AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

A State Dependent Regime Switching model of Dynamic Correlations

Hernan A. Tejeda

Graduate Research Assistant
North Carolina State University
email: hatejeda@ncsu.edu

Barry K. Goodwin

William Neal Reynolds Distinguished Professor
North Carolina State University
email: barry_goodwin@ncsu.edu

Denis Pelletier

Assistant Professor
North Carolina State University
email: denis_pelletier@ncsu.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, Wisconsin, July 26-29, 2009

Copyright 2009 by Hernan A. Tejeda, Barry K. Goodwin, and Denis Pelletier. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

A State Dependent Regime Switching model of Dynamic Correlations

Hernan A. Tejeda, Barry K. Goodwin, and Denis Pelletier

Abstract

We extend the Regime Switching for Dynamic Correlations (RSDC) model by Pelletier (Journal of Econometrics, 2006), to determine the effect of underlying fundamental variables in the evolution of the dynamic correlations between multiple time series. By introducing state dependent transition probabilities to the switching process between different regimes - governed by a Markov chain, we are able to identify potential thresholds and spillover effects in the dynamic process. In addition, asymmetric correlations between the series are determined. We simulate data for multiple series and find an initial better fit of state dependent transition probabilities, versus constant transition probabilities, for the regime switching model. Capturing more precisely the dynamic interrelationships between multiple series or markets conveys many benefits including - potential efficiency gains from related operations, determining the effects of shocks from related variables, as well as improvement in hedging operations.

Key Words: dynamic correlations, regime switching, state dependent probabilities, thresholds, spillovers

---

Hernan A. Tejeda is a Graduate Research Assistant in the Departments of Economics and Agricultural and Resource Economics at North Carolina State University.

Barry K. Goodwin is William Neal Reynolds Distinguished Professor in the Departments of Economics and Agricultural and Resource Economics at North Carolina State University.

Denis Pelletier is Assistant Professor in the Department of Economics at North Carolina State University.

Introduction

Financial instruments and more specifically commodity prices for agricultural products – such as crops and/or livestock, constitute an important source of risk for agricultural producers, consumers, investors, etc. Fluctuation in these prices that include sharp spikes and plunges or slight increases and drops, generates risk to producers and consumers called volatility. This volatility is steadily changing through time as new information arrives. This updated information may be directly related to the product(s) or general economic news – causing the time series to have heteroskedastic properties.

Measurement of this risk gave rise to the ARCH and GARCH models by Engle (1982) and Bollerslev (1986), respectively. These models are estimated via MLE, and a Lagrange Multiplier test is used to measure the significance of autocorrelation between these squared residuals, Engle (2004). These ARCH/GARCH models are univariate, permitting parameter estimation of single time series. In addition, multivariate GARCH models have been developed that enable the estimation of time varying correlations between many instruments. i.e. multivariate GARCH models permit parameter estimation of conditional correlations among multiple time series

These multivariate models have the restriction of requiring a positive semi-definite (PSD) correlation (or variance) matrix at every period. i.e. during parameter estimation - for each period the models must incorporate this PSD correlation matrix condition. In addition, some multivariate models may allow estimations for only a few time series, thus having a dimensionality curse. i.e. they can't be estimated in two steps as a set of SURE equations, which enables to deal with this dimensionality inconvenience. An extensive amount of univariate and multivariate ARCH /GARCH models have been developed, with an earlier paper by Bollerslev et al. (1992) addressing broad model developments including empirical financial applications. Recently, Bauwens et al. (2006) presents a broad survey of the multivariate models in existence.

A simple workhorse for the multivariate GARCH model has been the Constant Conditional Correlation (CCC) model from Bollerslev (1990). In this model the correlations are constant and so the conditional covariances are linearly proportional across time - as they are product of these constant correlations and the corresponding conditional volatilities. Nonetheless, it has been shown empirically for stock markets that this condition doesn't hold, since correlations tend to increase in periods of higher volatility as per Longin and Solnik (1995) and Ramchmand and Susmel (1998). On the other hand, from Engle (2002) the Dynamic Conditional Correlation (DCC) model considers the case of varying conditional correlations through each period. For this model, the correlations between the different assets or markets considered

follow a parsimonious parametric model – specifically a GARCH-type dynamic. i.e. there is a different correlation matrix at each period, built through exponential smoothing.

Another model which considers multivariate dynamic correlations is the Regime Switching Dynamic Correlation (RSDC) model from Pelletier (2006). This model considers the correlations as constant within a regime, yet they vary in value from one regime to another. The switch between regimes, producing a change of correlation values, is governed by a Markov chain. In this model the series may remain in a specific regime for ensuing periods before switching to a different regime, and only at this different regime is there a change in the correlation values. This is in contrast to the DCC model, which may have different correlation values for each subsequent period. As mentioned previously, in the RSDC model the transition between different states (or regimes) responds to a Markov chain - with transition probabilities between each state being constant.

We extend and develop a dynamic threshold model based on the RSDC model, by modifying the transition probabilities that govern the switching process between regimes - from constant probabilities to state dependent or time-varying probabilities. In this sense, we introduce weakly exogenous variables in the probabilities that determine the switch from one regime to another. These new transitional or regime switching probabilities, now state dependent or time-varying, incorporate underlying fundamental variables that are directly related to the evolution of the series being studied. These underlying related variables in our regime switching probabilities constitute specific threshold levels which have particular effects in our dynamic process. i.e. direct impact on being at one regime of correlation or another.

By introducing variables related to the time series in the regime switching process, we seek to capture the particular effect these variables may have on the dynamic correlation process. That is, we aim to determine the impact that these related underlying variables, which are weakly exogenous, have on the evolution of the dynamic process. The state dependent probabilities that now govern our regime switching process are established following Diebold et al. (1994). Our results are compared to the initial (nested) case where switching between different states or regimes responds solely to a completely unaccounted exogenous situation. i.e. by constant transition probabilities. The specific effect that each underlying variable related to the series has on the evolution of the series is considered a threshold level. This threshold level becomes proportional to the estimated parameter of our underlying variable - in the state dependent transition probabilities as will be seen ahead in the estimation process.

We introduce underlying related variables in the regime switching process, and test the significance of their relation to the series being studied. In this way, we determine significant factors that produce an impact on the evolution of the correlations among the series that are studied. These relevant underlying

related factors may produce asymmetries in the correlations between the series. i.e., the dynamic correlations calculated would result in different values for particular positive or negative shocks of the significant related variables. These significant related variables represent specific thresholds, which are unaccounted for in the case of constant transition probabilities.

Persistence in the Markov chain may vary as a function of these weakly exogenous variables. These weakly exogenous variables may form threshold levels that inhibit switching to a different regime, i.e. correlation values are maintained as the series stay in a certain regime. Conversely, these exogenous variables may prompt switching to a different correlation level than if they had otherwise not been considered. i.e. they may catalyze switching to a different correlation value by moving from one regime to another. The identification of these variables and their role in the evolution of the correlations is compared to the case where constant probabilities govern the Markov chain. Differences in the evolution of the dynamic correlation values are obtained by contrasting the state dependent probabilities over the constant transition probabilities governing the regime switching process. In this way, we determine the impact these weakly exogenous variables have on the dynamic process.

During periods of increasing changes in price levels and rising volatilities, it is especially relevant to determine both the dynamic market linkages among related assets or markets and the evolution of the transmission of price changes between these related markets. With this model we determine the dynamic correlation values between multiple related markets, considering the particular effect that underlying fundamental variables have in the evolution of these markets. Capturing the impact of these underlying variables enables to determine the response effect that specific shocks on these variables may produce on the dynamic process of these series. The model is also applicable for managing risk through hedging and assists in increasing efficiency in the operation of related markets.

Literature Review

Several studies have been conducted regarding asymmetric price adjustments, including threshold behavior. A paper by Goodwin and Holt (1999) analyzing the dynamic relation and transmission of market prices among marketing channels in the beef sector, used a threshold error correction model accounting for both the non-stationary nature of each of these prices and the long-run stationary equilibrium or co-integration among these prices. The paper considered the asymmetric effects produced and found significance for three different regimes, i.e. threshold behavior. Additionally, and in response

to price shocks, lags were found in the adjustment period between each channel. A subsequent study by Goodwin and Harper (2000) for the pork sector arrived at similar results.

