwise, inappropriate policy analysis may be concluded. So as to prevent these inappropriate results to occur, we aim to modify Kalman Filter estimation. This modification suggests replacing T_t with an endogenously generated penalty term in the Kalman filter estimation. In order to do this, we use the results we obtain in the estimation of Kalman filter with known system matrices. In Kalman Filter estimation with known system matrices, we are confronted with a strong direct relation between T_t and the predictive power of Kalman Filter. In other words, the predictive power of Kalman filter will fall as T_t moves from 1 to 0. In our model, the proposed penalty term exploits this pattern. Furthermore, Cochrane's variance ratio test is utilized so as to create a proper penalty term, which is a sequence with the same length as the data. This sequence is shown to have different patterns in different data generation processes. Finally, these series are replaced with T_t and the Kalman Filter recursions are initialized. However, the estimation results with these penalty terms are not very effective in terms of obtaining good prediction of true parameter process in models with related series. Nonetheless, we successfully detect spurious regression when the independent random walk series are used in the estimation. In spite of some minor distortions in estimation of related series, Kalman Filter also performs well in predicting the dependent variable. In conclusion, our attempt may draw attention to the spurious regression in Kalman Filter estimation of the time varying parameter models.

SELECTED BIBLIOGRAPHY

Box, G.E.P. and P. Newbold, 1971. Some comments on a paper of Coen, Gomme and Kendall. Journal of the Royal Statistical Society. Vol. 134, No. 2 ,pp. 229-240.

Canarella, G. , S. K. Pollard and K.S. Lai, 1990. Cointegration between Exchange rates and Relative Prices: Another View. European Economic Review,Vol. 34, pp1303-1322.

Cooley T. F. and E C. Prescott, 1976. Estimation in the Presence of Stochastic Parameter Variation. Econometrica, Vol. 44, No. 1 , pp. 167-184

Durlauf, S. N. and P. C. B. Phillips, 1988. Trends versus Random Walks in Time Series Analysis. Econometrica, Vol. 56, No. 6 , pp. 1333-1354

Engle, R. F. , C. W. J. Granger, 1991. Long-run Economic Relationships: Readings in Cointegration, Oxford University Press, New York.

Entorf, H. , 1997. Random Walks With Drifts: Nonsense Regression and Spurious Fixed-effect Estimation. Journal of Econometrics, Vol. 80, pp. 287-296

Evans, M., 1991. Discovering the Link between Inflation Rates and Inflation Uncertainty. Journal of Money, Credit and Banking, Vol. 23, pp.169-184.

Granger, C. W. J. , 1981. Some Properties of Time Series Data and Their Use in Econometric Model Specification. Journal of Econometrics, Vol.16, pp. 121-130.

Granger, C.W.J. and P. Newbold, 1974. Spurious Regression in Econometrics. Journal of Econometrics, Vol. 2, pp. 111-120

Granger, C. W. J . , N. Hyung and Y. Jeon, 2001. Spurious regressions with stationary series. Applied Economics, Vol. 33, pp. 899- 904

Hamilton, J. D. 1994a. State-Space Models. Handbook of Econometrics, Vol.4, No. 50, pp.3041-3077

Hamilton, J. D. 1994b. Time Series Analysis, Princeton University Press, New Jersey.

Hanohan, Patrick, 1993. Cointegration and Time Varying Parameters: A Comment, European Economic Review, vol. 37, pp. 1279-1281.

K., Tanaka, 1981. An Alternative Approach to the Asymptotic Theory of Spurious Regression, Cointegration, and near Cointegration. Econometric Theory, Vol. 9, No. 1, pp. 36-61

Kim C. and C. R. Nelson, 1989. The Time-Varying-Parameter Model for Modeling Changing Conditional Variance: The Case of the Lucas Hypothesis. Journal of Business & Economic Statistics, Vol. 7, No. 4, pp. 433-440

Kim,T., Y. Lee, P. Newbold, 2003. Spurious regressions with stationary processes around linear trends. Economics Letters, Vol. 83, pp. 257–262

Kitawa, G., 1987. Non-Gaussian State Space Modeling of Non-stationary Time Series. Journal of the American Statistical Association, Vol. 82, No. 400, pp. 1032-1041

Liu, Cristina Y., J. He, 1991. A Variance-Ratio Test of Random Walks in Foreign Exchange Rates. Journal of Finance, Vol. 46, no: 2, pp. 773-785.

Lucas, Robert E., Jr., 1976.Econometric Policy Evaluation: A Critique, Carnegie-Rochester Conference Series on Public Policy, Vol. 1, pp. 19-46.

Maddala, G.S. and I. Kim, 1998. Unit roots, Cointegration and Structural Change. Cambridge university press.UK.

Nelson, C. R. and H. Kang, 1984. Pitfalls in the Use of Time as an Explanatory Variable in Regression. Journal of Business & Economic Statistics, Vol. 2, No. 1, pp. 73-82

Philips, P.C.B, 1986. Understanding the Spurious Regression in Econometrics. Journal of Econometrics, Vol. 33, pp.311-340

Phillips, P. C. B, 1988. Regression Theory for Near-Integrated Time Series. Econometrica, Vol. 56, No. 5 , pp. 1021-1043

Ramajo, J. , 2001. Time-varying parameter error correction models: the demand for money in Venezuela, 1983:I-1994:IV. Applied Economics, Vol. 33, pp. 771 782

Tanizaki H. , 2000. Time-varying parameter model revisited, Kobe University Economic Review, Vol. 45, pp. 41-57

Tsay, W. , C. Chung, 1999. The spurious regression of fractionally integrated processes. Journal of Econometrics, Vol. 96, pp. 155-182

Yule ,G. U. , 1926. Why do we Sometimes get Nonsense-Correlations between Time-Series?: A Study in and the Nature of Time-Series. Journal of the Royal Statistical Society, Vol. 89, No. 1, pp. 1-63

WanChun, F. and B. Lun, 2009. Time-varying parameter auto-regressive models for autocovariance nonstationary time series. Science in China Series A: Mathematics, Vol. 52, No. 3, pp. 577–584

APPENDIX A: FIGURES

Figure 1. Identification of the Problem: Noncointegrated series without Drift T=1(Graph of actual and predicted y) N=100

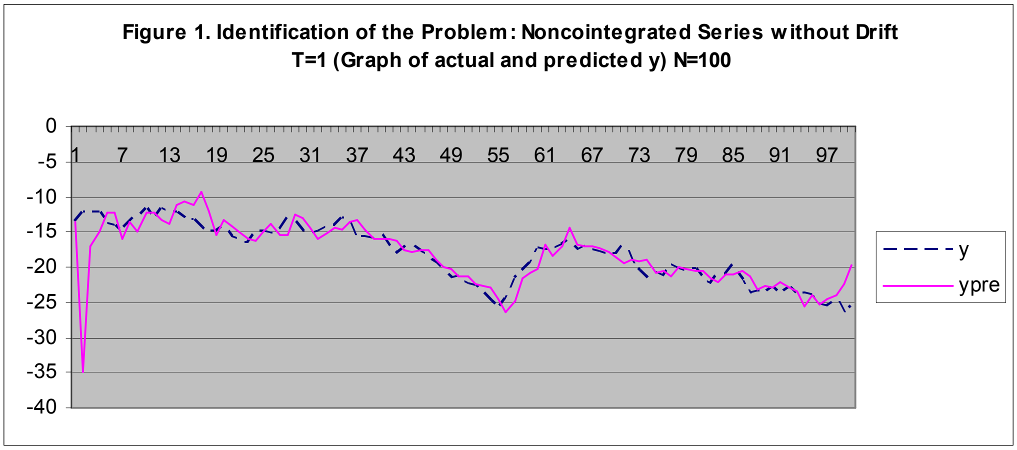


Figure 2. Identification of the Problem: Noncointegrated series without drift T=1 (Graph of z(t), at_1 and at) N=100

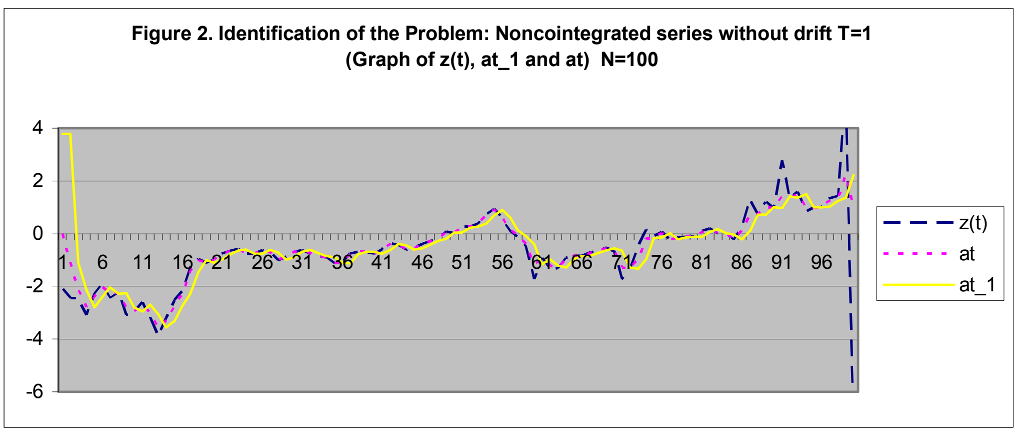


Figure 3. Identification of the Problem: Noncointegrated Series without Drift T=0.7 (Graph of actual and predicted y) N=100

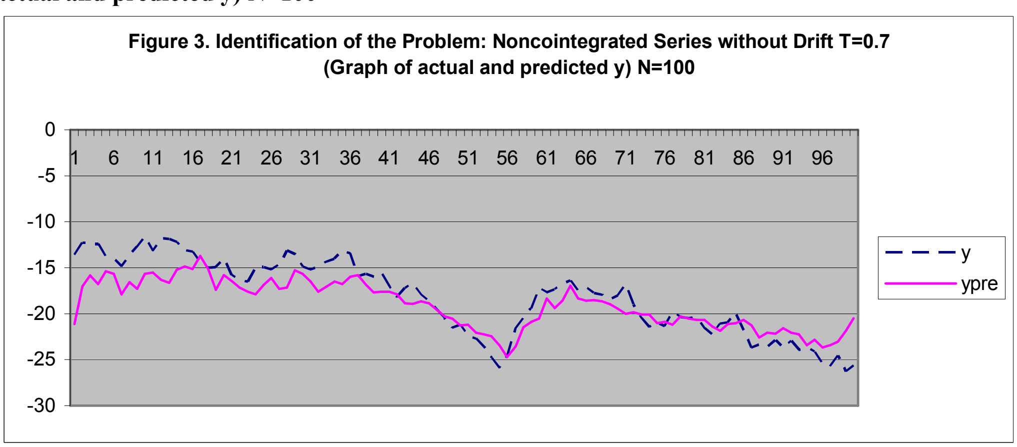


Figure 4. Identification of the Problem: Noncointegrated series without drift T=0.7 (Graph of z(t), at_1 and at) N=10

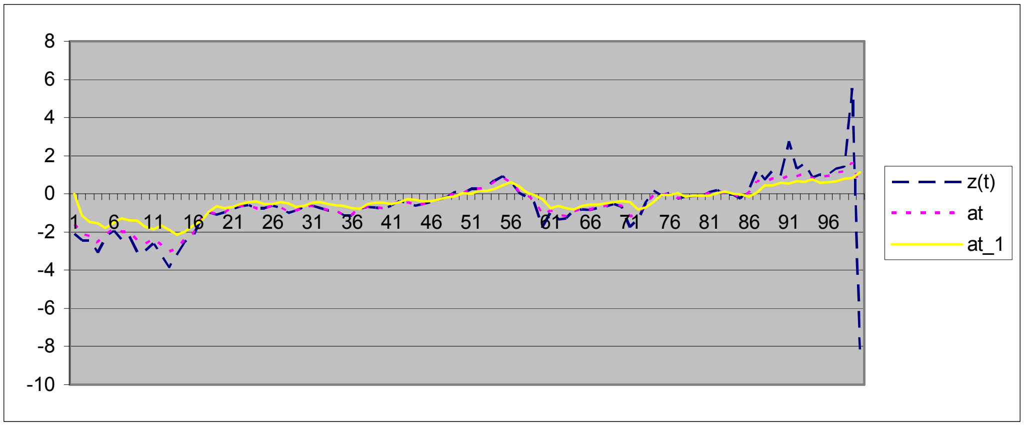


Figure 5. Identification of the Problem: Noncointegrated Series without Drift T=0.3 (Graph of actual and predicted y) N=100

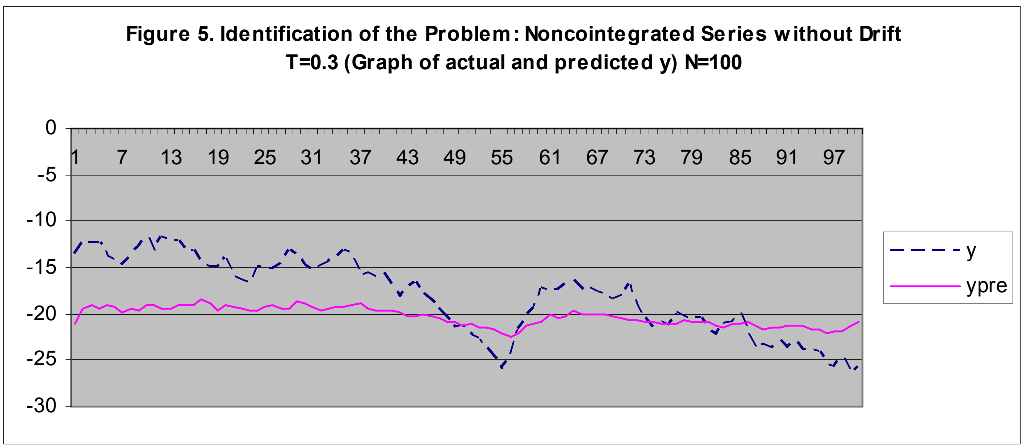


Figure 6. Identification of the Problem: Noncointegrated series without drift T=0.3 (Graph of z(t), at_1 and at) N=100

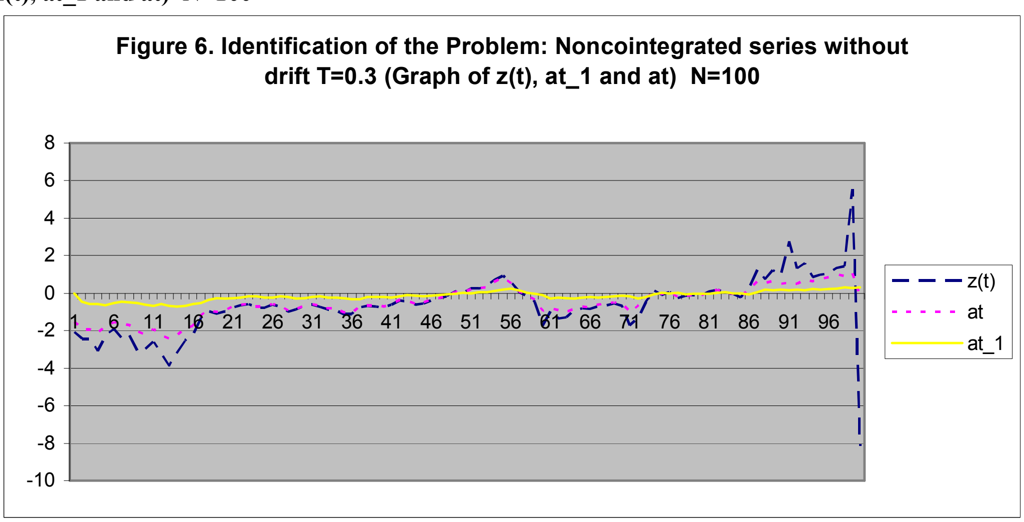


Figure 7. Identification of the Problem: Noncointegrated Series with Drift T=1 (Graph of actual and predicted y) N=100

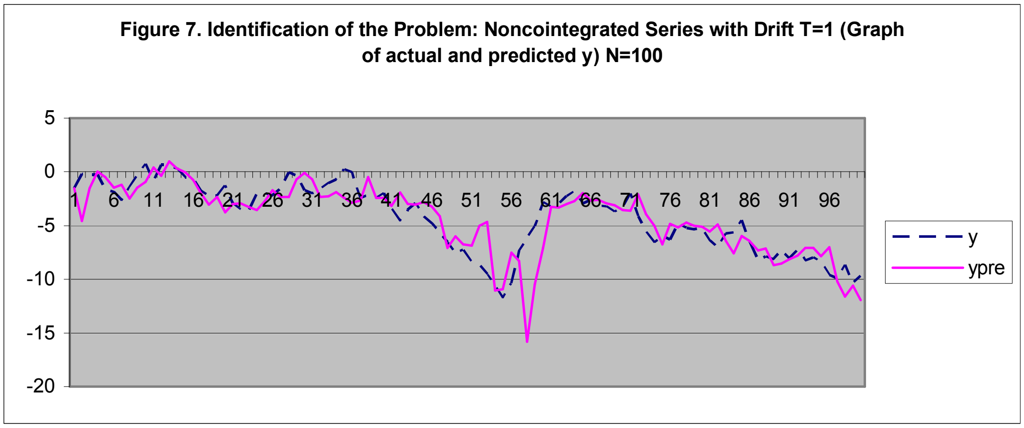


Figure 8. Identification of the Problem: Noncointegrated Series with drift T=1 (Graph of z(t), at_1 and at) N=100