Earlier papers by Boyd and Brorsen (1988) - studying market channels in the pork sector, and Hanh (1990) - studying market channels both in the pork and beef sector, also found significant lags during the adjustment of price variations. These models considered different parameters for both lags and speed of transmission, yet did not account for the price's non-stationary nature mentioned above. Boyd and Brorsen (1988) found symmetric response to price changes, supporting later findings mentioned above. Yet Hanh (1990) found some asymmetric response to price changes within the different market channels.

Another result from Goodwin and Holt (1999) and Goodwin and Harper (2000) confirmed that price changes within market channels mainly propagated in one direction. i.e. response to price shocks were generally found to produce adjustments when these shocks were applied at the farm markets and from there the adjustments were passed on to the wholesale markets, and then to the retail markets. This result corroborated earlier findings by Boyd and Brorsen (1985) and Schroeder (1988). A paper by Bailey and Brorsen (1989) regarding three major cattle markets (i.e. feedlot operators and packers) in different states – found that there was asymmetric spatial adjustment for price variations. This was reflected in a difference in speed of adjustment for price changes – responding quicker to price increases than price decreases.

Goodwin and Piggott (2001) studied market integration in spatially separate regional grain markets within a state, through price linkages. They incorporated in their analysis thresholds that account for transaction costs, which delay price adjustments. Their results indicated that the markets are well integrated, and also confirm the existence of thresholds points for price adjustments. Once these thresholds are accounted for in the model, the speed of adjustment for price variations is higher than when they are not considered. For a study conducting an extensive survey regarding asymmetric price transmission, see Meyer and von Cramon-Taubedel (2004).

Regarding the study of hedge ratios - these are defined as the covariance between the futures and spot price divided by the variance of the future price, as per Benningna et al. (1984). It has been noted by Myers and Thompson (1989) that both variances and covariances are dynamic, changing through time. Hence this property must be taken into account for proper determination of hedge ratios. Time-varying optimal hedge ratios have been proposed using GARCH models. Many studies have estimated time-varying optimal hedge ratios for commodity futures through bivariate GARCH models, such as Baillie and Myers (1991), Myers (1991), Bera et al. (1996) and Lien et al. (2002).

Recently Lee and Yoder (2007a) incorporate regime switching following a Markov process, into a bivariate GARCH (BEKK) approach. This method considers the optimal hedge ratios being both time dependent and state dependent. Their model is a bivariate extension of Gray (1996), which is a univariate generalized regime switching model with conditional transition probabilities as function of conditional cumulative normal distributions.

Regime switching models that incorporate GARCH models have a path dependency problem. This occurs because for *N* different regimes (each regime denoted by Δ_t, and $\Delta_t = 1,2 \ldots N$), each conditional variance - $\sigma_t^2 = w(\Delta_t) + \alpha(\Delta_t)y_{t-1}^2 + \beta(\Delta_t)\sigma_{t-1}^2$ and conditional covariance - $H_t(\Delta_t)$, depend on the entire number or sequence of different regimes (Δ_t) being considered. This makes calculation of the likelihood practically intractable, since a sample of length T would require integration over all the possible N^T paths. To circumvent this problem, Gray assumes the conditional variance is function of a weighted sum of the previous conditional variances over all the possible regimes, therefore producing only one conditional variance for each period.

This method of combining the previous conditional variances of all possible regimes into a single one in a univariate setting is extended by Lee and Yoder (2007a) to a bivariate setting - in their regime switching bivariate BEKK model. i.e., they use a similar path dependency solution for both conditional variance and conditional covariance of the spot and futures prices of the series. However, by using the BEKK model from Engle and Kroner (1995), Lee and Yoder are prevented from potentially extending the application to multiproduct hedging of many related series, since the model incurs in the dimensionality problem mentioned previously and it cannot be estimated in two steps.

Our extension of the Regime Switching Dynamic Correlations model is similar to the original RSDC - being a mixture model made of a correlations (not variance) matrix along with an ARMACH type equation to model both the variances and covariances of multiple series. The dynamics of the different regimes considered through a Markov chain $(\Delta_t = 1,2 \ldots . N)$ are introduced directly in the correlations matrix, as will be seen in detail ahead. An advantage of our model over the previous BEKK model is that we may estimate the parameters of the model in two steps. In a first step, estimation of the models for the variances of our series can be performed series by series (or asset by asset), and these equations need not be dependent or function of each different regime (Δ_t). Subsequently, the calculations for correlations are made accounting for each different regime considered in the Markov chain.

In another paper, Lee and Yoder (2007b) develop a Markov regime switching frame as a generalization of Tse & Tsui (2002), which considers time varying correlations within a GARCH model for estimating optimal hedge ratios. The model shows hedging improvement over the CCC model and a time varying correlation GARCH model, previously developed by Lee et al. (2006). Yet the time varying correlations here do not consider specific underlying fundamental variables and the effect these may have in the evolution of the process.

Many studies have analyzed correlation coefficients of financial instruments during periods of low volatility versus high volatility. Longin & Solnik (1995) and Ramchmand & Susmel (1998) specifically use multivariate GARCH models to test for different correlations when in periods of changing volatility. The first case uses monthly data and considers a bivariate Constant Conditional Correlations for the base specification. During periods of high volatility, they insert an exogenous threshold at the contemporaneous value of the volatility, resulting in a different correlation during those volatile times. The second case uses weekly data (stock returns from Thursday to Thursday) and introduces correlations as function of variance regimes, using a Markov regime switching ARCH model (SWARCH). This model estimates the regime of volatility completely exogenously, having constant transition probabilities within regimes. Both studies find that for higher conditional volatilities, market correlations increase.

A study by Forbes & Rigobon (2002) finds that correlation coefficients among markets are conditional on market volatility, and this characteristic induces a conditioning bias in the estimates of correlations. For example, a bivariate normal distribution with an unconditional correlation value is different than a conditional correlation when markets are in upswing or downswing moves. i.e. the conditional correlations calculated are biased when these specific moves are not considered. A study by Ang & Chen (2002) takes this factor into account when comparing U.S. stocks versus the aggregate market, and develops a statistic for measuring, comparing and testing asymmetries. They find that regime-switching models have better performance at capturing correlation asymmetries. This result reaffirms our model selection and development. A recent study by Tejeda and Goodwin (2009) apply the restricted version of the model to estimate dynamic correlations between corn, soybeans and cattle markets, determining the effect of ethanol driven corn consumption among these markets. Results determined are consistent with the literature.

Econometric Model

I. RSDC model

The dynamic covariances between series are partitioned into standard deviations and correlations, with different correlation values switching between different regimes through a Markov chain.

If we consider a K - multivariate time process, i.e. K: number of time series:

$Y_t = H_t^{1/2} U_t$ with $U_t \sim i.i.d.(0, I_K)$ - where Y_t may be a filtered process. (1.)

The time varying covariance matrix H_t is decomposed into standard deviations and correlations:

$H_t \equiv S_t \Gamma_t S_t$ where S_t is a Diagonal matrix with standard deviations: $s_{k,t}$ $k = 1 \ldots\ldots K$ (2.)

Γ_t is the correlations matrix

(this decomposition of covariance matrix has been previously used by Bollerslev 1990, Tse & Tsui 2002, Engle 2002, and Pelletier 2006).

The standard deviations $s_{k,t}$ for each time series k - from the diagonal matrix S_t, are assumed to follow an ARMACH model. The correlation matrix Γ_t follows a Markov chain, with different values for different regimes, i.e. for particular t periods it may be in one regime with a certain set of correlation values, and for other t periods it may be in another regime, with a different set of correlation values.

I.a. Markov Process for Regime Switching between Correlations:

The RSDC model considers multiple series transitioning between regimes of different correlation levels according to a Markov chain process, generating dynamic correlations. The switch from one regime to another is determined by transition probabilities, which in this extended model are time-varying transitional probabilities. i.e. constant transitional probabilities are nested (special case) within our model.

The time-varying correlation matrix Γ_t is defined as:

$$\Gamma_t = \sum_{n=1}^{N} \mathbf{1}_{\{\Delta_t = n\}} \Gamma_n \quad (2.1)$$

where Δ_t is an unobserved Markov chain process independent of U_t, taking N possible regimes or values ($\Delta_t = 1,2, \ldots N$). And **1** is an indicator function.

The $K\ x\ K$ matrices Γ_n $(n = 1 \ldots \ldots N$: # regimes) are correlation matrices - symmetric, PSD, ones on the diagonal, off-diagonal elements between -1 and 1, with $\Gamma_n \neq \Gamma_{n'}$ for $n \neq n'$.

The 'probability law' governing the Markov chain process Δ_t is defined by its state dependent transition probability matrix: Π_t.

The probability of going from regime i in period $t-1$ to regime j in period t is denoted by $\pi_t^{i,j}$

The limiting probability of being in regime n is π^n, as per the ergodic property of a Markov chain.

Matrix Π_t has elements of row i and column j : $\pi_t^{i,j}$ i.e. $\pi_t^{i,j} = P(\Delta_t = j | \Delta_{t-1} = i)$

The Markov chain fits standard assumptions - aperiodic, irreducible and ergodic - Ross (1993), Chapter 4.

In equation (2.1) we impose Γ_n to be a correlation matrix, and then Γ_t will also be a correlation matrix. In Γ_n the diagonal elements are 1 and the off-diagonal elements are between $[-1, 1]$; yet this does not imply the matrix necessarily being PSD (could have extreme cases of tri-variate time series with off-diagonals at -.99, leading to non-PSD).

Hence the Choleski decomposition is used in helping to impose PSD for Γ_n.

i.e.: $\Gamma_n = P_n P_n'$ with P_n as the lower triangular matrix, and restricting its first diagonal value equal to 1. This secures off-diagonal elements between [-1, 1].

Despite an increasing number of parameters coming from each Γ_n that are to be estimated, the Expectation maximization (EM) algorithm may be used from Dempster et al. (1977), as presented in Diebold et al. (1994) and Hamilton (1994 – Chapter 22), and thus handling this matter effectively. This will be duly addressed ahead in the estimation section. Also, as mentioned before, this model is linear due to the Markov chain, and therefore able to estimate directly multi-step ahead conditional expectations of the correlation matrix. This is in contrast to the DCC model, which introduces non-linearities by rescaling the covariances to obtain correlations between -1 and 1; hence direct conditional expectation computation is not possible.

Another relevant factor is that if persistence exists in the Markov chain (staying in the same regime in subsequent time periods), it will lead to smoother time-varying correlations. This could have an influence for VaR computations and dynamic portfolio allocation, since the benefits of portfolio diversification would be less volatile, as per Pelletier (2006).

There have previously been numerous regime switching models for univariate heteroskedastic time series, including Garcia and Perron (1996), Gray (1996), Dueker (1997). In addition, some of these models have been extended to consider bivariate time series, such as previously mentioned Lee and Yoder (2007a). In general, these models consider GARCH equations which must also incorporate regime switching parameters for the variances, as well as 'normalization' of the covariances to become correlations.

Our model doesn't need parameters for the correlations of the same variables, as all these correlations must be equal to one. At the same time, the model directly estimates the correlations without need to normalize, this in contrast to the case of models that estimate covariances.

I.b. Restricted Model

The parsimonious or restricted model for the time-varying correlation matrix Γ_t is similar to Pelletier (2006). That is:

$$\Gamma_t = \Gamma\lambda(\Delta_t) + I_K\big(1 - \lambda(\Delta_t)\big) \qquad (2.2)$$

with:

Γ being a fixed KxK correlation matrix – for every state or regime considered.

I_K is a KxK identity matrix.

$\lambda(\Delta_t) \epsilon$ [0,1] (for assurance of eliminating possibilities of non-PSD correlation matrix) is a univariate random process governed by an unobserved Markov chain process Δ_t that takes N possible values $(\Delta_t = 1,2 \ldots N)$, and is independent of U_t.

The 'probability law' governing the Markov chain process Δ_t is defined by its state dependent transition probability matrix: Π_t which can be a function of either weakly exogenous variables, or exogenous indices, or a mix between them as will be noted. These state dependent (time varying) transition probabilities serve to determine the impact of underlying related factors (prices and/or indexes) in the change of these dynamic correlations. That is, by the use of variables such as associated prices and/or indexes in these time varying transition probabilities, we are able to assess their impact on the switch between one state of dynamic correlation and another.

The correlation matrix at time t (i.e. Γ_t) is a weighted average of two extreme states or regimes – uncorrelated returns by $\lambda(\Delta_t) = 0$, or totally correlated returns at $\lambda(\Delta_t) = 1$. Changes among correlations of different regimes are strictly proportional to $\lambda(\Delta_t)$, allowing for regimes of higher or lower correlations since the diagonals (own-correlations) are left at one.

This model enables calculation of dynamic correlations in a context of time-varying transition probabilities, without the need for expectation maximization since less number of parameters are to be estimated. In other words, through maximum likelihood and using a correlation targeting method described next, we are able to estimate dynamic correlations between regimes when considering state dependent transition probabilities.

Because in (2.2) only the product of Γ and λ is identifiable for off-diagonal constraints (diagonal elements are equal to 1), we use the same correlation targeting method followed by Pelletier (2006). That is by using either of the following sets of constraints during our estimation:

i. $\lambda(1) = 1, \lambda(1) > \lambda(2) \dots \lambda(N-1) > \lambda(N)$ (2.2.1)

which fixes the first λ value at its highest i.e.1, obtaining directly Γ from $\Gamma\lambda$; also the rest of $\lambda's$ are in diminishing order to eliminate the possibility of relabeling regime i as j and viceversa.

ii. $\max_{i \neq j} |\Gamma_{ij}| = 1$; with $1 > \lambda(1) > \lambda(2) \dots, \lambda(N-1) > \lambda(N)$ (2.2.2)

which fixes an off-diagonal element at 1; but it does not mean the correlation may be 1 (or -1), since the correlation is actually the product given by $\Gamma_{ij}\lambda(\Delta_t)$, and $\lambda(1) < 1$.

I.c. Univariate Model

The time-varying standard deviations are modeled directly by using the ARMACH process as per Taylor (1986). Since covariances are the product of correlations and standard deviations, here there is no need to model the variance through GARCH, and then using a non-linear square root to obtain the standard deviation.

In ARMACH, the conditional standard deviation follows:

$$s_t = \omega + \sum_{i=1}^{q} \tilde{\alpha}_i \, |y_{t-i}| + \sum_{j=1}^{p} \beta_j \, s_{t-j} \quad \text{with } \tilde{\alpha}_i = \alpha_i / E|\tilde{u}_t|, \text{ for stationarity purposes} \qquad (2.3)$$

As mentioned previously since the expectation enters as a linear operator, it provides ease of use for multi-step ahead computation of conditional expectations.

II. State Dependent, Time-varying Probabilities

In order to introduce state dependent, time-varying transition probabilities within regimes into the dynamic correlations model, we use a procedure proposed by Diebold et al. (1994). As noted, there are two regime switching models that will be extended. The first one is the parsimonious (restricted) model (2.2), where the Correlation matrix Γ_t is obtained by correlation targeting using equation (2.2.2), though we would arrive at similar results if using equation (2.2.1). In this case parameter estimation is obtained through maximum likelihood, and we need to obtain previously the filtered/smoothed probabilities because of the unobserved Markov chain. That is, we require the filtered and the subsequent smoothed probabilities in order to estimate our dynamic correlations. On the other hand, for the general regime switching model (2.1), and its estimation via maximization of parameters, we apply the EM maximization procedure.