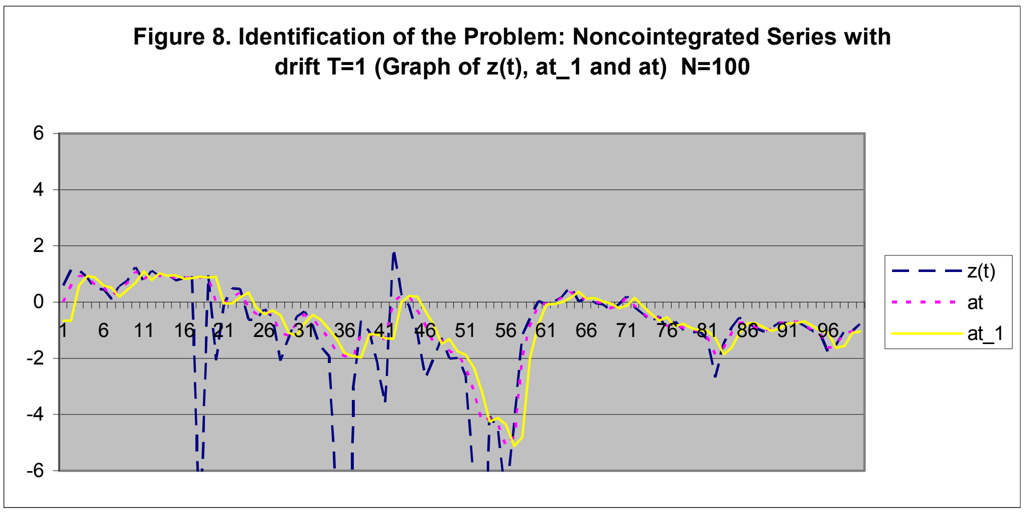


Figure 9. Identification of the Problem: Noncointegrated Series with Drift T=0.7 (Graph of actual and predicted y) N=100

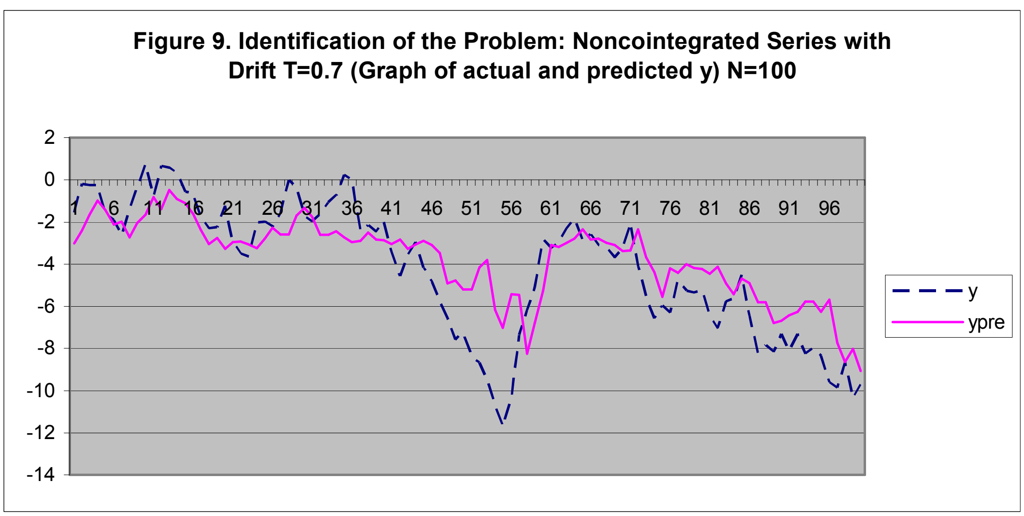


Figure 10. Identification of the Problem: Noncointegrated Series with drift T=0.7 (Graph of z(t), at_1 and at) N=100

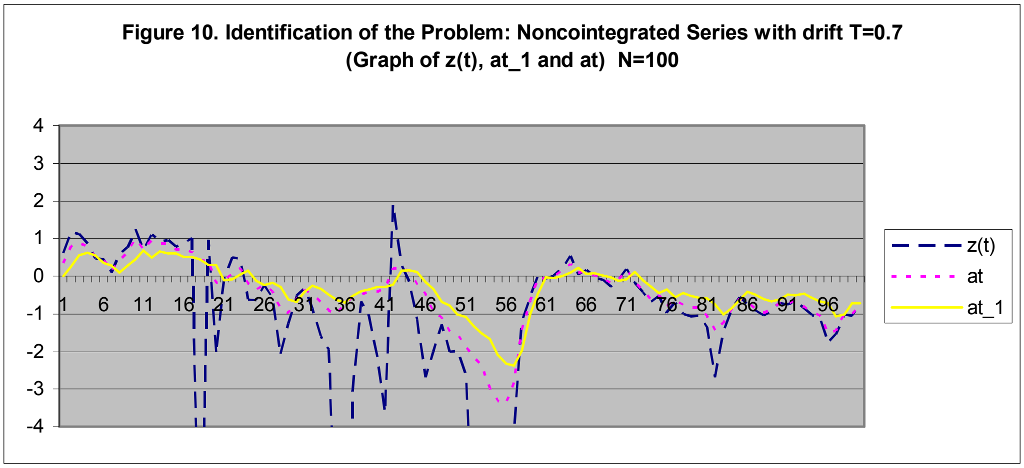


Figure 11. Identification of the Problem: Noncointegrated Series with Drift T=0.3 (Graph of actual and predicted y) N=100

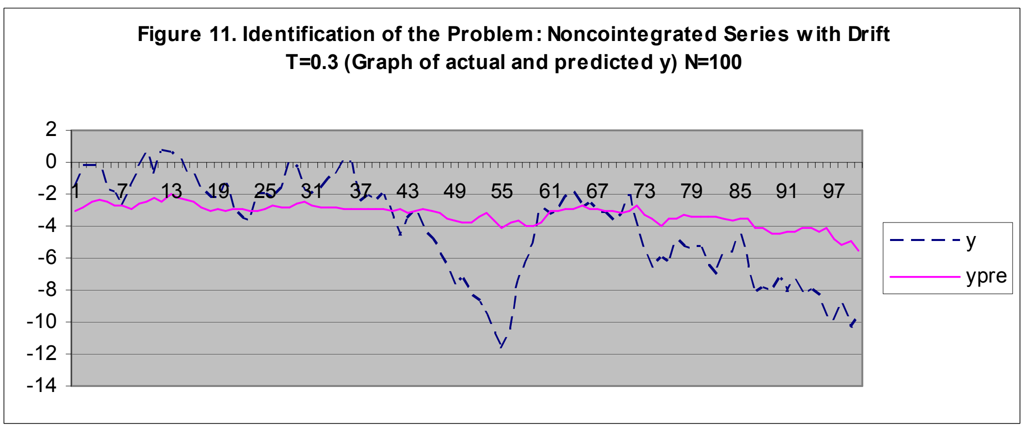


Figure 12. Identification of the Problem: Noncointegrated Series with drift T=0.3 (Graph of z(t), at_1 and at) N=100

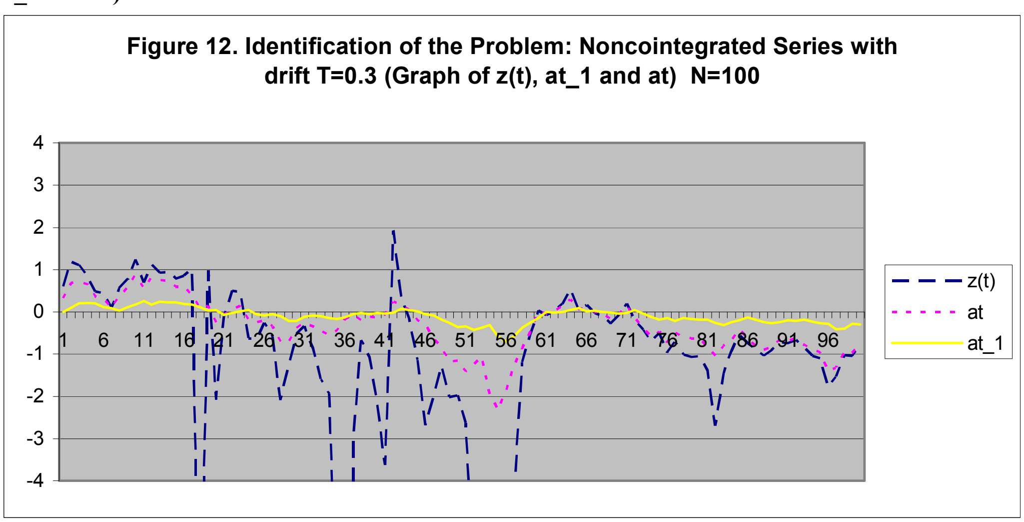


Figure 13. Identification of the Problem: Linearly Related Series with Drift T=1 (Graph of actual and predicted y) N=100

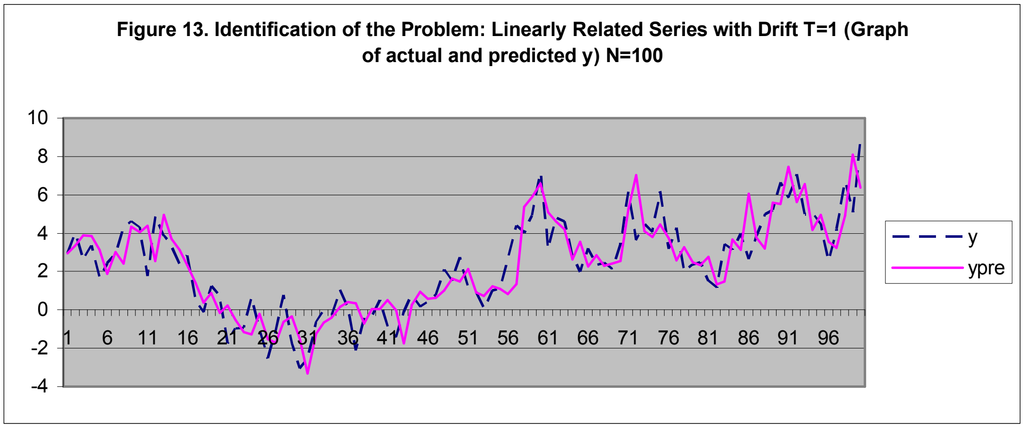


Figure 14. Identification of the Problem: Linearly Related Series with drift T=1 (Graph of z(t), at_1 and at) N=100

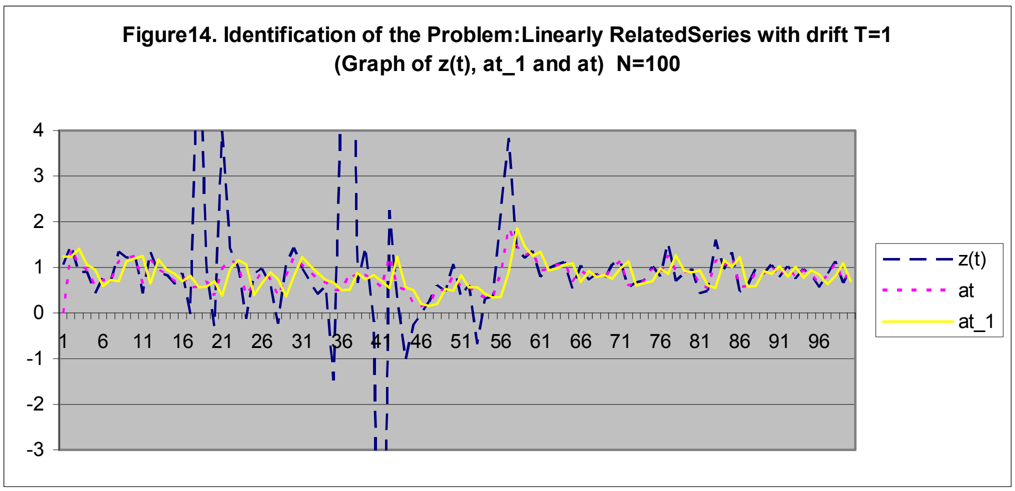


Figure 15. Identification of the Problem: Linearly Related Series with Drift T=0.7 (Graph of actual and predicted y) N=100

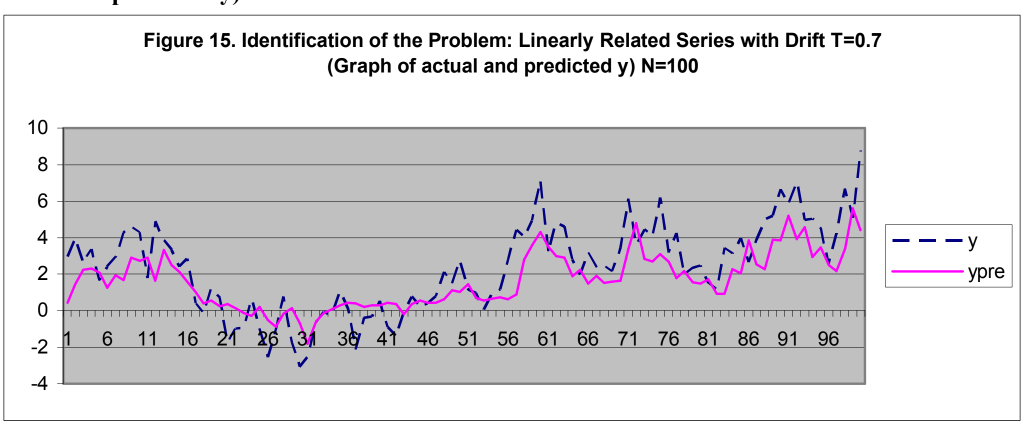


Figure 16. Identification of the Problem: Linearly Related Series with drift T=0.7 (Graph of z(t), at_1 and at)

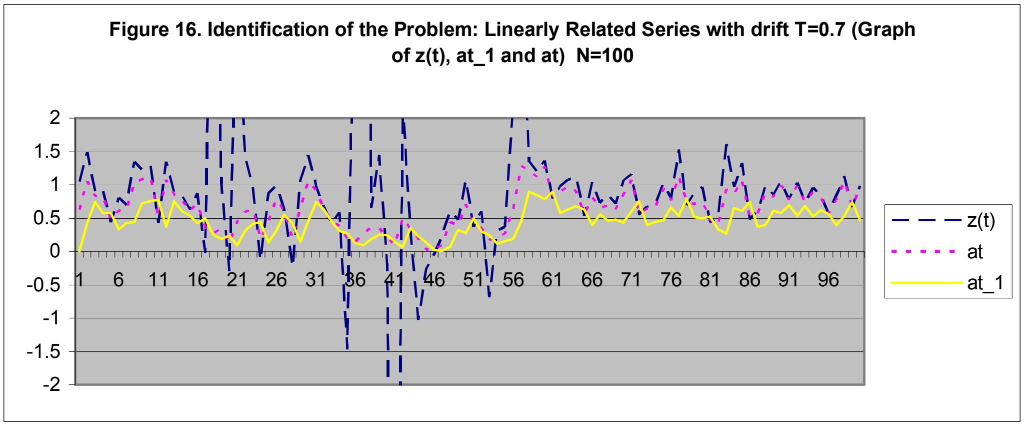


Figure 17. Identification of the Problem: Linearly Related Series with Drift T=0.3 (Graph of actual and predicted y) N=100

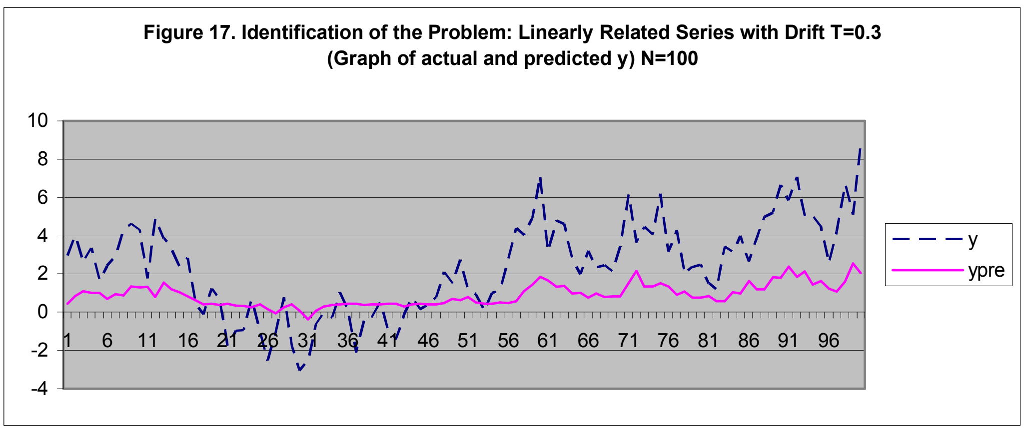


Figure 18. Identification of the Problem: Linearly Related Series with drift T=0.3 (Graph of z(t), at_1 and at) N=100