The state dependent probability matrix Π_t has elements of row i and column j:

$$\pi_t^{i,j} = P(\Delta_t = j | \Delta_{t-1} = i).$$

The probability of transitioning from regime i at period $t-1$, to regime j at period t is: $\pi_t^{i,j}$

Such that:

$$\pi_t^{i,i} = P(\Delta_t = i \mid \Delta_{t-1} = i\ , x_{t-1}; \beta_i) = \frac{\exp(x'_{t-1}\beta_i)}{1+\exp(x'_{t-1}\beta_i)}\ ; \quad \text{and}$$

$$\pi_t^{i,j} = P(\Delta_t = j \mid \Delta_{t-1} = i\ , x_{t-1}; \beta_i) = \quad 1 - \frac{\exp(x'_{t-1}\beta_i)}{1+\exp(x'_{t-1}\beta_i)}$$

conversely,

$$\pi_t^{j,j} = P(\Delta_t = j \ \mid \Delta_{t-1} = j\ , x_{t-1}; \beta_j) = \frac{exp(x'_{t-1}\beta_j)}{1+exp(x'_{t-1}\beta_j)}\ ; \quad \text{and}$$

$$\pi_t^{j,i} = P(\Delta_t = i \mid \Delta_{t-1} = j\ , x_{t-1}; \beta_j) = \quad 1 - \frac{exp(x'_{t-1}\beta_j)}{1+exp(x'_{t-1}\beta_j)}$$

The variables x_{t-1} are weakly exogenous (or fully exogenous) variables.

For the specific case of two regimes:

i.e. $\Delta_t = 1\ \boldsymbol{or}\ \Delta_t = 2$

$$P(\Delta_t = 1\ \boldsymbol{or}\ \Delta_t = 2 \mid \Delta_{t-1} = 1\ , x_{t-1}; \beta_1) = \frac{\exp(x'_{t-1}\beta_1)}{1+\exp(x'_{t-1}\beta_1)} \quad \boldsymbol{or} \quad 1 - \frac{\exp(x'_{t-1}\beta_1)}{1+\exp(x'_{t-1}\beta_1)} \quad \text{and}$$

$$P(\Delta_t = 1\ \boldsymbol{or}\ \Delta_t = 2 \mid \Delta_{t-1} = 2\ , x_{t-1}; \beta_2) = \frac{exp(x'_{t-1}\beta_2)}{1+exp(x'_{t-1}\beta_2)} \quad \boldsymbol{or} \quad 1 - \frac{exp(x'_{t-1}\beta_2)}{1+exp(x'_{t-1}\beta_2)}$$

The transition probability matrix Π_t: (3.1)

		Time t	
	State 1		**State 2**
State 1	π_t^{11} $P(\Delta_t = 1 \mid \Delta_{t-1} = 1, x_{t-1}; \gamma_1)$ $= \frac{\exp(x'_{t-1}\gamma_1)}{1+\exp(x'_{t-1}\gamma_1)}$		$\pi_t^{12} = (1 - \pi_t^{11})$ $P(\Delta_t = 2 \mid \Delta_{t-1} = 1, x_{t-1}; \gamma_1)$ $= 1 - \frac{\exp(x'_{t-1}\gamma_1)}{1+\exp(x'_{t-1}\gamma_1)}$
Time t-1			
State 2	$\pi_t^{21} = (1 - \pi_t^{22})$ $P(\Delta_t = 1 \mid \Delta_{t-1} = 2, x_{t-1}; \gamma_2)$ $= 1 - \frac{\exp(x'_{t-1}\gamma_2)}{1+\exp(x'_{t-1}\gamma_2)}$		π_t^{22} $P(\Delta_t = 2 \mid \Delta_{t-1} = 2, x_{t-1}; \gamma_2)$ $= \frac{\exp(x'_{t-1}\gamma_2)}{1+\exp(x'_{t-1}\gamma_2)}$

where $x_{t-1} = (1, x_{1,t-1}, \ldots.., x_{(m-1),t-1})'$ & $\gamma_i{}^1 = (\gamma_{i1}, \gamma_{i2}, \ldots\ldots\ldots, \gamma_{i(m-1)})$;

III. Estimation:

From equations (1) and (2), the log-likelihood can be written as:

$$L = -\frac{1}{2}\sum_{t=1}^{T}[Klog(2\pi) + \log(|H_t|) + Y_t' H_t^{-1} Y_t]$$

$$= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}[Klog(2\pi) + \log(|S_t \Gamma_t S_t|) + Y_t' S_t^{-1} \Gamma_t^{-1} S_t^{-1} Y_t]$$

$$L = -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}[Klog(2\pi) + 2\log(|S_t|) + \log(|\Gamma_t|) + \widetilde{U}_t' \Gamma_t^{-1} \widetilde{U}_t] \quad (4.1.)$$

where $\widetilde{U}_t = S_t^{-1} Y_t$ and $\widetilde{U}_t = [\tilde{u}_{1,t} \ldots\ldots\ldots\ldots \tilde{u}_{K,t}]'$ is a zero mean process with covariance matrix Γ_t; also $|H_t| = \det(H_t)$.

[1] Here we use $\gamma_i$same as is if it was β_i of previous page.

Estimation of the model parameters is made in two steps, with the assurance that the variance/covariance matrix is PSD (positive semi-definite). First the standard deviations are obtained, and then the correlations are calculated. Standard deviations are obtained directly via the ARMACH model in (2.3), permitting a smooth linear calculation for correlations versus the use of a GARCH type model, which introduces nonlinearities when going from covariances to correlations since it requires the square root of the variances. This arrangement also enables the possibility of calculating analytically multi-step ahead conditional expectations, due to its linear properties; yet using a GARCH type model does not permit this.

The estimation method calculates both filtered and smoothed transition probabilities between regimes, and subsequently makes use of an Expectation Maximization (EM) algorithm for the unrestricted model. i.e. the model estimation involves two main parts. The first part is the expectation step, which estimates the expectation of the complete data log-likelihood. This involves calculating the filtered probabilities for the complete data log-likelihood conditional on data observed, and then obtaining back the smoothed probabilities. The second part is the maximization step, which considers the use of these smoothed probabilities in our expected complete-data log likelihood function and maximizes directly with respect to the parameters.

When there are a large number of parameters being estimated we can use a two step estimation procedure as in Engle (2002) and Pelletier (2006). In the first step, univariate volatility models are estimated independently for each asset. In the second step, the correlation matrix is estimated conditional on the first step estimates. This procedure helps in avoiding the dimensionality curse. However, if the number of parameters is not too great, then only a one step procedure may be necessary for the estimation of all the parameters. In our empirical application, we will use the two-step estimation procedure. Both procedures are described below.

III. a. **One-step estimate**:

For maximum likelihood, we evaluate:

$$QL(\theta; Y) = \sum_{t=1}^{T} \log\ f(Y_t | \overrightarrow{Y_{t-1}}) \qquad (4.1.1)$$

with $\overrightarrow{Y_{t-1}} = \{Y_{t-1}, Y_{t-2}, \dots\}$ and θ is the vector of unknown parameters.

This equation comes from (4.1.).

We use a Diebold et al. (1994) procedure that considers a filter for time-varying probabilities - creating smoothed probabilities, and then estimating through maximum likelihood for the restricted model and through expected maximization for the general model. We will present some general steps for this procedure, with full details of the process in Appendix 1.

To simplify, we let $\{\Delta_t\}_{t=1}^{T}$ be the sample path of a first order, two state Markov chain process i.e. $\Delta_t = 1$ or $\Delta_t = 2$, with time varying transition probabilities according to the matrix Π_t (3.1) above. In general, the matrix Π_t considers a x_{t-1} vector ($mx1$) for the underlying related variable(s), i.e. a set of weakly exogenous variables that affect the state dependent transition probabilities. Therefore parameters γ_i with $i = 1, 2$ constitute the two regimes i.e. γ_i is a $(2mx1)$ vector such that $\gamma = (\gamma_1', \gamma_2')'$, and the constant transition probabilities are nested within these state dependent probabilities. i.e. the first element of $mx1$ is a constant.

In our first and unrestricted model (2.1), we may estimate the case of two regimes considering state dependent transition probabilities[2], by:

$$\hat{\Gamma}_t = \hat{\Gamma}_1 * P(\Delta_t = 1 | x_{t-1}; \theta) + \hat{\Gamma}_2 * P(\Delta_t = 2 | x_{t-1}; \theta) \tag{4.1.2}$$

In our restricted or parsimonious model from (2.2), we estimate the case of our two regimes dynamic correlations by:

$$\hat{\Gamma}_t = \hat{\Gamma} * \hat{\lambda}(1) * P(\Delta_t = 1 | x_{t-1}; \theta) + \hat{\Gamma} * \hat{\lambda}(2) * P(\Delta_t = 2 | x_{t-1}; \theta) \tag{4.1.3}$$

Let $\{y_t\}_{t=1}^{T}$ be a time series that evolves according to the Markov chain $\{\Delta_t\}_{t=1}^{T}$. In our case, the series are previously standardized according to their volatility obtained through the ARMACH model (2.3). Then we have:

$$(y_t | \Delta_t = i; \alpha_i) \overset{iid}{\sim} \mathrm{N}(\mu_i, \sigma_i^2), \text{ with } \alpha_i = (\mu_i, \sigma_i^2)' \text{ and } i = 1, 2$$

with the conditional density given by:

$$f(y_t | \Delta_t = i; \alpha_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(\frac{-(y_t - \mu_i)^2}{2\sigma_i^2}\right) \quad \text{with } \alpha = (\alpha_1', \alpha_2')' \tag{4.2}$$

[2] Here $P(\Delta_t = 1 | x_{t-1}; \theta) = P(\Delta_t = 1, \Delta_t = 1 | x_{t-1}; \theta) + P(\Delta_t = 1, \Delta_t = 2 | x_{t-1}; \theta)$ - further details at end of Appendix 1.2.

The restricted or parsimonious model enables calculation of dynamic correlations within a time-varying transition probabilities context by maximum likelihood. In other words, through the use of correlation targeting described below, we are able to estimate dynamic correlations between regimes when considering state dependent transition probabilities.

For the general model (2.1), the Expectation Maximization (EM) algorithm is a robust and stable method to maximize the incomplete data log-likelihood, via iterative maximization of the expected complete-data log likelihood, conditional upon the observed data set. For the long-run probability of the first (starting) state at $t = 1$, being at state 1, we use $P(\Delta_1 = 1) = \rho$.

In the case of our restricted model, instead of having to estimate this extra parameter ρ, we proceed in similar way to Pelletier (2006) and consider the limiting probabilities for the Markov process: π^1, π^2 and solve for:

$\binom{\pi^1}{\pi^2} = \Pi'_t \binom{\pi^1}{\pi^2}$; $\pi^1 + \pi^2 = 1$ - as per Ross (1993) - Ch.4.4

Such that $\hat{\pi}^1 = \hat{\pi}_t^{21} \big/ (\hat{\pi}_t^{12} + \hat{\pi}_t^{21})$ and $\hat{\pi}^2 = \hat{\pi}_t^{12} \big/ (\hat{\pi}_t^{12} + \hat{\pi}_t^{21})$

Now the complete-data likelihood – assuming all regimes Δ_t are observed, and letting $\theta = (\alpha', \gamma', \rho)'$ (excluding ρ in the case of the restricted model, as mentioned previously), would be:

$$f\left(\vec{y}_T, \vec{\Delta}_T \middle| \vec{x}_T;\ \theta\right) = f(y_1, \Delta_1 | \vec{x}_T;\ \theta) \prod_{t=2}^{T} f(\, y_t, \Delta_t |\, \vec{y}_{t-1}, \vec{\Delta}_{t-1}, \vec{x}_T; \theta)$$

$$= f(y_1 | \Delta_1, \vec{x}_T;\ \theta) P(\Delta_1) \textstyle\prod_{t=2}^{T} f(y_t | \Delta_t, \vec{y}_{t-1}, \vec{\Delta}_{t-1}, \vec{x}_T; \theta) P(\Delta_t | \vec{y}_{t-1}, \vec{\Delta}_{t-1}, \vec{x}_t; \theta)$$

$$= f(y_1 | \Delta_1;\ \alpha) P(\Delta_1) \textstyle\prod_{t=2}^{T} f(y_t | \Delta_t;\ \alpha) P(\Delta_t |\, \Delta_{t-1}, x_{t-1}; \beta)$$

Introducing indicator functions for convenience:

$$f\left(\vec{y}_T, \vec{\Delta}_T \middle| \vec{x}_T;\ \theta\right) = [\mathrm{I}(\Delta_1 = 1) f(\mathrm{y}_1 | \Delta_1 = 1;\ \alpha_1)\rho + \mathrm{I}(\Delta_1 = 0) f(\mathrm{y}_1 | \Delta_1 = 0;\ \alpha_0)(1 - \rho)] \times$$

$$\times \textstyle\prod_{t=2}^{T} \{\, \mathrm{I}(\Delta_t = 1, \Delta_{t-1} = 1) f(\mathrm{y}_t | \Delta_t = 1;\ \alpha_1) \pi_t^{11} +$$

$$+\ \mathrm{I}(\Delta_t = 0, \Delta_{t-1} = 1) f(\mathrm{y}_t | \Delta_t = 0;\ \alpha_0)(1 - \pi_t^{11}) +$$

$$+\ \mathrm{I}(\Delta_t = 1, \Delta_{t-1} = 0) f(\mathrm{y}_t | \Delta_t = 1;\ \alpha_1)(1 - \pi_t^{00}) +$$

$$+ \quad \mathrm{I}(\Delta_t = 0, \Delta_{t-1} = 0) f(\mathrm{y}_t | \Delta_t = 0;\ \alpha_0) \pi_t^{00} \}$$

Taking logs:

$$\log f\left((\vec{y}_T, \vec{\Delta}_T | \vec{x}_T;\ \theta\right) = \mathrm{I}(\Delta_1 = 1)[\log f(y_1 | \Delta_1 = 1;\ \propto_1) + \log \rho] +$$

$$+ \mathrm{I}(\Delta_1 = 0)[\log f(y_1 | \Delta_1 = 0;\ \propto_1) + \log(1 - \rho)] +$$

$$+ \sum_{t=2}^{T} \{ \mathrm{I}(\Delta_t = 1) \log f(y_t | \Delta_t = 1; \alpha_1) + \mathrm{I}(\Delta_t = 0) \log f((y_t | \Delta_t = 0; \alpha_0) +$$

$$+ \mathrm{I}(\Delta_t = 1, \Delta_{t-1} = 1) \log(\pi_t^{11}) + \mathrm{I}(\Delta_t = 0, \Delta_{t-1} = 1) \log(1 - \pi_t^{11}) +$$

$$+ \mathrm{I}(\Delta_t = 1, \Delta_{t-1} = 0) \log(1 - \pi_t^{00}) + \mathrm{I}(\Delta_t = 0, \Delta_{t-1} = 0) \log(\pi_t^{00}) \qquad (4.3)$$

Therefore a basic algorithm procedure for parameter estimation consists of the following sequence. For $\theta = (\alpha', \gamma', \rho)'$ - excluding ρ in the case of the restricted model, as mentioned previously.