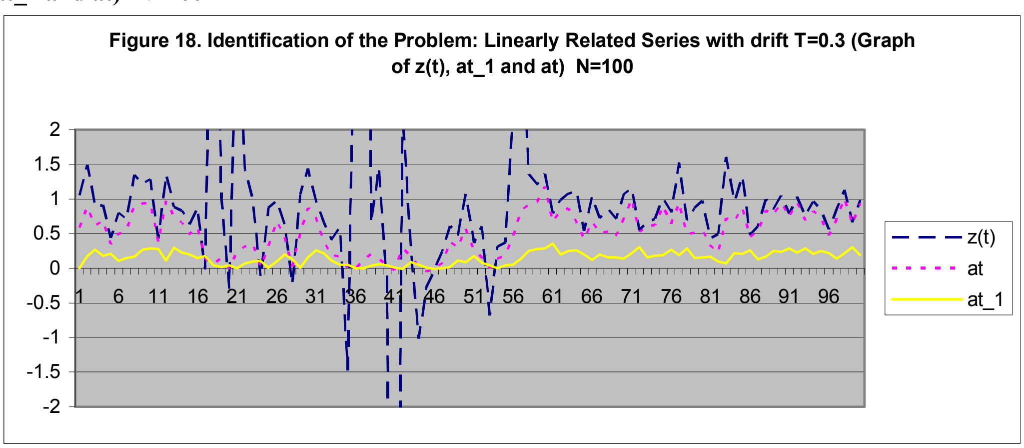


Figure 19. Identification of the Problem: Noncointegrated Series without Drift MLE Estimation (Graph of actual and predicted y) N=100

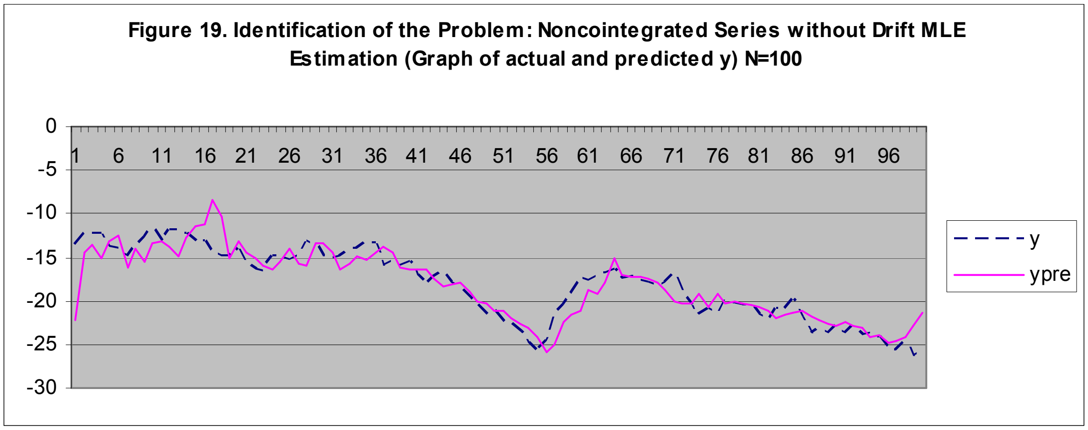


Figure 20. Identification of the Problem: Noncointegrated Series without drift MLE estimation (Graph of z(t), at_1 and at) N=100

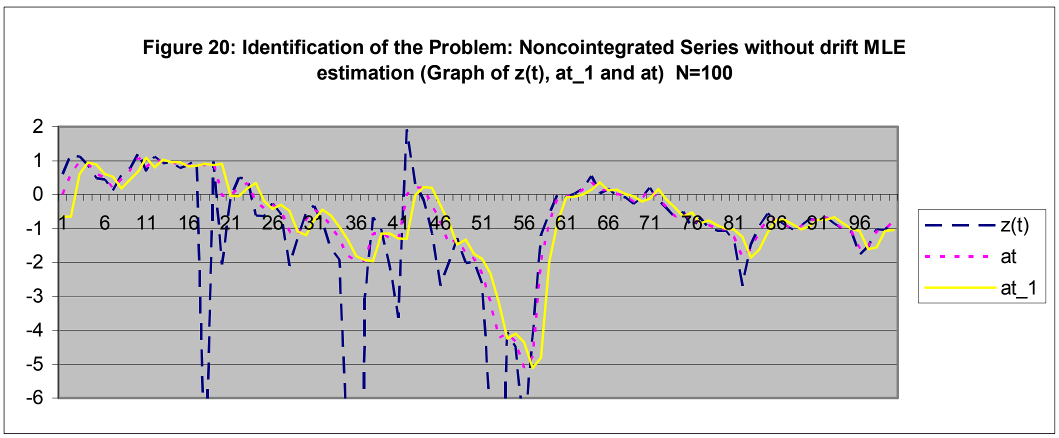


Figure 21. Identification of the Problem: Noncointegrated Series with Drift MLE Estimation (Graph of actual and predicted y) N=100

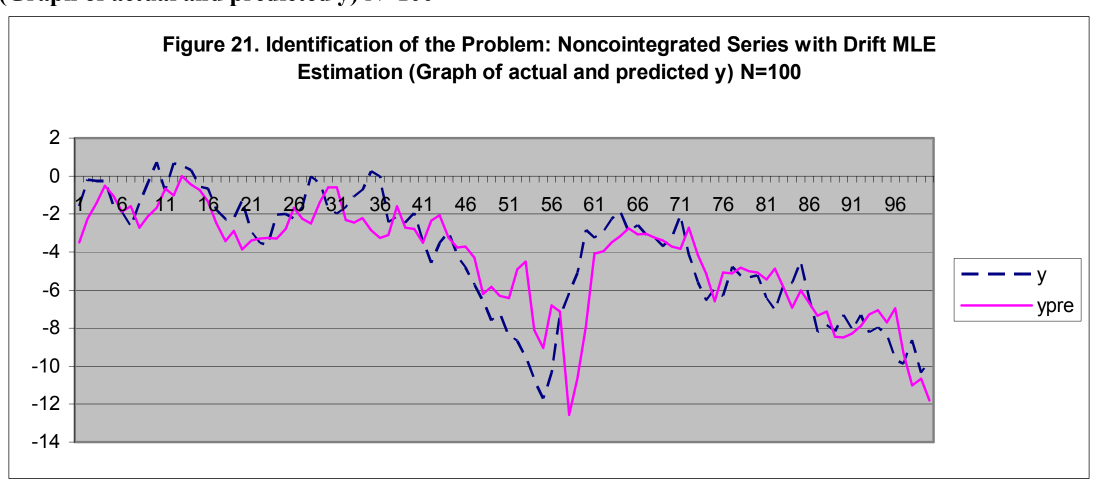


Figure 22. Identification of the Problem: Noncointegrated Series with Drift MLE Estimation (Graph of z(t), at_1 and at) N=100

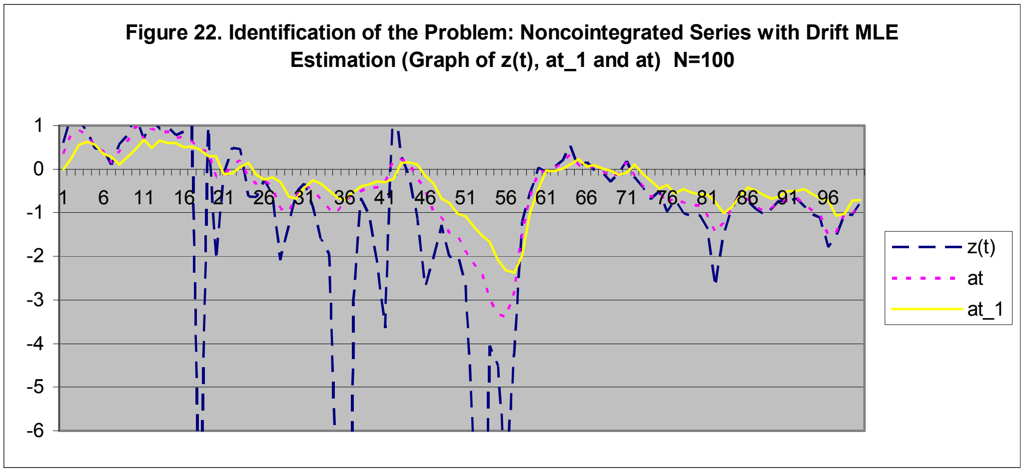


Figure 23. Identification of the Problem: Linearly Related Series MLE Estimation (Graph of actual and predicted y) N=100

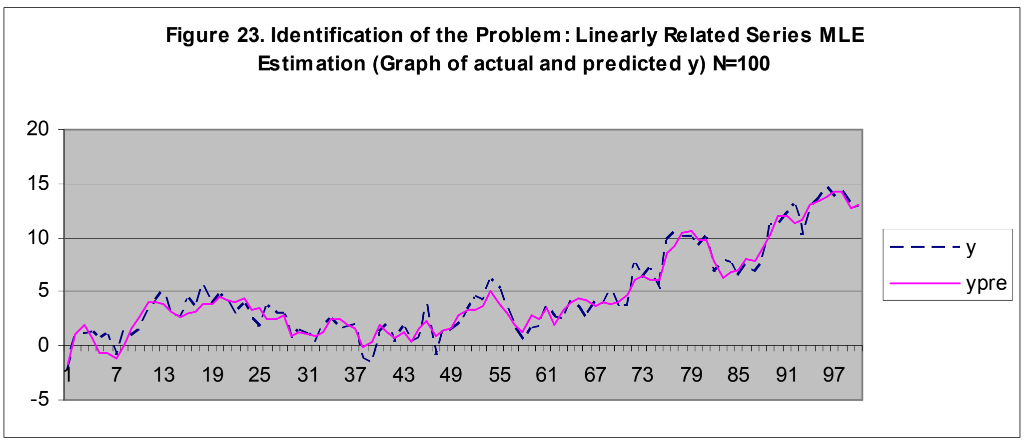


Figure 24. Identification of the Problem: Linearly Related Series MLE Estimation (Graph of z(t), at_1 and at) N=100

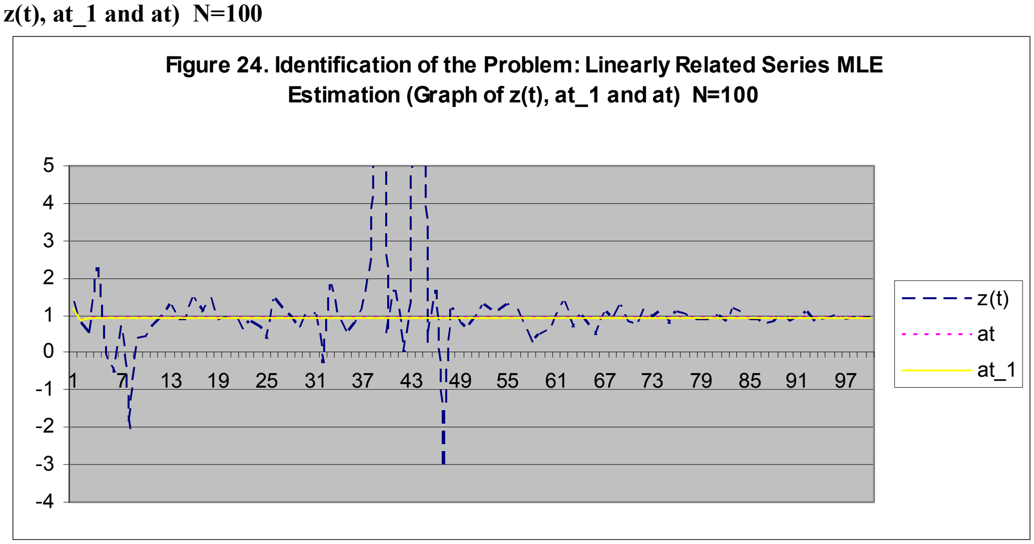


Figure 25. Identification of the Problem: TVP with κ=0.7 and var (ξ)=0.2 MLE Estimation (Graph of y and ypre) Sample=100

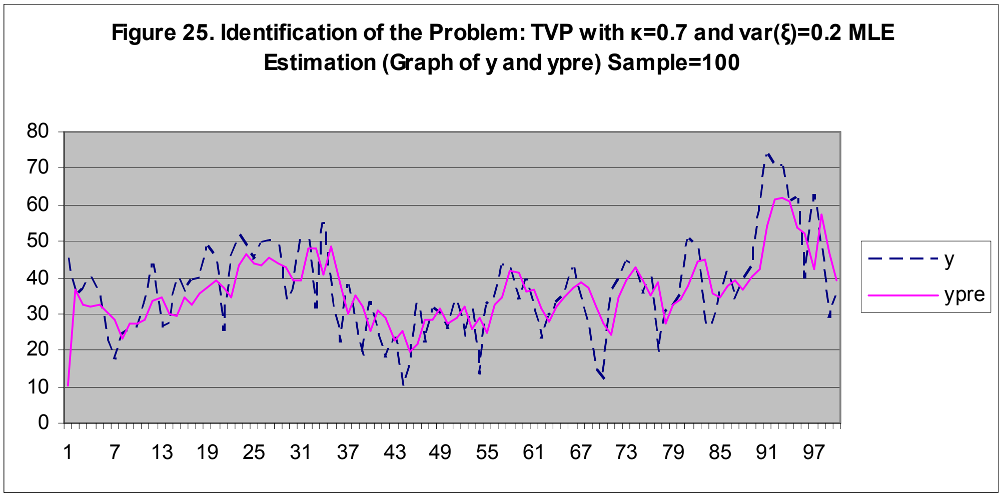


Figure 26. Identification of the Problem: TVP with κ=0.7 and var(ξ)=0.2 MLE Estimation (Graph of z(t), at_1 and β(t)) N=100

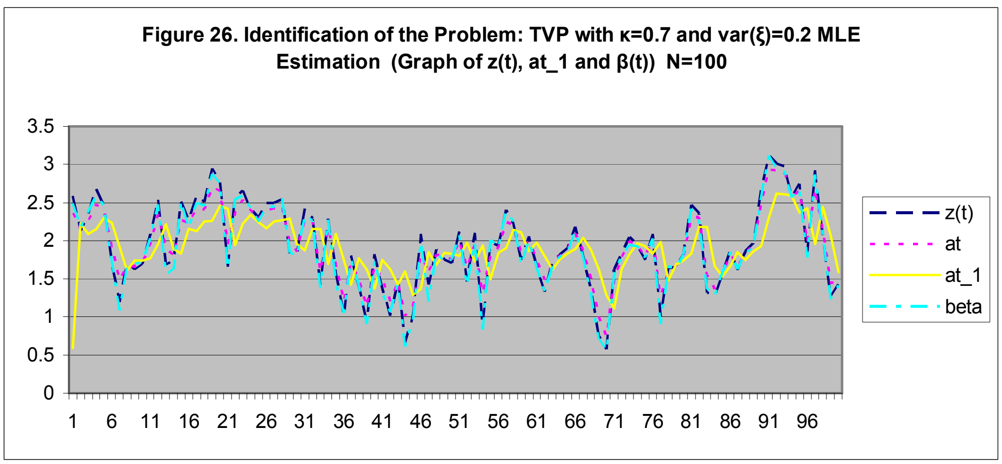


Figure 27. Identification of the Problem: TVP with κ=0.7 and var(ξ)=0.02 MLE Estimation (Graph of y and ypre) N=100

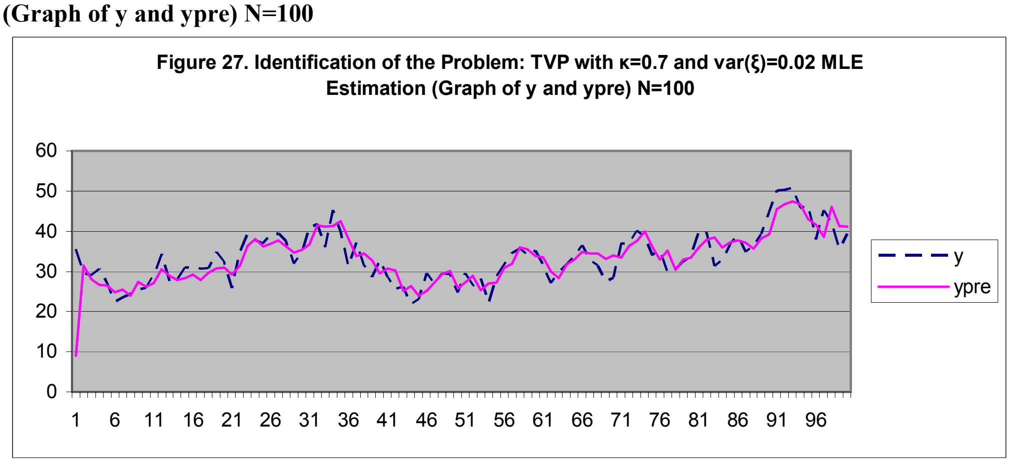


Figure 28. Identification of the Problem: TVP with κ=0.7 and var(ξ)=0.02 MLE Estimation (Graph of z(t), at_1 and β(t)) N=100