1. Pick $\theta^{(0)}$

2. Obtain filtered and smoothed probabilities for the following – see Appendix 1.2 for equation details regarding the marginal and joint probabilities[3], conditional on $\theta^{(0)}$:

$$P(\Delta_t = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})^3 \quad \forall_t,\ P(\Delta_t = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})^3 \quad \forall_t,$$

$$P(\Delta_t = 1, \Delta_{t-1} = 1 |\ \vec{y}_T, \vec{x}_T;\ \theta^{(0)})^3 \quad \forall_t, \quad P(\Delta_t = 0, \Delta_{t-1} = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})^3 \quad \forall_t$$

$$P(\Delta_t = 1, \Delta_{t-1} = 0 |\ \vec{y}_T, \vec{x}_T;\ \theta^{(0)})^3 \quad \forall_t, \quad P(\Delta_t = 0, \Delta_{t-1} = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})^3 \quad \forall_t$$

3. Construct: E $\log f\left(\vec{y}_t, \vec{\Delta}_t \middle|\ \vec{x}_t;\ \theta^{(0)}\right)$ - the hypothetical (i.e. assuming all $\{y_t\}$ and $\{\Delta_t\}$ are observed) complete data likelihood (see Appendix 1.1 below) by replacing the indicator functions (I's) from (4.3) with smoothed probabilities (*P*'s) obtained in previous step 2, obtaining the following:

$$\mathrm{E}[\log f\left(\vec{y}_t, \vec{\Delta}_t \middle|\ \vec{x}_t;\ \theta^{(0)}\right) = \rho^{(0)}[\log f\left(y_1 \middle| \Delta_1 = 1;\ \alpha_1^{(0)}\right) + \log \rho^{(0)}] +$$

$$(1 - \rho^{(0)})[\log f\left(y_1 \middle| \Delta_1 = 0;\ \alpha_0^{(0)}\right) + \log(1 - \rho^{(0)}) +$$

$$\sum_{t=2}^{T} \{ \mathrm{P}(\Delta_t = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})[\log f\left(y_t \middle| \Delta_t = 1;\ \alpha_1^{(0)}\right) +$$

$$\mathrm{P}(\Delta_t = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})[\log f\left(y_t \middle| \Delta_t = 0;\ \alpha_0^{(0)}\right) +$$

$$\mathrm{P}(\Delta_t = 1, \Delta_{t-1} = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})\log(\pi_t^{11}) +$$

$$\mathrm{P}(\Delta_t = 0, \Delta_{t-1} = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})\ \log(1 - \pi_t^{11}) +$$

$$+ \mathrm{P}(\Delta_t = 1, \Delta_{t-1} = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})\log(1 - \pi_t^{00}) +$$

$$\mathrm{P}(\Delta_t = 0, \Delta_{t-1} = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(0)})\ \log(\pi_t^{00})\ \}$$

4. Set $\theta^{(1)} = \arg\max_{\theta} \mathrm{E}[\log f((\vec{y}_t, \vec{\Delta}_t \middle| \vec{x}_t;\ \theta^{(0)})]$

5. Iterate to convergence.

This last convergence criterion may be obtained by assuming a change of likelihood from one iteration to the next, or for a change in the gradient vector, or for $\left|\left|\theta^{(j)} - \theta^{(j-1)}\right|\right|$ that is smaller than a certain minimum value.

For the correlation targeting method used in the restricted model, we just obtain the smoothed probabilities as per previous step 2. i.e., once these smoothed probabilities are estimated, they may be used directly for estimation of the parameters in the restricted model (2.2), via maximum likelihood.

III. b. Threshold levels

Persistence of the dynamic correlations in the Markov chain may vary as a function of the weakly exogenous variables. i.e. these variables or underlying related factors of the multiple series may produce persistence at a specific regime's correlation value along the Markov chain. Thus a threshold is identified when it maintains the series in a particular regime instead of switching to a different regime and correlation value, had the related factor not been taken into account. This type of threshold may appear for example as an effect or consequence of the steady increasing or decreasing level of underlying related market factors such as prices, price ratios or changes in prices – which are explicitly included as a weakly exogenous variable.

The scenario may be the case of continuous rising or decreasing price levels in certain commodities responding to market fundamentals, thus producing higher correlation levels among related markets. On the other hand, there may be the case of continuous decreasing prices in financial assets, leading to anticipated increases in correlation levels among markets, a situation in accordance with the literature.

Conversely, if the related variable favors the switch from one regime to another regime in contrast to the case where the factor had not been considered, i.e. in this latter case the regime switch occurs due to unaccounted exogenous factors, then the resulting correlation level may be more prevalent when the related variable is taken into account. Then this related factor is the opposite of a threshold and may be considered a catalyst. This case where the underlying factor prompts a regime switch, producing a larger switch than if the factor had not previously considered, may respond to explicitly considering the volatility of related market factor(s) which had previously not been directly considered.

The threshold levels are a function of weakly exogenous variables x_{t-1} and their coefficients β_i (or γ_i) for the different regimes considered. The case of two regimes in our dynamic correlations model and for simplicity we consider only one weakly exogenous variable besides the constant factor, results in the following coefficient(s) β_1 and β_2 in the previous transition probabilities:

Coefficients - b_{11} for the constant and $\boldsymbol{b_{12}}$ for the weakly exogenous variable - at Regime ***1***,
Coefficients - b_{21} for the constant and $\boldsymbol{b_{22}}$ for the weakly exogenous variable – at Regime ***2***.
The nested case of constant transition probabilities considers $\boldsymbol{b_{12}} = 0$ and $\boldsymbol{b_{22}} = 0$.
To assess the impact of a significant coefficient of a weakly exogenous variable i.e. what we consider a threshold level, we do a first order Taylor approximation for this probability at a small value around our weakly exogenous variable x_{t-1} valued at zero.

For example at $t-1$ and being at regime 1, i.e. $\Delta_{t-1} = 1$, and remaining at regime 1 for the next period, i.e. $\Delta_t = 1$, and for a small value of x_{t-1} around zero:

$$P(\Delta_t = 1 \mid \Delta_{t-1} = 1, x_{t-1}; \beta_0) =$$

$$= P(\Delta_t = 1 \mid \Delta_{t-1} = 1, x_{t-1}; \beta_0)|_{x_{t-1}=0} + \frac{\partial P(\%)}{\partial x}|_{x_{t-1}=0} * (x_{t-1} - 0)$$

$$= P(\Delta_t = 1 \mid \Delta_{t-1} = 1; \beta_0) + \boldsymbol{b}_{12} \frac{\exp(x'_{t-1}\beta_0)}{[1+\exp(x'_{t-1}\beta_0)]^2} |_{x_{t-1}=0} * (x_{t-1} - 0)$$

results in:

$$P(\Delta_t = 1 \mid \Delta_{t-1} = 1\,, x_{t-1}; \beta_0) - P(\Delta_t = 1 \mid \Delta_{t-1} = 1\,; \beta_0) = \boldsymbol{b}_{12} \frac{\exp(\mathrm{b}_{11})}{[1+\exp(\mathrm{b}_{11})]^2} * x_{t-1}$$

Or

$$\Delta\, P(\Delta_t = 1 \mid \Delta_{t-1} = 1\,, \boldsymbol{x_{t-1}}; \beta_0) = \boldsymbol{b}_{12} \frac{\exp(\mathrm{b}_{11})}{[1+\exp(\mathrm{b}_{11})]^2} * \boldsymbol{x_{t-1}} \qquad (4.3)$$

That is, a small change in the probability of remaining in regime 1 (spillover effect), resulting from a small change in the weakly exogenous variable $\boldsymbol{x_{t-1}}$**,** is equal to the product of three terms. These terms are - the coefficient $(\boldsymbol{b_{12}})$ of the weakly exogenous variable, the constant coefficient $(\boldsymbol{b_{11}})$ being a function of its exponential and the squared inverse of one plus its exponential, and the small change of this variable $\boldsymbol{x_{t-1}}$. There are three basic cases which may occur regarding a small change in the weakly exogenous variable $\boldsymbol{x_{t-1}}$**.**

If the coefficient $\boldsymbol{b_{12}}$ of our weakly exogenous variable is insignificant, then this underlying related variable would not form a threshold for price variations among markets. i.e. changes in this weakly exogenous variable would not make a difference in the evolution of our correlation values, and these market correlations would evolve completely exogenous to this particular factor. A different situation takes place if this coefficient $\boldsymbol{b_{12}}$ is significant and positive, where two cases emerge. One case occurs when the product of $\boldsymbol{b_{12}}$ with the constant coefficient's function $\boldsymbol{b_{11}}$ of exponentials is large, then positive variations of our weakly exogenous variable $(\boldsymbol{x_{t-1}})$ will lead to a higher probability of remaining at regime 1 (i.e. longer spillover effect). The other case results in no further effect from positive variations of our related variable had this former product been zero or small. In other words for this second case - the coefficient $\boldsymbol{b_{12}}$ being positive and significant determines an existing threshold from this related variable. Yet the effect of increases in this weakly exogenous variable are dampened because the terms multiplied to this variable, specifically the product of $\boldsymbol{b_{12}}$ and exponential functions of $\boldsymbol{b_{11}}$ are small or negligible. i.e. the threshold identified may produce spillover effects, yet additional spillover effects from increases in the related variable coming from shocks would not be produced.