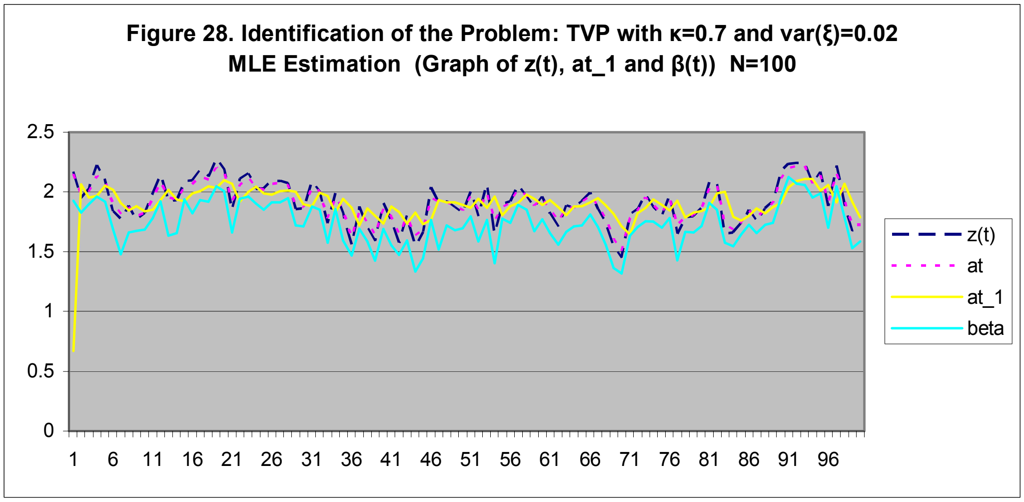


Figure 29. Identification of the Problem: TVP with κ=0.3 and var(ξ)=0.2 MLE Estimation (Graph of y and ypre) N=100

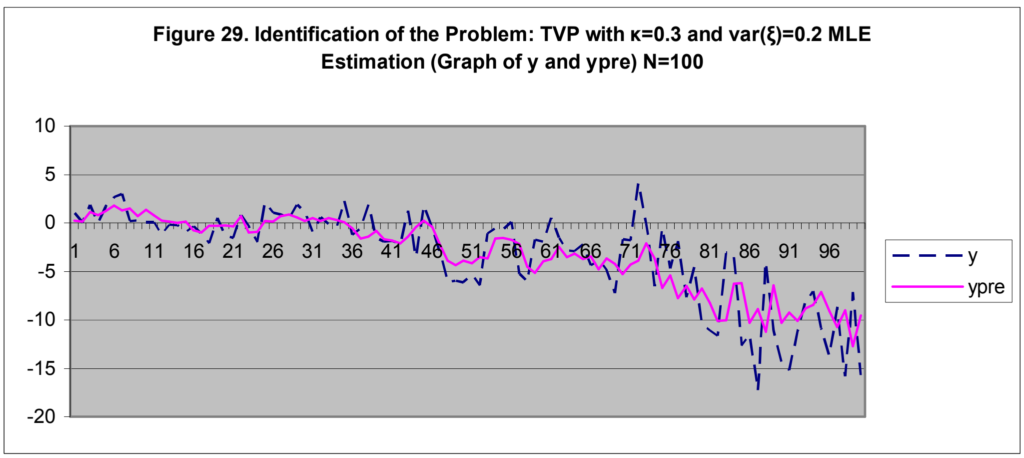


Figure 30. Identification of the Problem: TVP with κ=0.3 and var(ξ)=0.2 MLE Estimation (Graph of z(t), at_1 and β(t)) N=100

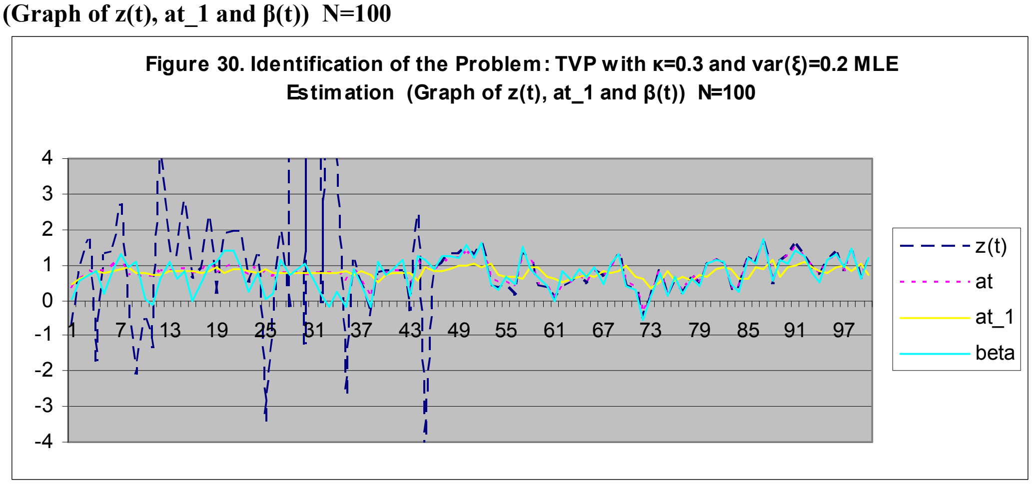


Figure 31. Identification of the Problem: TVP with κ=0.3 and var(ξ)=0.02 MLE Estimation (Graph of y and ypre) N=100

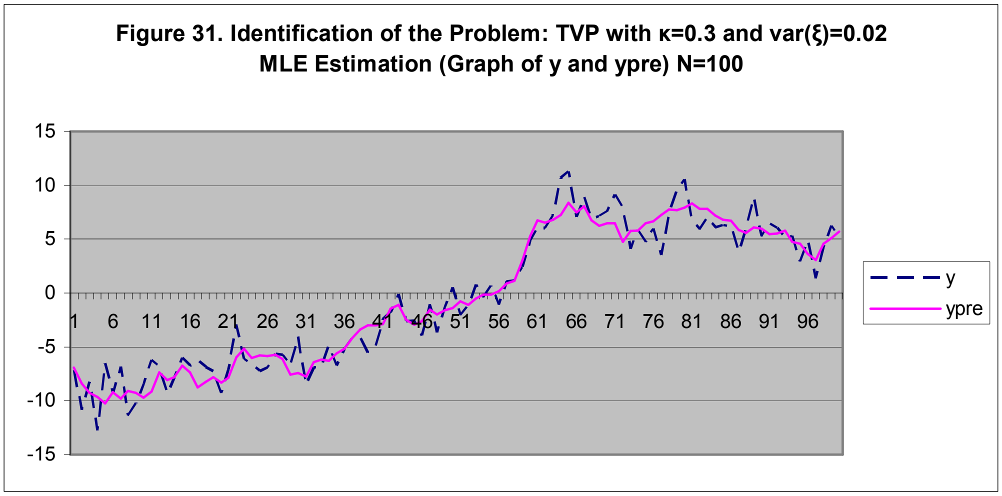


Figure 32. Identification of the Problem: TVP with κ=0.3 and var(ξ)=0.02 MLE Estimation (Graph of z(t), at_1 and β(t)) N=100

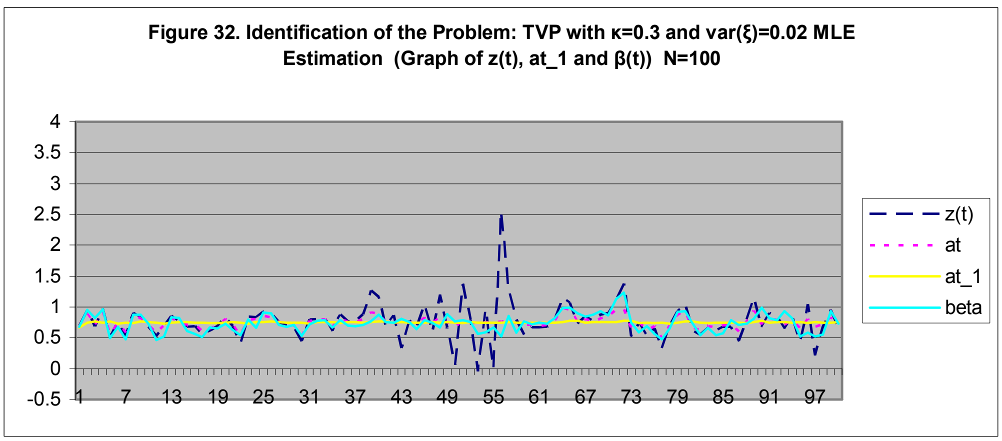


Figure 33. Identification of the Problem: TVP Smooth Transition with γ=0.5 (Graph of actual and predicted y) N=100

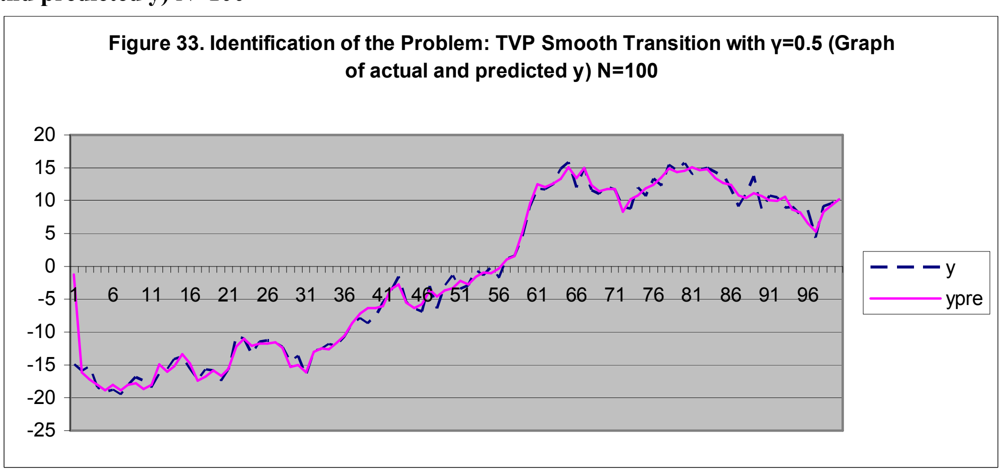


Figure 34. Identification of the Problem: TVP Smooth Transition with γ=0.5 (Graph of z(t), at_1 and β(t)) N=100

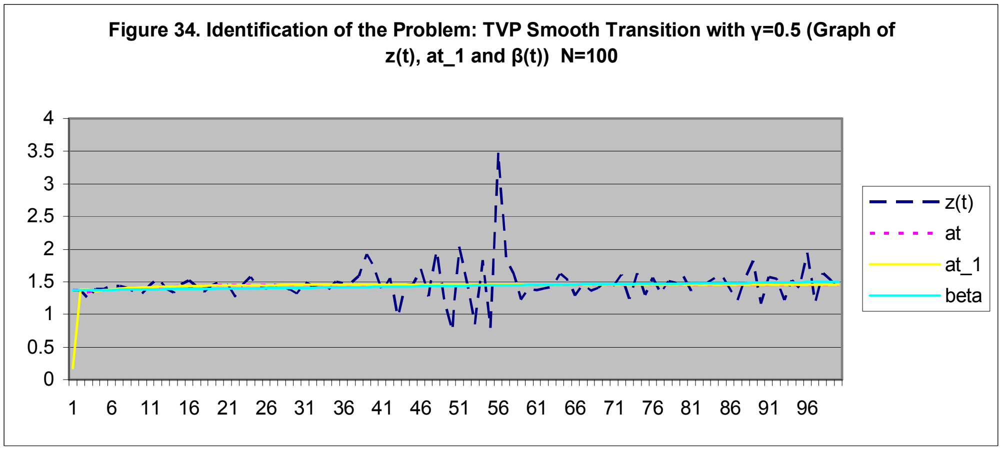


Figure 35. Identification of the Problem: TVP Smooth Transition with γ=1.5 (Graph of actual and predicted y) N=100

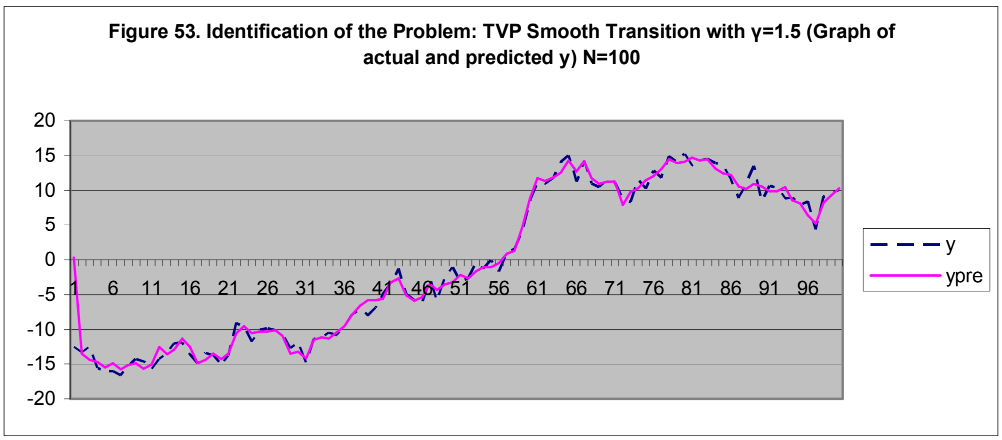


Figure 36. Identification of the Problem: TVP Smooth Transition with γ=1.5 (Graph of z(t), at_1 and β(t)) N=100

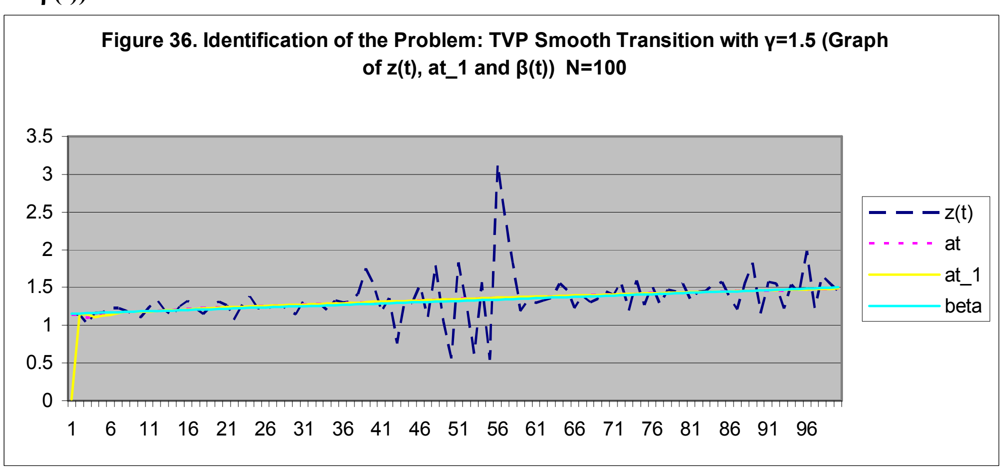


Figure 37. Noncointegrated Series without Drift MLE Estimation Penalty Applied (Graph of actual and predicted y) N=100

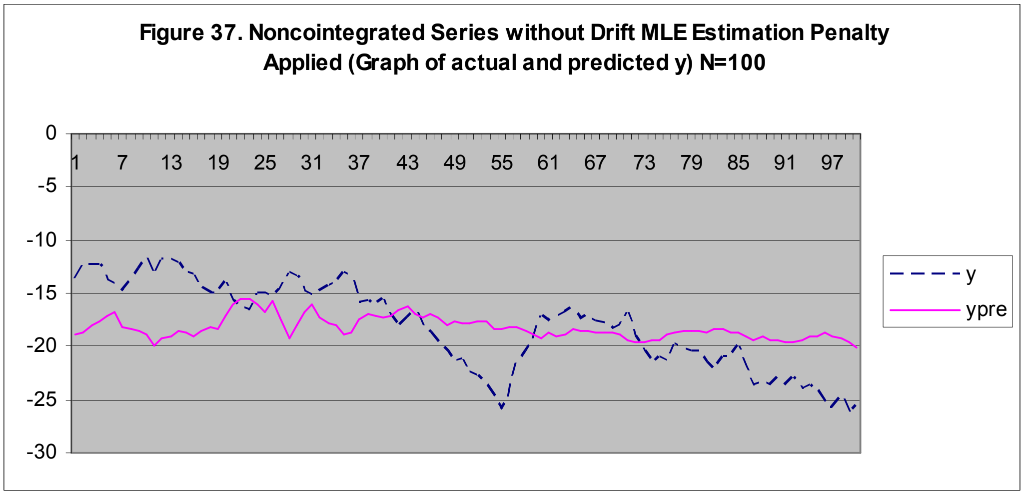


Figure 38. Noncointegrated Series without Drift MLE Estimation Penalty Applied (Graph of z(t), at_1 and at) N=100

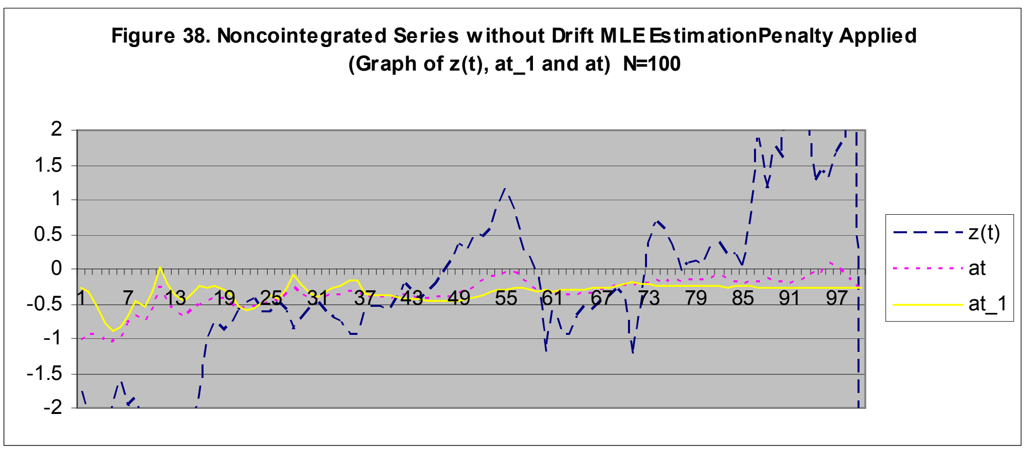


Figure 39. Linearly Related Series MLE Estimation Penalty Applied (Graph of actual and predicted y) N=100

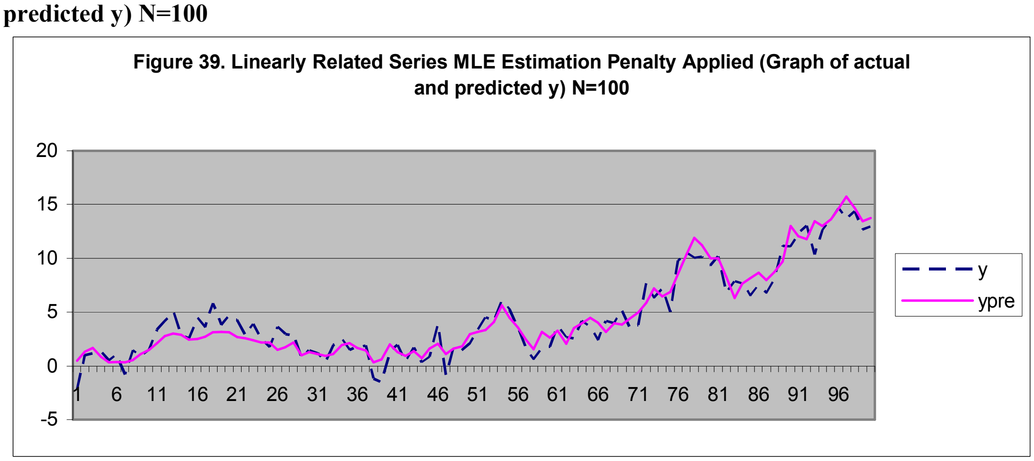


Figure 40. Linearly Related Series MLE Estimation Penalty Applied (Graph of z(t), at_1 and at) N=100

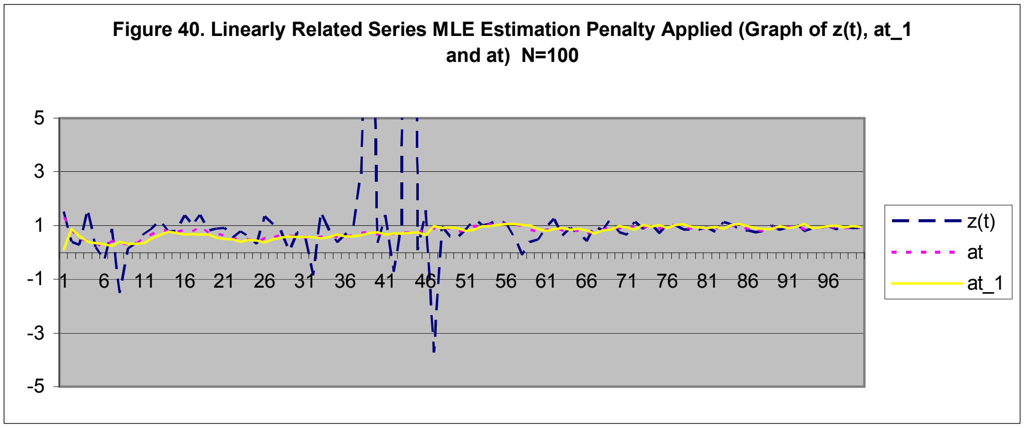


Figure 41. TVP with κ=0.7 and var(ξ)=0.2 MLE Estimation Penalty Applied (Graph of y and ypre) N=100

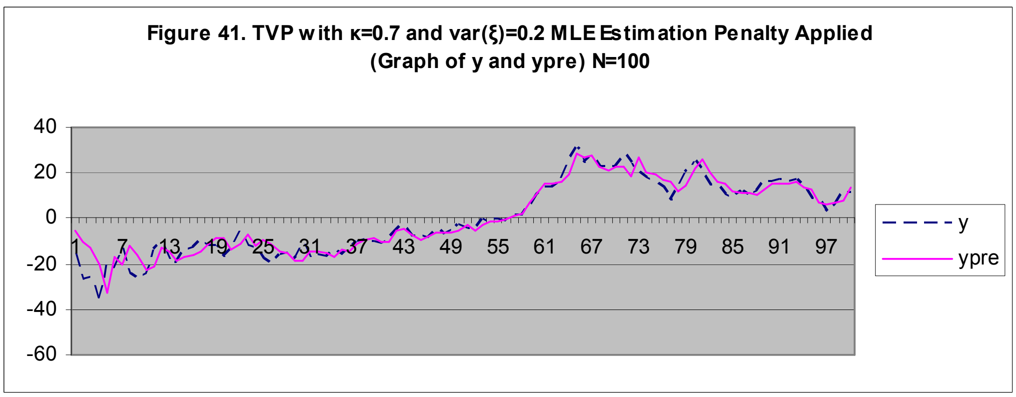


Figure 42. TVP with κ=0.7 and var(ξ)=0.2 MLE Estimation Penalty Applied (Graph of z(t), at_1 and β(t)) N=100

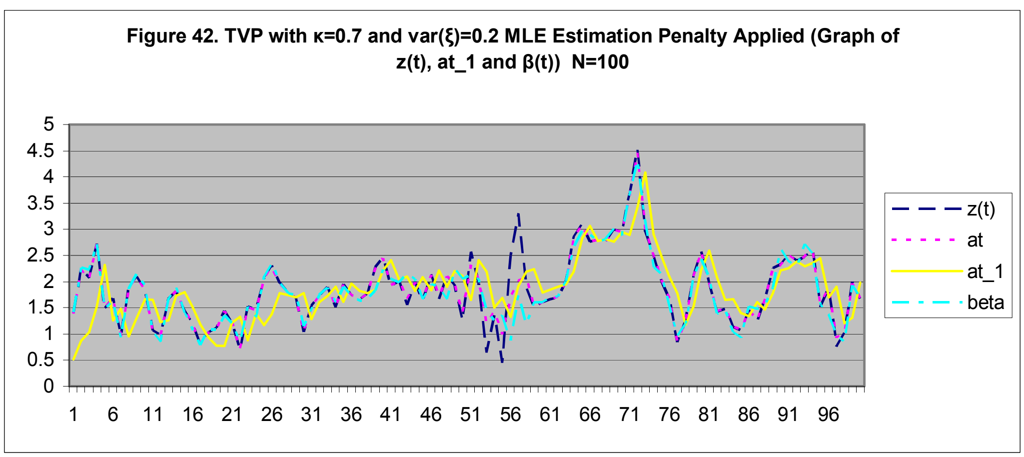


Figure 43. TVP with κ=0.7 and var(ξ)=0.02 MLE Estimation Penalty Applied (Graph of y and ypre) N=100

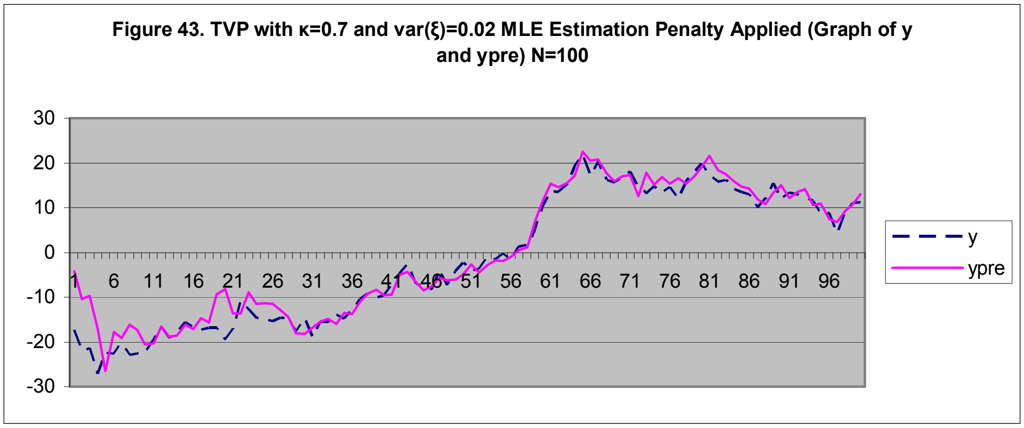


Figure 44. TVP with κ=0.7 and var(ξ)=0.02 MLE Estimation Penalty Applied (Graph of z(t), at_1 and β(t)) N=100

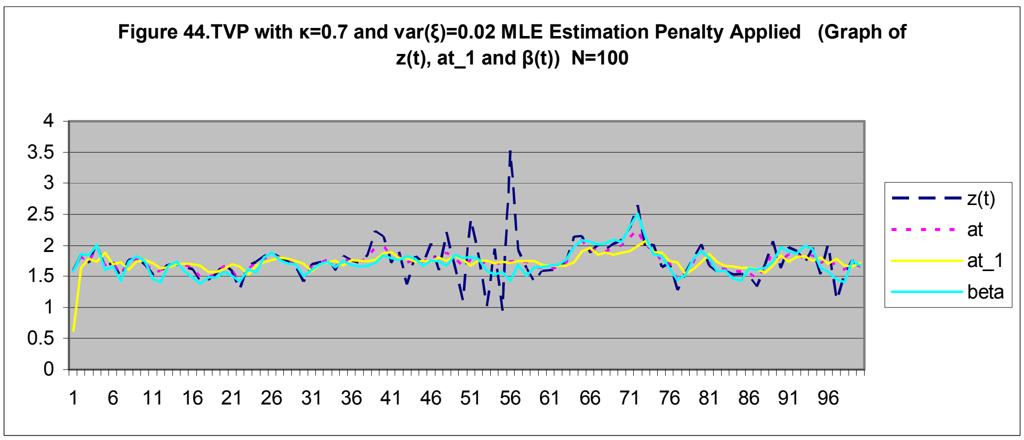


Figure 45. TVP with κ=0.3 and var(ξ)=0.2 MLE Estimation Penalty Applied (Graph of y and ypre) N=100

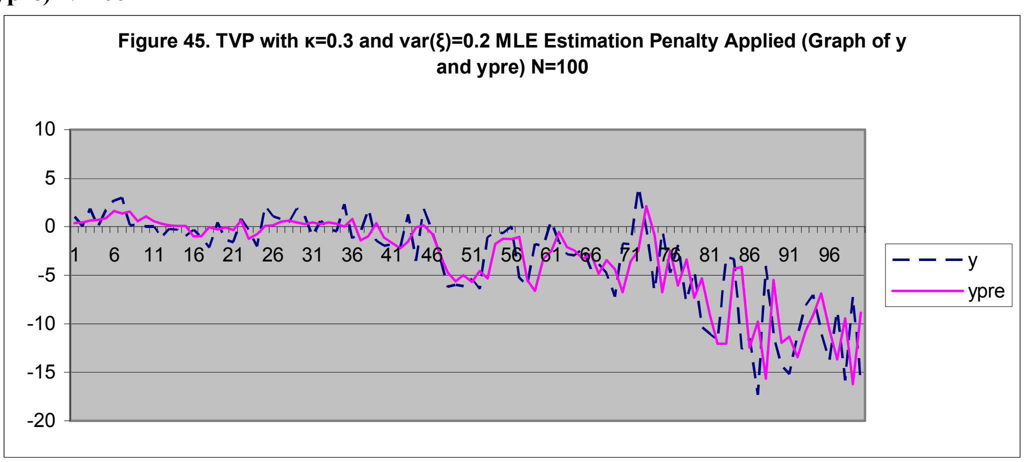


Figure 46. TVP with κ=0.3 and var(ξ)=0.2 MLE Estimation Penalty Applied (Graph of z(t), at_1 and β(t)) N=100

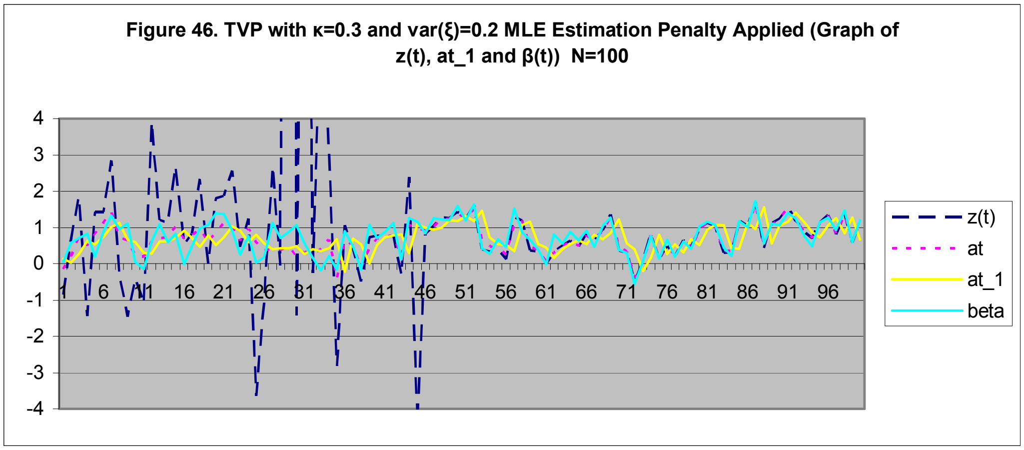


Figure 47. TVP with κ=0.3 and var(ξ)=0.02 MLE Estimation Penalty Applied (Graph of y and ypre) N=100

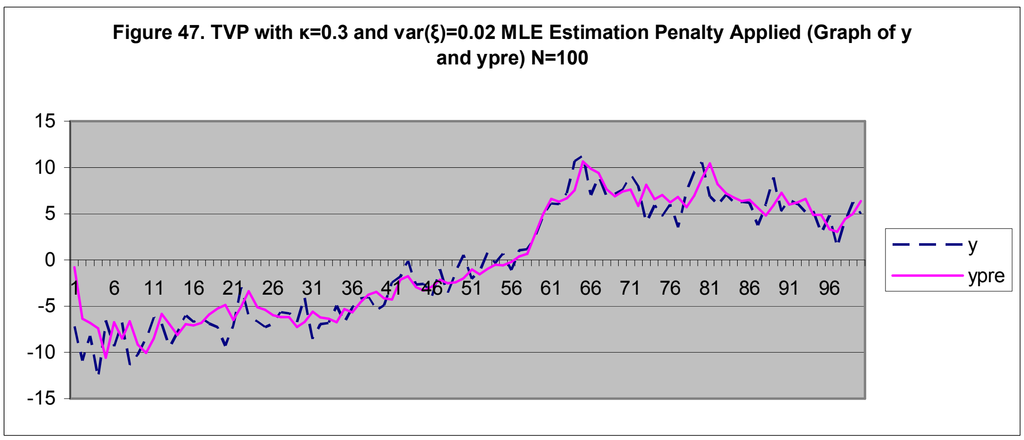


Figure 48. TVP with κ=0.3 and var(ξ)=0.02 MLE Estimation Penalty Applied (Graph of z(t), at_1 and β(t)) N=100