Conversely, if the coefficient $\boldsymbol{b_{12}}$ is significant and negative at (4.3), and its product with the constant coefficient's $\boldsymbol{b_{11}}$ function of exponentials is large, then positive variations of our weakly exogenous variable $(\boldsymbol{x_{t-1}})$ will lead to a lower probability of remaining at regime 1 or higher probability of switching to a different regime, in this case switching to regime 2. This is in comparison to the case of the product of the two former factors being negative yet negligible. In other words, positive shocks from our

related variable will increase the probability of regime switching resulting in a larger correlation level at the new regime, versus the case where positive shocks from the related variables have no effect. This latter case is a result of the terms that are multiplied to the related variable, i.e. $\boldsymbol{b_{12}}$ and exponential functions of $\boldsymbol{b_{12}}$ being once again very small or negligible.

IV. **Two-step estimate:**

Once again we make use of a similar procedure in Pelletier (2006), yet make the necessary modifications involved for considering state dependent transition probabilities instead of having constant transition probabilities. The parameter space θ is partitioned into θ_1- parameters from the univariate volatility model for each time series (i.e. ARMACH model), and θ_2- parameters from the correlation model.

The first step from (4.1), the likelihood assumes the correlation matrix is an identity matrix (i.e. $\Gamma_t = \mathrm{I}_t$):

$$QL_1(\theta_1; Y) = -\frac{1}{2}\sum_{t=1}^{T}(Klog(2\pi) + 2\log(|S_t|) + U'_t U_t)$$

This does not require the use of a filter since the series' volatility is not governed by a Markov chain, and the parameters for each series are estimated separately.

For the second step, the likelihood from (4.1) is estimated given θ_1 and taking out S_t previously estimated, which becomes:

$$QL_2(\theta_2; Y, \theta_1) = -\frac{1}{2}\sum_{t=1}^{T}(Klog(2\pi) + \log(|\Gamma_t|) + U'_t \Gamma_T^{-1} U_t)$$

This second estimation does require the use of filtering because Δ_t (Markov chain) is not observed. This filtering has been obtained previously from step 2 for the one-step estimation case mentioned before.

When considering the non-restricted model (2.1), estimation is done through Expectation Maximization (EM) following steps 3. and 4. from the algorithm detailed previously in the one-step estimation. Iteration of the maximization process is continued until new vectors $\theta^{(j)}$ computed have a difference with subsequent vectors that becomes arbitrarily small.

As mentioned before, the case of the restricted model (2.2) considers correlation targeting. This involves first the calculation of unconditional expectation of the correlation matrix. i.e.

$$E[\Gamma_t] = \Gamma \sum_{n=1}^{N} \lambda(n)\pi^n + I_K \sum_{n=1}^{N} (1-\lambda(n))\,\pi^n$$

The off-diagonal terms are part of the first term, i.e. matrix Γ multiplied by the scalar $\sum_{n=1}^{N} \lambda(n)\pi^n$. This means that an estimation of Γ $\left(\hat{\Gamma}\right)$, computed with the standardized residuals from the first step, will have off-diagonal elements 'scaled' by $\sum_{n=1}^{N} \lambda(n)\pi^n$; thus these will be re-scaled according to the natural constraint posed before in (2.2.2). In other words, divide the off-diagonal elements of $\hat{\Gamma}$ by the highest absolute value among them, to obtain -1 or 1. This secures $\lambda(1) < 1$ and the number of parameters (λ) to be non-linearly estimated increases with the number of different regimes, and not with the number of time series.

In our case of two regimes, we have 1 > $\lambda(1) > \lambda(2)$, therefore:

$$E[\Gamma_t] = \Gamma(\lambda(1)\pi^1 + \lambda(2)\pi^2) + I_k[\left(1-\lambda(1)\right)\pi^1 + \left(1-\lambda(2)\right)\pi^2]$$

These two-step estimates are consistent, and their asymptotic distribution follows a Normal (0, *V*), i.e.:

$$\sqrt{T}\left(\begin{bmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{bmatrix} - \begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}\right) \rightarrow N(0;V) \qquad \text{with}$$

$$V = \begin{bmatrix} G_{\theta_1}^{-1} & -G_{\theta_1}^{-1}G_{\theta_2}M^{-1} \\ 0 & M^{-1} \end{bmatrix} E\left[\frac{\partial lnf}{\partial \theta}\frac{\partial lnf}{\partial \theta'}\right] \begin{bmatrix} G_{\theta_1}^{-1} & -G_{\theta_1}^{-1}G_{\theta_2}M^{-1} \\ 0 & M^{-1} \end{bmatrix}'$$

such that:

$$G_{\theta_1} = E\left[\frac{\partial g(Y,\boldsymbol{\theta_1},\boldsymbol{\theta_2})}{\partial \boldsymbol{\theta_1'}}\right],\ G_{\theta_2} = E\left[\frac{\partial g(Y,\boldsymbol{\theta_1},\boldsymbol{\theta_2})}{\partial \boldsymbol{\theta_2'}}\right],\quad M = E\left[\frac{\partial m(Y,\boldsymbol{\theta_2})}{\partial \boldsymbol{\theta_2'}}\right],$$

$$g(Y,\theta_1,\theta_2) = \frac{\partial lnf(Y_t|\overrightarrow{Y_{t-1}})}{\partial \boldsymbol{\theta_1}},\ \ m(Y,\theta_2) = \frac{\partial lnf(Y_t|\overrightarrow{Y_{t-1}})}{\partial \boldsymbol{\theta_2}}$$

The proof is in Pelletier (2006).

IV. Simulation

We simulated four series, considering a Regime 1 level of correlation (Γ) of 0.85 and a Lambda proportion (λ) of 0.3. Hence the Regime 2 level of correlations ($\Gamma * \lambda$) is 0.255. Results of the simulated data from our restricted model - in Appendix_2, indicate that we have obtained good approximations of the original 'true' parameters. That is, our restricted state dependent model is providing good estimates of the simulated data. The coefficients estimated for the correlations, including the parameters in our state dependent transition probabilities and for the lambda proportion are significantly equal to the original 'true' values which simulated the series of data.

We also include results of the model estimated by considering constant transition probabilities. This was done by making the coefficients of the weakly exogenous variable equal to zero. In this case, we obtain good estimates of the parameters for correlation and lambda proportion - being significantly equal to the original 'true' values. However, we do not obtain these good results for the coefficients of the constants in our transition probabilities. In other words, these parameters are not significantly equal to the original ones. Perhaps this should be the case, since the switching probabilities are now not depending on any related variable, but just the simulated series themselves. Nonetheless, when we compare these models by likelihood, we obtain better results with our state dependent transition probabilities (less likelihood preferred, yet we also include two extra parameters being estimated), than for constant transition probabilities. Charts showing the correlation series and results of different number of simulations considering state dependent transition probabilities and constant transition probabilities are in Appendix_3. In addition, direct comparison between the state dependent transition probabilities and constant transition probabilities along these simulated series are presented.

Discussion

Through the use of multiple series of simulated data, our model was able to capture the effect or impact of underlying variables specifically related to the evolution of the multiple series. The simulated data was particularly constructed with a related variable, a difference between two of series, which determined its evolution process. This related variable in turn, establishes a threshold level in transitioning from one regime of correlation to another through the coefficient it has in the state dependent transition probability. In other words, in our simulated case for the restricted model, the threshold levels are proportional to the coefficients of these related variables, such that there may be a higher probability of staying in one regime when these related variables are determined. These threshold levels are estimated by computing a Taylor approximation of these related variables, as per previous section III. b.