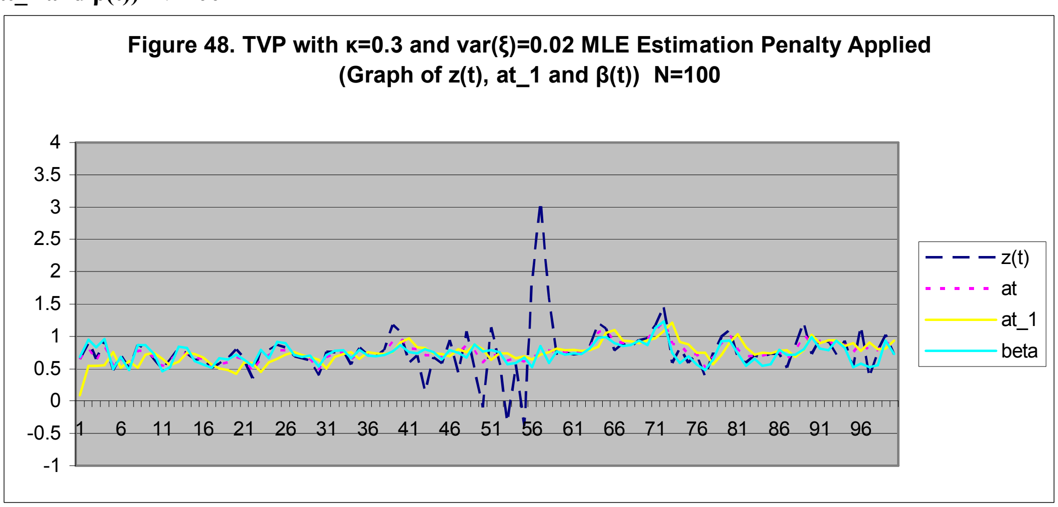


Figure 49. TVP Smooth Transition with γ=0.5 MLE Penalty Applied (Graph of actual and predicted y) N=100

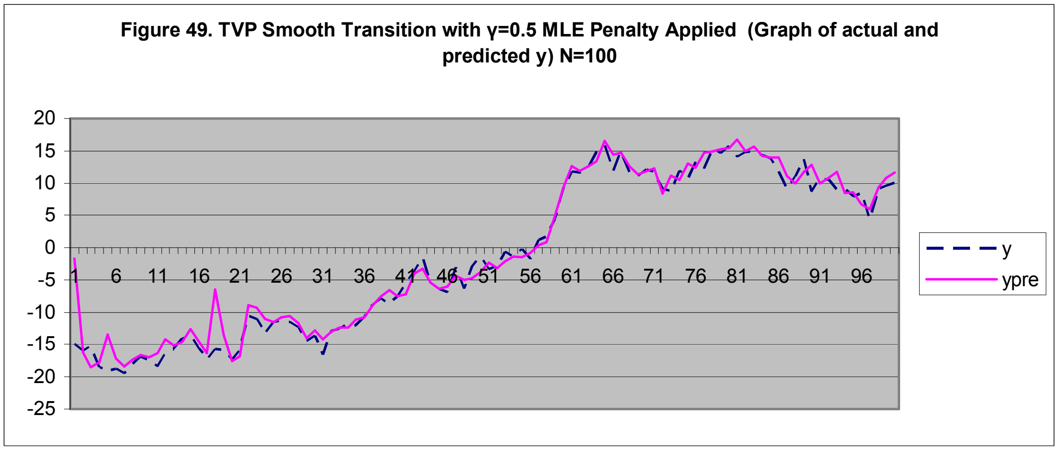


Figure 50. TVP Smooth Transition with γ=0.5 MLE Penalty Applied (Graph of z(t), at_1 and β(t)) N=100

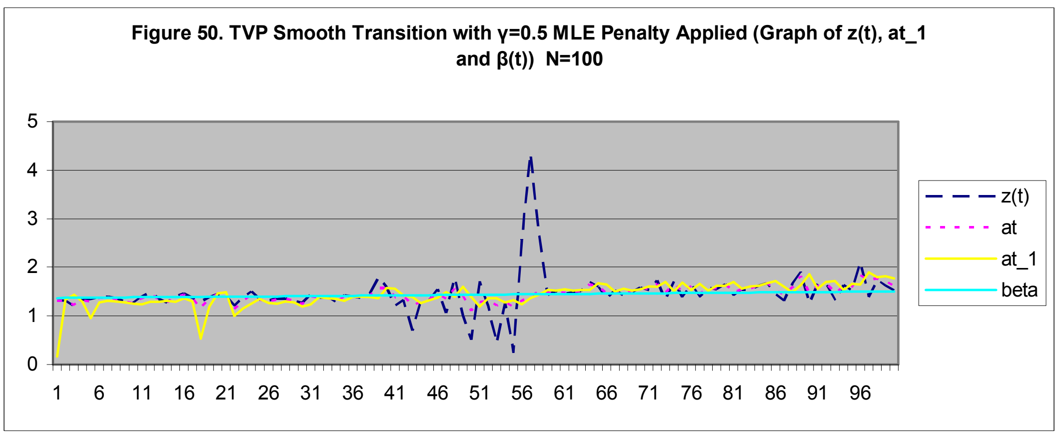


Figure 51. TVP Smooth Transition with γ=1.5 MLE Penalty Applied (Graph of actual and predicted y) N=100

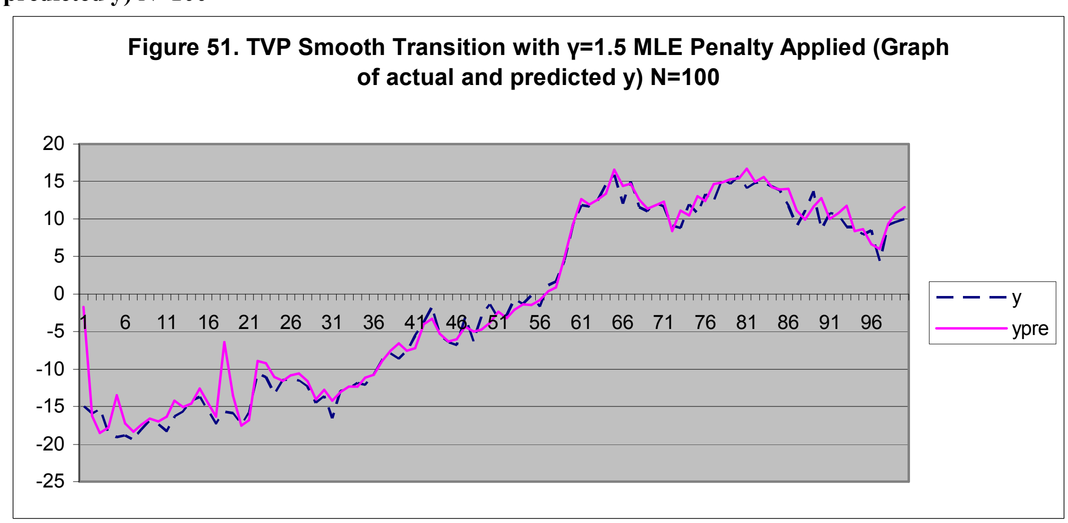
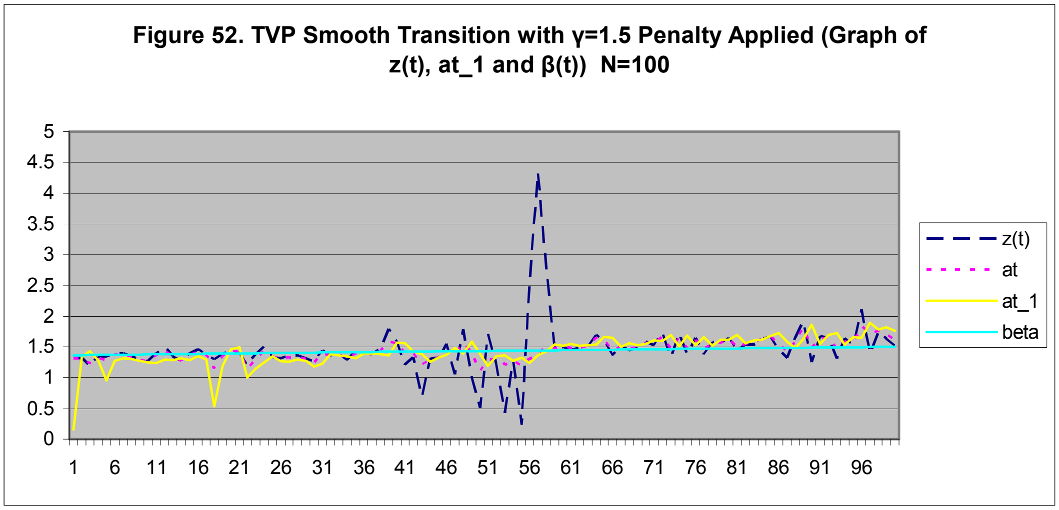


Figure 52. TVP Smooth Transition with γ=1.5 Penalty Applied (Graph of z(t), at_1 and β(t)) N=100



APPENDIX B: TABLES

Table 20. Histograms of MLE Estimator of Parameters in Linearly Related Random Walk without Drift Sample Size=400

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	0	-1.2	0	-1.2	0	-1.2	0
-1.1	0	-1.1	0	-1.1	0	-1.1	0
-1	0	-1	0	-1	0	-1	0
-0.9	0	-0.9	0	-0.9	0	-0.9	0
-0.8	0	-0.8	5	-0.8	0	-0.8	0
-0.7	0	-0.7	7	-0.7	0	-0.7	0
-0.6	1	-0.6	13	-0.6	0	-0.6	0
-0.5	0	-0.5	16	-0.5	0	-0.5	0
-0.4	1	-0.4	23	-0.4	0	-0.4	0
-0.3	2	-0.3	30	-0.3	0	-0.3	0
-0.2	5	-0.2	27	-0.2	0	-0.2	0
-0.1	8	-0.1	107	-0.1	0	-0.1	0
0	10	0	1123	0	0	0	0
0.1	26	0.1	1162	0.1	3	0.1	2999
0.2	55	0.2	199	0.2	15	0.2	1
0.3	110	0.3	85	0.3	26	0.3	0
0.4	329	0.4	54	0.4	42	0.4	0
0.5	975	0.5	43	0.5	44	0.5	0
0.6	946	0.6	35	0.6	56	0.6	0
0.7	313	0.7	35	0.7	80	0.7	0
0.8	117	0.8	23	0.8	196	0.8	0
0.9	59	0.9	11	0.9	1162	0.9	0
1	22	1	2	1	1170	1	0
1.1	12	1.1	0	1.1	91	1.1	0
1.2	4	1.2	0	1.2	35	1.2	0
More	5	More	0	More	80	More	0
Average	0.4978		0.0304		0.873		0.007891
Variance	0.0257		0.0341		0.0278		0.000126
# of times the null that ML estimate is insignificant is rejected							
501		2880		1896		2578	

Table 21. Histograms of MLE Estimator of Parameters in Linearly Related Random Walk without Drift Sample Size=100

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	35	-1.2	0	-1.2	0	-1.2	0
-1.1	9	-1.1	0	-1.1	0	-1.1	0
-1	11	-1	0	-1	0	-1	0
-0.9	6	-0.9	2	-0.9	0	-0.9	0
-0.8	14	-0.8	5	-0.8	0	-0.8	0
-0.7	12	-0.7	6	-0.7	0	-0.7	0
-0.6	11	-0.6	14	-0.6	0	-0.6	0
-0.5	27	-0.5	18	-0.5	0	-0.5	0
-0.4	28	-0.4	28	-0.4	0	-0.4	0
-0.3	24	-0.3	34	-0.3	0	-0.3	0
-0.2	37	-0.2	44	-0.2	0	-0.2	0
-0.1	55	-0.1	113	-0.1	0	-0.1	0
0	93	0	1111	0	0	0	0
0.1	81	0.1	1186	0.1	4	0.1	2946
0.2	149	0.2	182	0.2	10	0.2	50
0.3	193	0.3	83	0.3	28	0.3	4
0.4	298	0.4	45	0.4	22	0.4	0
0.5	422	0.5	44	0.5	45	0.5	0
0.6	410	0.6	31	0.6	55	0.6	0
0.7	285	0.7	22	0.7	89	0.7	0
0.8	187	0.8	22	0.8	204	0.8	0
0.9	132	0.9	8	0.9	1190	0.9	0
1	121	1	2	1	1074	1	0
1.1	66	1.1	0	1.1	140	1.1	0
1.2	64	1.2	0	1.2	43	1.2	0
More	230	More	0	More	96	More	0
Average	0.5012		0.0206		0.8807		0.018422
Variance	0.3353		0.0332		0.0277		0.000728
# of times the null that ML estimate is insignificant is rejected							
1814		2877		1784		2557	

Table 22. Histograms of MLE Estimator of Parameters: TVP with T=0.7 and var=0.2 Sample Size=100

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	740	-1.2	0	-1.2	1	-1.2	0
-1.1	14	-1.1	0	-1.1	0	-1.1	0
-1	13	-1	0	-1	0	-1	0
-0.9	9	-0.9	0	-0.9	0	-0.9	0
-0.8	15	-0.8	0	-0.8	0	-0.8	0
-0.7	9	-0.7	0	-0.7	0	-0.7	0
-0.6	15	-0.6	0	-0.6	0	-0.6	0
-0.5	15	-0.5	0	-0.5	0	-0.5	0
-0.4	20	-0.4	0	-0.4	0	-0.4	0
-0.3	23	-0.3	0	-0.3	2	-0.3	0
-0.2	28	-0.2	0	-0.2	6	-0.2	0
-0.1	40	-0.1	0	-0.1	23	-0.1	0
0	29	0	2	0	33	0	0
0.1	45	0.1	0	0.1	93	0.1	0
0.2	59	0.2	1	0.2	147	0.2	0
0.3	116	0.3	6	0.3	237	0.3	28
0.4	139	0.4	36	0.4	434	0.4	611
0.5	156	0.5	143	0.5	549	0.5	2091
0.6	185	0.6	543	0.6	515	0.6	262
0.7	145	0.7	1203	0.7	366	0.7	7
0.8	127	0.8	926	0.8	261	0.8	1
0.9	64	0.9	137	0.9	151	0.9	0
1	50	1	3	1	89	1	0
1.1	36	1.1	0	1.1	39	1.1	0
1.2	21	1.2	0	1.2	29	1.2	0
More	887	More	0	More	25	More	0

Table 23. Histograms of MLE Estimator of Parameters: TVP with T=0.7 and var=0.02 Sample Size=100

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	499	-1.2	0	-1.2	0	-1.2	0
-1.1	9	-1.1	0	-1.1	0	-1.1	0
-1	22	-1	0	-1	0	-1	0
-0.9	23	-0.9	0	-0.9	0	-0.9	0
-0.8	17	-0.8	0	-0.8	0	-0.8	0
-0.7	22	-0.7	2	-0.7	0	-0.7	0
-0.6	21	-0.6	1	-0.6	0	-0.6	0
-0.5	26	-0.5	0	-0.5	0	-0.5	0
-0.4	27	-0.4	3	-0.4	0	-0.4	0
-0.3	27	-0.3	3	-0.3	0	-0.3	0
-0.2	40	-0.2	4	-0.2	0	-0.2	0
-0.1	47	-0.1	7	-0.1	0	-0.1	0
0	41	0	41	0	2	0	0
0.1	69	0.1	55	0.1	8	0.1	173
0.2	69	0.2	45	0.2	30	0.2	2726
0.3	138	0.3	59	0.3	128	0.3	99
0.4	173	0.4	114	0.4	331	0.4	2
0.5	240	0.5	262	0.5	520	0.5	0
0.6	246	0.6	634	0.6	540	0.6	0
0.7	194	0.7	928	0.7	420	0.7	0
0.8	114	0.8	666	0.8	324	0.8	0
0.9	81	0.9	164	0.9	215	0.9	0
1	53	1	12	1	141	1	0
1.1	36	1.1	0	1.1	78	1.1	0
1.2	38	1.2	0	1.2	50	1.2	0
More	728	More	0	More	213	More	0

Table 24. Histograms of MLE Estimator of Parameters: TVP with T=0.3 and var=0.2 Sample Size=100

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	629	-1.2	0	-1.2	1	-1.2	0
-1.1	14	-1.1	0	-1.1	0	-1.1	0
-1	15	-1	0	-1	0	-1	0
-0.9	17	-0.9	0	-0.9	0	-0.9	0
-0.8	13	-0.8	0	-0.8	0	-0.8	0
-0.7	26	-0.7	0	-0.7	1	-0.7	0
-0.6	17	-0.6	0	-0.6	1	-0.6	0
-0.5	13	-0.5	0	-0.5	1	-0.5	0
-0.4	28	-0.4	0	-0.4	4	-0.4	0
-0.3	24	-0.3	2	-0.3	11	-0.3	0
-0.2	26	-0.2	8	-0.2	14	-0.2	0
-0.1	43	-0.1	14	-0.1	28	-0.1	0
0	36	0	38	0	34	0	0
0.1	48	0.1	170	0.1	55	0.1	8
0.2	69	0.2	545	0.2	109	0.2	25
0.3	123	0.3	929	0.3	188	0.3	101
0.4	154	0.4	830	0.4	345	0.4	642
0.5	206	0.5	342	0.5	687	0.5	2015
0.6	234	0.6	75	0.6	660	0.6	206
0.7	167	0.7	34	0.7	426	0.7	3
0.8	134	0.8	7	0.8	217	0.8	0
0.9	66	0.9	4	0.9	103	0.9	0
1	52	1	2	1	50	1	0
1.1	48	1.1	0	1.1	31	1.1	0
1.2	26	1.2	0	1.2	15	1.2	0
More	772	More	0	More	19	More	0

Table 25. Histograms of MLE Estimator of Parameters: TVP with T=0.3 and var=0.02 Sample Size=100

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	356	-1.2	0	-1.2	0	-1.2	0
-1.1	13	-1.1	0	-1.1	0	-1.1	0
-1	25	-1	0	-1	0	-1	0
-0.9	12	-0.9	0	-0.9	0	-0.9	0
-0.8	20	-0.8	2	-0.8	0	-0.8	0
-0.7	26	-0.7	6	-0.7	0	-0.7	0
-0.6	22	-0.6	5	-0.6	0	-0.6	0
-0.5	29	-0.5	5	-0.5	0	-0.5	0
-0.4	39	-0.4	6	-0.4	0	-0.4	0
-0.3	40	-0.3	7	-0.3	0	-0.3	0
-0.2	34	-0.2	36	-0.2	1	-0.2	0
-0.1	46	-0.1	56	-0.1	0	-0.1	0
0	56	0	195	0	2	0	0
0.1	60	0.1	390	0.1	7	0.1	350
0.2	86	0.2	678	0.2	24	0.2	2595
0.3	122	0.3	749	0.3	111	0.3	55
0.4	208	0.4	471	0.4	309	0.4	0
0.5	305	0.5	213	0.5	617	0.5	0
0.6	331	0.6	119	0.6	801	0.6	0
0.7	218	0.7	42	0.7	621	0.7	0
0.8	130	0.8	17	0.8	342	0.8	0
0.9	80	0.9	1	0.9	106	0.9	0
1	65	1	2	1	32	1	0
1.1	60	1.1	0	1.1	9	1.1	0
1.2	48	1.2	0	1.2	9	1.2	0
More	569	More	0	More	9	More	0

Table 26. Histograms of MLE Estimator of Parameters: TVP with T=0.7 and var=0.2 Sample Size=400

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	387	-1.2	0	-1.2	0	-1.2	0
-1.1	9	-1.1	0	-1.1	0	-1.1	0
-1	10	-1	0	-1	0	-1	0
-0.9	14	-0.9	0	-0.9	0	-0.9	0
-0.8	9	-0.8	0	-0.8	0	-0.8	0
-0.7	16	-0.7	0	-0.7	0	-0.7	0
-0.6	11	-0.6	0	-0.6	0	-0.6	0
-0.5	20	-0.5	0	-0.5	0	-0.5	0
-0.4	13	-0.4	0	-0.4	0	-0.4	0
-0.3	15	-0.3	0	-0.3	0	-0.3	0
-0.2	16	-0.2	0	-0.2	0	-0.2	0
-0.1	20	-0.1	0	-0.1	0	-0.1	0
0	35	0	0	0	0	0	0
0.1	61	0.1	0	0.1	2	0.1	0
0.2	71	0.2	0	0.2	8	0.2	0
0.3	147	0.3	0	0.3	42	0.3	0
0.4	242	0.4	0	0.4	290	0.4	70
0.5	351	0.5	0	0.5	1077	0.5	2919
0.6	345	0.6	53	0.6	1189	0.6	11
0.7	265	0.7	1679	0.7	339	0.7	0
0.8	190	0.8	1268	0.8	48	0.8	0
0.9	72	0.9	0	0.9	4	0.9	0
1	55	1	0	1	1	1	0
1.1	36	1.1	0	1.1	0	1.1	0
1.2	32	1.2	0	1.2	0	1.2	0
More	558	More	0	More	0	More	0

Table 27. Histograms of MLE Estimator of Parameters: TVP with T=0.7 and var=0.02 Sample Size=400

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	200	-1.2	0	-1.2	0	-1.2	0
-1.1	13	-1.1	0	-1.1	0	-1.1	0
-1	16	-1	0	-1	0	-1	0
-0.9	14	-0.9	0	-0.9	0	-0.9	0
-0.8	15	-0.8	0	-0.8	0	-0.8	0
-0.7	18	-0.7	0	-0.7	0	-0.7	0
-0.6	11	-0.6	0	-0.6	0	-0.6	0
-0.5	15	-0.5	0	-0.5	0	-0.5	0
-0.4	21	-0.4	0	-0.4	0	-0.4	0
-0.3	24	-0.3	0	-0.3	0	-0.3	0
-0.2	26	-0.2	0	-0.2	0	-0.2	0
-0.1	30	-0.1	0	-0.1	0	-0.1	0
0	40	0	0	0	0	0	0
0.1	46	0.1	0	0.1	0	0.1	3
0.2	85	0.2	0	0.2	0	0.2	2997
0.3	120	0.3	0	0.3	8	0.3	0
0.4	290	0.4	3	0.4	177	0.4	0
0.5	511	0.5	11	0.5	1028	0.5	0
0.6	522	0.6	182	0.6	1207	0.6	0
0.7	261	0.7	1647	0.7	470	0.7	0
0.8	124	0.8	1140	0.8	87	0.8	0
0.9	59	0.9	17	0.9	13	0.9	0
1	47	1	0	1	8	1	0
1.1	51	1.1	0	1.1	2	1.1	0
1.2	36	1.2	0	1.2	0	1.2	0
More	405	More	0	More	0	More	0

Table 28. Histograms of MLE Estimator of Parameters: TVP with T=0.3 and var=0.2 Sample Size=400

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	259	-1.2	0	-1.2	0	-1.2	0
-1.1	14	-1.1	0	-1.1	0	-1.1	0
-1	10	-1	0	-1	0	-1	0
-0.9	9	-0.9	0	-0.9	0	-0.9	0
-0.8	15	-0.8	0	-0.8	0	-0.8	0
-0.7	7	-0.7	0	-0.7	0	-0.7	0
-0.6	20	-0.6	0	-0.6	0	-0.6	0
-0.5	27	-0.5	0	-0.5	0	-0.5	0
-0.4	22	-0.4	0	-0.4	0	-0.4	0
-0.3	24	-0.3	0	-0.3	0	-0.3	0
-0.2	30	-0.2	0	-0.2	0	-0.2	0
-0.1	29	-0.1	0	-0.1	0	-0.1	0
0	37	0	0	0	0	0	0
0.1	71	0.1	2	0.1	6	0.1	0
0.2	67	0.2	125	0.2	7	0.2	0
0.3	127	0.3	1459	0.3	18	0.3	0
0.4	269	0.4	1325	0.4	185	0.4	123
0.5	434	0.5	89	0.5	1271	0.5	2868
0.6	479	0.6	0	0.6	1287	0.6	9
0.7	254	0.7	0	0.7	194	0.7	0
0.8	158	0.8	0	0.8	24	0.8	0
0.9	67	0.9	0	0.9	5	0.9	0
1	47	1	0	1	0	1	0
1.1	47	1.1	0	1.1	2	1.1	0
1.2	25	1.2	0	1.2	0	1.2	0
More	452	More	0	More	1	More	0

Table 29. Histograms of MLE Estimator of Parameters: TVP with T=0.3 and var=0.02 Sample Size=400

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	77	-1.2	0	-1.2	0	-1.2	0
-1.1	4	-1.1	0	-1.1	0	-1.1	0
-1	10	-1	0	-1	0	-1	0
-0.9	12	-0.9	0	-0.9	0	-0.9	0
-0.8	10	-0.8	0	-0.8	0	-0.8	0
-0.7	15	-0.7	0	-0.7	0	-0.7	0
-0.6	12	-0.6	0	-0.6	0	-0.6	0
-0.5	10	-0.5	0	-0.5	0	-0.5	0
-0.4	23	-0.4	0	-0.4	0	-0.4	0
-0.3	22	-0.3	1	-0.3	0	-0.3	0
-0.2	29	-0.2	0	-0.2	0	-0.2	0
-0.1	24	-0.1	1	-0.1	0	-0.1	0
0	35	0	12	0	0	0	0
0.1	61	0.1	39	0.1	0	0.1	5
0.2	75	0.2	293	0.2	0	0.2	2995
0.3	131	0.3	1408	0.3	2	0.3	0
0.4	293	0.4	1073	0.4	92	0.4	0
0.5	657	0.5	161	0.5	1200	0.5	0
0.6	648	0.6	12	0.6	1491	0.6	0
0.7	288	0.7	0	0.7	190	0.7	0
0.8	112	0.8	0	0.8	21	0.8	0
0.9	65	0.9	0	0.9	3	0.9	0
1	48	1	0	1	1	1	0
1.1	40	1.1	0	1.1	0	1.1	0
1.2	40	1.2	0	1.2	0	1.2	0
More	259	More	0	More	0	More	0

Table 30. Histograms of MLE Estimator of Parameters: TVP Generated as Smooth Transition γ=0.5 Sample Size=100

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	135	-1.2	0	-1.2	0	-1.2	0
-1.1	18	-1.1	0	-1.1	0	-1.1	0
-1	27	-1	0	-1	0	-1	0
-0.9	39	-0.9	0	-0.9	0	-0.9	0
-0.8	35	-0.8	3	-0.8	0	-0.8	0
-0.7	28	-0.7	6	-0.7	0	-0.7	0
-0.6	42	-0.6	9	-0.6	0	-0.6	0
-0.5	38	-0.5	6	-0.5	0	-0.5	0
-0.4	49	-0.4	14	-0.4	0	-0.4	0
-0.3	35	-0.3	22	-0.3	0	-0.3	0
-0.2	58	-0.2	31	-0.2	0	-0.2	0
-0.1	64	-0.1	77	-0.1	0	-0.1	0
0	69	0	363	0	240	0	0
0.1	71	0.1	628	0.1	1027	0.1	2931
0.2	113	0.2	191	0.2	54	0.2	66
0.3	153	0.3	62	0.3	35	0.3	3
0.4	202	0.4	44	0.4	31	0.4	0
0.5	280	0.5	45	0.5	22	0.5	0
0.6	279	0.6	63	0.6	39	0.6	0
0.7	242	0.7	42	0.7	41	0.7	0
0.8	168	0.8	40	0.8	39	0.8	0
0.9	127	0.9	58	0.9	23	0.9	0
1	66	1	1013	1	29	1	0
1.1	67	1.1	283	1.1	33	1.1	0
1.2	63	1.2	0	1.2	77	1.2	0
More	532	More	0	More	1310	More	0

Table 31. Histograms of MLE Estimator of Parameters: TVP Generated as Smooth Transition γ=1.5 Sample Size=100

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	107	-1.2	0	-1.2	0	-1.2	0
-1.1	18	-1.1	0	-1.1	0	-1.1	0
-1	14	-1	0	-1	0	-1	0
-0.9	20	-0.9	0	-0.9	0	-0.9	0
-0.8	18	-0.8	1	-0.8	0	-0.8	0
-0.7	28	-0.7	2	-0.7	0	-0.7	0
-0.6	24	-0.6	1	-0.6	0	-0.6	0
-0.5	35	-0.5	3	-0.5	0	-0.5	0
-0.4	32	-0.4	7	-0.4	0	-0.4	0
-0.3	42	-0.3	10	-0.3	0	-0.3	0
-0.2	55	-0.2	9	-0.2	0	-0.2	0
-0.1	51	-0.1	22	-0.1	0	-0.1	0
0	86	0	62	0	562	0	0
0.1	109	0.1	124	0.1	1869	0.1	2905
0.2	129	0.2	86	0.2	26	0.2	89
0.3	178	0.3	46	0.3	30	0.3	6
0.4	228	0.4	26	0.4	28	0.4	0
0.5	279	0.5	26	0.5	21	0.5	0
0.6	274	0.6	31	0.6	24	0.6	0
0.7	261	0.7	29	0.7	22	0.7	0
0.8	169	0.8	34	0.8	16	0.8	0
0.9	134	0.9	38	0.9	22	0.9	0
1	121	1	1407	1	27	1	0
1.1	67	1.1	1036	1.1	52	1.1	0
1.2	66	1.2	0	1.2	90	1.2	0
More	455	More	0	More	211	More	0

Table 32. Histograms of MLE Estimator of Parameters: TVP Generated as Smooth Transition γ=0.5 Sample Size=400

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	0	-1.2	0	-1.2	0	-1.2	0
-1.1	0	-1.1	0	-1.1	0	-1.1	0
-1	1	-1	0	-1	0	-1	0
-0.9	1	-0.9	0	-0.9	0	-0.9	0
-0.8	4	-0.8	0	-0.8	0	-0.8	0
-0.7	2	-0.7	0	-0.7	0	-0.7	0
-0.6	3	-0.6	1	-0.6	0	-0.6	0
-0.5	9	-0.5	0	-0.5	0	-0.5	0
-0.4	6	-0.4	1	-0.4	0	-0.4	0
-0.3	14	-0.3	0	-0.3	0	-0.3	0
-0.2	11	-0.2	1	-0.2	0	-0.2	0
-0.1	17	-0.1	1	-0.1	0	-0.1	0
0	32	0	2	0	72	0	0
0.1	60	0.1	5	0.1	2909	0.1	2999
0.2	69	0.2	2	0.2	1	0.2	1
0.3	155	0.3	1	0.3	0	0.3	0
0.4	331	0.4	0	0.4	1	0.4	0
0.5	796	0.5	0	0.5	2	0.5	0
0.6	768	0.6	1	0.6	0	0.6	0
0.7	331	0.7	2	0.7	1	0.7	0
0.8	147	0.8	1	0.8	0	0.8	0
0.9	86	0.9	1	0.9	0	0.9	0
1	62	1	2764	1	0	1	0
1.1	30	1.1	217	1.1	1	1.1	0
1.2	22	1.2	0	1.2	2	1.2	0
More	43	More	0	More	11	More	0

Table 33. Histograms of MLE Estimator of Parameters: TVP Generated as Smooth Transition γ=1.5 Sample Size=400

α		T		μ		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	0	-1.2	0	-1.2	0	-1.2	0
-1.1	0	-1.1	0	-1.1	0	-1.1	0
-1	1	-1	0	-1	0	-1	0
-0.9	3	-0.9	0	-0.9	0	-0.9	0
-0.8	1	-0.8	0	-0.8	0	-0.8	0
-0.7	3	-0.7	0	-0.7	0	-0.7	0
-0.6	4	-0.6	0	-0.6	0	-0.6	0
-0.5	4	-0.5	0	-0.5	0	-0.5	0
-0.4	8	-0.4	0	-0.4	0	-0.4	0
-0.3	11	-0.3	0	-0.3	0	-0.3	0
-0.2	20	-0.2	0	-0.2	0	-0.2	0
-0.1	19	-0.1	0	-0.1	0	-0.1	0
0	31	0	0	0	182	0	0
0.1	45	0.1	0	0.1	2818	0.1	3000
0.2	77	0.2	0	0.2	0	0.2	0
0.3	146	0.3	0	0.3	0	0.3	0
0.4	341	0.4	0	0.4	0	0.4	0
0.5	808	0.5	0	0.5	0	0.5	0
0.6	767	0.6	0	0.6	0	0.6	0
0.7	330	0.7	0	0.7	0	0.7	0
0.8	147	0.8	0	0.8	0	0.8	0
0.9	79	0.9	0	0.9	0	0.9	0
1	51	1	181	1	0	1	0
1.1	41	1.1	2819	1.1	0	1.1	0
1.2	18	1.2	0	1.2	0	1.2	0
More	45	More	0	More	0	More	0

Table 34. Distribution of Mean of dv_t over 3000 Sample N=100 (1)

Noncointegrated		linearly Related		TVP T=0.3	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	30	-1.2	0	-1.2	0
-1.1	12	-1.1	0	-1.1	0
-1	16	-1	0	-1	0
-0.9	16	-0.9	0	-0.9	0
-0.8	21	-0.8	0	-0.8	0
-0.7	48	-0.7	0	-0.7	0
-0.6	52	-0.6	0	-0.6	0
-0.5	64	-0.5	0	-0.5	0
-0.4	61	-0.4	0	-0.4	0
-0.3	108	-0.3	0	-0.3	0
-0.2	130	-0.2	0	-0.2	0
-0.1	168	-0.1	0	-0.1	0
0	201	0	0	0	0
0.1	212	0.1	0	0.1	0
0.2	301	0.2	0	0.2	0
0.3	345	0.3	0	0.3	0
0.4	347	0.4	0	0.4	0
0.5	369	0.5	0	0.5	0
0.6	278	0.6	0	0.6	0
0.7	162	0.7	5	0.7	0
0.8	57	0.8	151	0.8	3
0.9	2	0.9	2385	0.9	2601
1	0	1	459	1	396
1.1	0	1.1	0	1.1	0
1.2	0	1.2	0	1.2	0
More	0	More	0	More	0

Table 35. Distribution of Mean of dv_t over 3000 Sample N=100 (2)

TVP T=0.7		Smooth γ=0.5		Smooth γ=1.5	
dv					
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	0	-1.2	0	-1.2	0
-1.1	0	-1.1	0	-1.1	0
-1	0	-1	0	-1	0
-0.9	0	-0.9	0	-0.9	0
-0.8	0	-0.8	0	-0.8	0
-0.7	0	-0.7	0	-0.7	0
-0.6	0	-0.6	0	-0.6	0
-0.5	0	-0.5	0	-0.5	0
-0.4	0	-0.4	0	-0.4	0
-0.3	0	-0.3	0	-0.3	0
-0.2	0	-0.2	0	-0.2	0
-0.1	0	-0.1	0	-0.1	0
0	0	0	0	0	0
0.1	0	0.1	0	0.1	1
0.2	0	0.2	0	0.2	1
0.3	0	0.3	0	0.3	5
0.4	0	0.4	0	0.4	4
0.5	2	0.5	0	0.5	18
0.6	33	0.6	1	0.6	31
0.7	502	0.7	6	0.7	114
0.8	1865	0.8	77	0.8	559
0.9	598	0.9	1513	0.9	2073
1	0	1	1403	1	194
1.1	0	1.1	0	1.1	0
1.2	0	1.2	0	1.2	0
More	0	More	0	More	0

Table 36. Distribution of Mean of dv_t over 3000 Sample N=400 (1)

Noncointegrated		Linearly Related		TVP T=0.3	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	1	-1.2	0	-1.2	0
-1.1	4	-1.1	0	-1.1	0
-1	3	-1	0	-1	0
-0.9	9	-0.9	0	-0.9	0
-0.8	7	-0.8	0	-0.8	0
-0.7	20	-0.7	0	-0.7	0
-0.6	27	-0.6	0	-0.6	0
-0.5	46	-0.5	0	-0.5	0
-0.4	75	-0.4	0	-0.4	0
-0.3	133	-0.3	0	-0.3	0
-0.2	190	-0.2	0	-0.2	0
-0.1	253	-0.1	0	-0.1	0
0	354	0	0	0	0
0.1	464	0.1	0	0.1	0
0.2	494	0.2	0	0.2	0
0.3	463	0.3	0	0.3	0
0.4	306	0.4	0	0.4	0
0.5	116	0.5	0	0.5	0
0.6	32	0.6	0	0.6	0
0.7	3	0.7	0	0.7	0
0.8	0	0.8	0	0.8	1
0.9	0	0.9	429	0.9	2142
1	0	1	2571	1	857
1.1	0	1.1	0	1.1	0
1.2	0	1.2	0	1.2	0
More	0	More	0	More	0

Table 37. Distribution of Mean of dv_t over 3000 Sample N=400 (2)

TVP T=0.7		Smooth γ=0.5		Smooth γ=1.5	
bin	*Frequency*	*bin*	*Frequency*	bin	Frequency
-1.2	0	-1.2	0	-1.2	0
-1.1	0	-1.1	0	-1.1	0
-1	0	-1	0	-1	0
-0.9	0	-0.9	0	-0.9	0
-0.8	0	-0.8	0	-0.8	0
-0.7	0	-0.7	0	-0.7	0
-0.6	0	-0.6	0	-0.6	0
-0.5	0	-0.5	0	-0.5	0
-0.4	0	-0.4	0	-0.4	0
-0.3	0	-0.3	0	-0.3	0
-0.2	0	-0.2	0	-0.2	0
-0.1	0	-0.1	0	-0.1	0
0	0	0	0	0	0
0.1	0	0.1	0	0.1	0
0.2	0	0.2	2	0.2	0
0.3	0	0.3	1	0.3	0
0.4	0	0.4	10	0.4	0
0.5	2	0.5	31	0.5	0
0.6	33	0.6	73	0.6	1
0.7	502	0.7	177	0.7	6
0.8	1865	0.8	459	0.8	77
0.9	598	0.9	1735	0.9	1513
1	0	1	512	1	1403
1.1	0	1.1	0	1.1	0
1.2	0	1.2	0	1.2	0
More	0	More	0	More	0

Table 38. Histogram of MLE estimates Independent Random Walks Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	1283	-1.2	9	-1.2	0	-1.2	0
-1.1	7	-1.1	4	-1.1	0	-1.1	0
-1	5	-1	9	-1	0	-1	0
-0.9	7	-0.9	7	-0.9	0	-0.9	0
-0.8	5	-0.8	21	-0.8	0	-0.8	0
-0.7	9	-0.7	27	-0.7	0	-0.7	0
-0.6	7	-0.6	31	-0.6	0	-0.6	0
-0.5	5	-0.5	38	-0.5	0	-0.5	0
-0.4	6	-0.4	86	-0.4	0	-0.4	0
-0.3	2	-0.3	117	-0.3	0	-0.3	0
-0.2	5	-0.2	194	-0.2	0	-0.2	0
-0.1	4	-0.1	344	-0.1	0	-0.1	0
0	7	0	649	0	0	0	0
0.1	9	0.1	616	0.1	931	0.1	1551
0.2	5	0.2	336	0.2	18	0.2	690
0.3	4	0.3	182	0.3	17	0.3	328
0.4	5	0.4	110	0.4	27	0.4	181
0.5	7	0.5	79	0.5	42	0.5	88
0.6	6	0.6	44	0.6	43	0.6	68
0.7	8	0.7	30	0.7	42	0.7	33
0.8	9	0.8	17	0.8	56	0.8	17
0.9	7	0.9	13	0.9	58	0.9	14
1	4	1	11	1	64	1	8
1.1	4	1.1	10	1.1	52	1.1	6
1.2	5	1.2	5	1.2	46	1.2	5
More	1575	More	11	More	1604	More	11

Table 39. Histogram of MLE Estimates of Linearly Related Random Walks Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	872	-1.2	0	-1.2	0	-1.2	0
-1.1	9	-1.1	0	-1.1	0	-1.1	0
-1	12	-1	0	-1	0	-1	0
-0.9	7	-0.9	0	-0.9	0	-0.9	0
-0.8	14	-0.8	0	-0.8	0	-0.8	0
-0.7	8	-0.7	0	-0.7	0	-0.7	0
-0.6	16	-0.6	0	-0.6	0	-0.6	0
-0.5	17	-0.5	0	-0.5	0	-0.5	0
-0.4	14	-0.4	0	-0.4	0	-0.4	0
-0.3	22	-0.3	0	-0.3	0	-0.3	0
-0.2	20	-0.2	0	-0.2	0	-0.2	0
-0.1	21	-0.1	0	-0.1	0	-0.1	0
0	32	0	106	0	0	0	0
0.1	39	0.1	2303	0.1	10	0.1	2319
0.2	75	0.2	584	0.2	1	0.2	582
0.3	77	0.3	7	0.3	6	0.3	88
0.4	97	0.4	0	0.4	6	0.4	11
0.5	110	0.5	0	0.5	23	0.5	0
0.6	122	0.6	0	0.6	64	0.6	0
0.7	95	0.7	0	0.7	126	0.7	0
0.8	102	0.8	0	0.8	329	0.8	0
0.9	58	0.9	0	0.9	645	0.9	0
1	50	1	0	1	868	1	0
1.1	35	1.1	0	1.1	588	1.1	0
1.2	24	1.2	0	1.2	235	1.2	0
More	1052	More	0	More	99	More	0

Table 40. Histograms of MLE Estimator of Parameters: TVP with T=0.7 and var=0.2 Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	906	-1.2	0	-1.2	0	-1.2	0
-1.1	7	-1.1	0	-1.1	0	-1.1	0
-1	6	-1	1	-1	0	-1	0
-0.9	9	-0.9	1	-0.9	0	-0.9	0
-0.8	7	-0.8	1	-0.8	0	-0.8	0
-0.7	10	-0.7	0	-0.7	0	-0.7	0
-0.6	10	-0.6	1	-0.6	0	-0.6	0
-0.5	10	-0.5	3	-0.5	0	-0.5	0
-0.4	14	-0.4	6	-0.4	0	-0.4	0
-0.3	12	-0.3	18	-0.3	0	-0.3	0
-0.2	20	-0.2	45	-0.2	0	-0.2	0
-0.1	27	-0.1	96	-0.1	0	-0.1	0
0	23	0	169	0	0	0	0
0.1	41	0.1	268	0.1	761	0.1	1
0.2	66	0.2	411	0.2	16	0.2	6
0.3	73	0.3	608	0.3	13	0.3	58
0.4	99	0.4	557	0.4	12	0.4	474
0.5	122	0.5	371	0.5	30	0.5	1434
0.6	116	0.6	215	0.6	50	0.6	720
0.7	115	0.7	105	0.7	80	0.7	165
0.8	86	0.8	47	0.8	131	0.8	82
0.9	52	0.9	37	0.9	174	0.9	25
1	44	1	20	1	214	1	15
1.1	33	1.1	7	1.1	229	1.1	9
1.2	22	1.2	4	1.2	137	1.2	4
More	1070	More	9	More	1153	More	7

Table 41. Histograms of MLE Estimator of Parameters: TVP with T=0.7 and var=0.02 Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	917	-1.2	0	-1.2	0	-1.2	0
-1.1	6	-1.1	0	-1.1	0	-1.1	0
-1	7	-1	0	-1	0	-1	0
-0.9	7	-0.9	0	-0.9	0	-0.9	0
-0.8	3	-0.8	0	-0.8	0	-0.8	0
-0.7	8	-0.7	0	-0.7	0	-0.7	0
-0.6	9	-0.6	0	-0.6	0	-0.6	0
-0.5	8	-0.5	1	-0.5	0	-0.5	0
-0.4	17	-0.4	3	-0.4	0	-0.4	0
-0.3	16	-0.3	1	-0.3	0	-0.3	0
-0.2	21	-0.2	2	-0.2	0	-0.2	0
-0.1	33	-0.1	36	-0.1	0	-0.1	0
0	33	0	152	0	0	0	0
0.1	47	0.1	414	0.1	631	0.1	53
0.2	66	0.2	834	0.2	14	0.2	1664
0.3	85	0.3	909	0.3	27	0.3	873
0.4	89	0.4	388	0.4	31	0.4	274
0.5	101	0.5	156	0.5	50	0.5	78
0.6	122	0.6	62	0.6	66	0.6	29
0.7	111	0.7	23	0.7	110	0.7	15
0.8	80	0.8	6	0.8	209	0.8	8
0.9	43	0.9	7	0.9	285	0.9	2
1	51	1	3	1	358	1	2
1.1	21	1.1	2	1.1	336	1.1	1
1.2	30	1.2	0	1.2	208	1.2	1
More	1069	More	1	More	675	More	0

Table 42. Histograms of MLE Estimator of Parameters: TVP with T=0.3 and var=0.2 Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	829	-1.2	0	-1.2	0	-1.2	0
-1.1	9	-1.1	0	-1.1	0	-1.1	0
-1	12	-1	0	-1	0	-1	0
-0.9	10	-0.9	0	-0.9	0	-0.9	0
-0.8	8	-0.8	0	-0.8	0	-0.8	0
-0.7	8	-0.7	0	-0.7	0	-0.7	0
-0.6	11	-0.6	0	-0.6	0	-0.6	0
-0.5	7	-0.5	0	-0.5	0	-0.5	0
-0.4	17	-0.4	0	-0.4	0	-0.4	0
-0.3	17	-0.3	0	-0.3	0	-0.3	0
-0.2	27	-0.2	3	-0.2	0	-0.2	0
-0.1	17	-0.1	14	-0.1	0	-0.1	0
0	32	0	387	0	0	0	0
0.1	37	0.1	1739	0.1	41	0.1	498
0.2	68	0.2	809	0.2	1	0.2	684
0.3	84	0.3	46	0.3	5	0.3	552
0.4	125	0.4	2	0.4	4	0.4	430
0.5	146	0.5	0	0.5	10	0.5	523
0.6	146	0.6	0	0.6	9	0.6	283
0.7	120	0.7	0	0.7	31	0.7	26
0.8	86	0.8	0	0.8	51	0.8	4
0.9	58	0.9	0	0.9	87	0.9	0
1	57	1	0	1	112	1	0
1.1	35	1.1	0	1.1	152	1.1	0
1.2	20	1.2	0	1.2	163	1.2	0
More	1014	More	0	More	2334	More	0

Table 43. Histograms of MLE Estimator of Parameters: TVP with T=0.3 and var=0.02 Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	855	-1.2	0	-1.2	0	-1.2	0
-1.1	8	-1.1	0	-1.1	0	-1.1	0
-1	8	-1	0	-1	0	-1	0
-0.9	10	-0.9	0	-0.9	0	-0.9	0
-0.8	5	-0.8	0	-0.8	0	-0.8	0
-0.7	11	-0.7	0	-0.7	0	-0.7	0
-0.6	5	-0.6	0	-0.6	0	-0.6	0
-0.5	16	-0.5	0	-0.5	0	-0.5	0
-0.4	14	-0.4	0	-0.4	0	-0.4	0
-0.3	15	-0.3	0	-0.3	0	-0.3	0
-0.2	22	-0.2	0	-0.2	0	-0.2	0
-0.1	23	-0.1	1	-0.1	0	-0.1	0
0	41	0	204	0	0	0	0
0.1	45	0.1	2346	0.1	7	0.1	1735
0.2	59	0.2	442	0.2	1	0.2	1120
0.3	101	0.3	6	0.3	2	0.3	134
0.4	117	0.4	1	0.4	0	0.4	10
0.5	117	0.5	0	0.5	4	0.5	1
0.6	143	0.6	0	0.6	5	0.6	0
0.7	122	0.7	0	0.7	16	0.7	0
0.8	81	0.8	0	0.8	35	0.8	0
0.9	51	0.9	0	0.9	102	0.9	0
1	35	1	0	1	227	1	0
1.1	44	1.1	0	1.1	320	1.1	0
1.2	22	1.2	0	1.2	288	1.2	0
More	1030	More	0	More	1993	More	0

Table 44. Histograms of MLE Estimator of Parameters: TVP Smooth Transition with γ=0.5 Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	898	-1.2	0	-1.2	0	-1.2	0
-1.1	7	-1.1	0	-1.1	0	-1.1	0
-1	7	-1	0	-1	0	-1	0
-0.9	7	-0.9	0	-0.9	0	-0.9	0
-0.8	5	-0.8	0	-0.8	0	-0.8	0
-0.7	11	-0.7	0	-0.7	0	-0.7	0
-0.6	13	-0.6	0	-0.6	0	-0.6	0
-0.5	14	-0.5	0	-0.5	0	-0.5	0
-0.4	14	-0.4	0	-0.4	0	-0.4	0
-0.3	19	-0.3	0	-0.3	0	-0.3	0
-0.2	27	-0.2	0	-0.2	0	-0.2	0
-0.1	27	-0.1	0	-0.1	0	-0.1	0
0	24	0	53	0	0	0	0
0.1	49	0.1	1195	0.1	85	0.1	1759
0.2	59	0.2	1449	0.2	12	0.2	928
0.3	86	0.3	289	0.3	17	0.3	245
0.4	103	0.4	13	0.4	40	0.4	51
0.5	104	0.5	1	0.5	78	0.5	15
0.6	110	0.6	0	0.6	132	0.6	2
0.7	99	0.7	0	0.7	216	0.7	0
0.8	94	0.8	0	0.8	379	0.8	0
0.9	54	0.9	0	0.9	583	0.9	0
1	52	1	0	1	634	1	0
1.1	30	1.1	0	1.1	494	1.1	0
1.2	30	1.2	0	1.2	179	1.2	0
More	1057	More	0	More	151	More	0

Table 45. Histograms of MLE Estimator of Parameters: TVP Smooth Transition with γ=1.5 Model Penalty Applied N=100

α		μ		H		σ_q	
bin	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*	*bin*	*Frequency*
-1.2	898	-1.2	0	-1.2	0	-1.2	0
-1.1	7	-1.1	0	-1.1	0	-1.1	0
-1	7	-1	0	-1	0	-1	0
-0.9	7	-0.9	0	-0.9	0	-0.9	0
-0.8	5	-0.8	0	-0.8	0	-0.8	0
-0.7	11	-0.7	0	-0.7	0	-0.7	0
-0.6	13	-0.6	0	-0.6	0	-0.6	0
-0.5	14	-0.5	0	-0.5	0	-0.5	0
-0.4	14	-0.4	0	-0.4	0	-0.4	0
-0.3	19	-0.3	0	-0.3	0	-0.3	0
-0.2	27	-0.2	0	-0.2	0	-0.2	0
-0.1	27	-0.1	0	-0.1	0	-0.1	0
0	24	0	53	0	0	0	0
0.1	49	0.1	1195	0.1	85	0.1	1759
0.2	59	0.2	1449	0.2	12	0.2	928
0.3	86	0.3	289	0.3	17	0.3	245
0.4	103	0.4	13	0.4	40	0.4	51
0.5	104	0.5	1	0.5	78	0.5	15
0.6	110	0.6	0	0.6	132	0.6	2
0.7	99	0.7	0	0.7	216	0.7	0
0.8	94	0.8	0	0.8	379	0.8	0
0.9	54	0.9	0	0.9	583	0.9	0
1	52	1	0	1	634	1	0
1.1	30	1.1	0	1.1	494	1.1	0
1.2	30	1.2	0	1.2	179	1.2	0
More	1057	More	0	More	151	More	0