In the case of our simulated data, the coefficients of our state dependent probability are $\boldsymbol{b_{12}}$ (or γ_{12}) equal to 2, and $\boldsymbol{b_{22}}$ (or γ_{22}) equal to -2. When the underlying related variable considered has a positive value, the first variable coefficient $\boldsymbol{b_{12}}$ (or γ_{12}) being positive increases the probability of remaining at regime 1 in case of previously being at regime 1. At the same time and also when the related variable considered has a positive value, the second variable coefficient $\boldsymbol{b_{22}}$ (or γ_{22}) being negative increases the probability of switching from regime 2 to regime 1, in case of previously being at regime 2. Hence for steady positive values of our underlying related variable, the first coefficients may become a threshold $\boldsymbol{b_{12}}$ (or γ_{12}) indicating spillover effects or persistence to remain in regime 1, than at regime 2. Additionally, the second variable $\boldsymbol{b_{22}}$ (or γ_{22}) makes the switch from regime 2 to regime 1 more prevalent. In this restricted model, regime 2 ($\Gamma * \lambda$) is at a lower correlation level than regime 1 (Γ). The converse case of spillover effect in regime 2 may be achieved, if the related variable considered has negative values. Thus there is an impact in an inverse manner than the previous case, through the probabilities of either remaining in regime 1 – with π_t^{11}, or of switching to regime 1 from regime 2 – with π_t^{21}. This latter effect is also obtained in our simulated data, as can be seen form charts in Appendix 3.

When modeling our simulated data with the original constant transition probabilities model, we are able to obtain the proper correlation values for both regimes. However, the dynamics of these correlations may be better determined when we include the underlying related variable to the process, as with the state dependent probabilities. This is further corroborated by an initial better fit to the simulated data through our likelihood values from both models.

Being able to determine fundamental underlying variables in the evolution of multiple series or markets enables us to analyze the impact of shocks from these related variables. In other words, the effect a shock from a related variable may have on the multiple series being considered, including the spillovers it produces, can be determined and also long-run implications may be assessed. In addition, having a better portrayal of the changes between different correlation states among multiple series may assist in efficiency of related operations, as well as for risk hedging improvements.

Conclusion

Proper assessment of the dynamic interrelationships among different time series and/or markets, especially during periods of high volatility, is critical for efficiency gains, management of risk and policy analysis. We extend a regime switching dynamic correlation model by including the possibility of

determining the effect of underlying related variables, to the evolution of the dynamic process of multiple series or markets. These related variables of the multiple series modeled, form part of a switching probability process for transitioning between regimes of different correlation values.

We simulate multiple series in two different regimes of correlation values, by explicitly considering a related series in the transition probability for being in one regime or another. We then model their dynamic correlations and find advantages of our model which includes state dependent transition probabilities, than if considering constant transition probabilities, between regimes. Determining the effect of the underlying related variable in the evolution of the dynamic process enables to identify thresholds and spillover effects between the series, as well as better portrayal of the dynamic process for efficiency gains, risk management or policy analysis. Further empirical applications with the unrestricted model will be conducted.

References

Ang, A., J. Chen, (2002), "Asymmetric correlations of equity portfolios", *Journal of Financial Economics*, 63: 443-494.

Baillie, R.T., R.J. Meyer (1991), "Bivariate GARCH estimation of the optimal commodity futures hedge", *Journal of Applied Econometrics*, 6: 109-124.

Bauwens, L., S. Laurent, J.K. Rombouts, (2006), "Multivariate GARCH Models: A Survey", *Journal of Applied Econometrics*, 21: 79 – 109.

Benninga, S., R. Eldor, and I. Zilcha. (1984), "The Optimal Hedge Ratio in Unbiased Futures Markets.", *The Journal of Futures Markets*. 4:155-59.

Bera, A.K., P. Garcia, and J.S. Roh. (1997), "Estimation of Time Varying Hedge Ratios for Corn and Soybeans: BGARCH and Random Coefficient Approaches", *The Indian Journal of Statistics* 59 (3):346-368

Bollerslev, T., (1986) "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31(3): 307-327

Bollerslev, T., (1990) "Modelling the coherence in short run nominal exchange rates: a multivariate generalized arch model." *Review of Economics and Statistics*, 72: 498-505.

Bollerslev, T., Chou, R.Y., Kroner, K.F., (1992) "ARCH modeling in finance", *Journal of Econometrics*, 52(1): 5-59

Boyd, M.S. and B.W. Brorsen (1985), "Dynamic Relationship of Weekly Prices in the United States Beef and Pork Marketing Channels", *Canadian Journal of Agricultural Economics,* 33: 331- 342.

Boyd, M.S. and B.W. Brorsen (1988) "Price Asymmetry in the U.S. Pork Marketing Channel", *North Central Journal of Agricultural Economics,* 10:103-109

Dempster, A.P., N.M. Laird, D.B.Rubin, (1977) "Maximum Likelihood from incomplete data via the EM Algorithm", *Journal of the Royal Statistical Society,* Series B 39 (1): 1-38.

Diebold, F. X., J.H. Lee, G.C.Weinbach, (1994) "Regime Switching with Time-Varying Transition Probabilities", *Non-Stationary Time Series Analysis and Co-Integration*, Advanced Texts in Econometrics, Hargreaves, C.P. (Ed.). The Clarendon Press & Oxford University Press.

Dueker, M., (1997), "Markov Switching in GARCH processes and mean-reverting stock market volatility", *Journal of Business and Economic Statistics*, 15: 26-34.

Engle, R.F. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation", *Econometrica*, 50: 987-1007.

Engel, R., K.F. Kroner, (1995),"Multivariate simultaneous generalized ARCH", *Econometric Theory*, 11: 122-150.

Engle, R., (2002) "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models", *Journal of Business & Economic Statistics*, 20(3): 339-350.

Engle, R., (2004) "Risk and Volatility: Econometric Models and Financial Practice", *The American Economic Review*, 94(3): 405 – 420.

Forbes, K.J., R. Rigobon, (2002) "No Contagion, Only Interdependence: Measuring Stock Market Comovements", *Journal of Finance*, 57(5): 2223-2261

Garcia, R., P. Perron, (1996), "An analysis of the real interest rate under regime shifts", *Review of Economics and Statistics* 78 (1): 111-125.

Goodwin, B.K. and M.T. Holt, (1999), "Price Transmission and Asymmetric Adjustment in the U.S. Beef Sector", *American Journal of Agricultural Economics*, 81: 630-637.

Goodwin, B.K. and D.C. Harper, (2000), "Price Transmission, Threshold Behavior and Asymmetric Adjustment in the U.S. Pork Sector", *Journal of Agricultural & Applied Economics*, 32:543-553.

Goodwin, B.K. and N.E. Piggott, (2001), "Spatial Market Integration in the Presence of Threshold Effects", *American Journal of Agricultural Economics*, 83: 302-317

Gray, S.F. (1996), "Modelling the conditional distribution of interest rates as a regime-switching process", *Journal of Financial Economics*, 42: 27-62

Hamilton, J.D. (1990), "Analysis of Time Series subject to Changes in Regime", *Journal of Econometrics*, 45: 39-70

Hamilton, J.D., (1994), "Time Series Analysis", Princeton University Press, Princeton.

Hanh, W.F. (1990), "Price Transmission Asymmetry in Pork and Beef Markets", *Journal of Agricultural Economics Research*, 42: 21-30

Lee, H.T., J.K. Yoder, R.C. Mittelhammer, and J.J. McCluskey (2006), "A Random Coefficient Autoregressive Markov Regime Switching model for Dynamic Futures hedging ", *Journal of Futures Markets*, 26:103 – 109

Lee, H.T., J.K. Yoder (2007a), "A Bivariate Markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios", *Applied Economics* 39: 1253 – 1265.

Lee, H.T., J.K. Yoder (2007b), "Optimal Hedging with a Regime-Switching Time-Varying Correlations GARCH model", *Journal of Futures Markets*, 27: 495 – 516.

Lien, D.Y., K. Tse, A.K.C. Tsui (2002), "Evaluating the hedging performance of the constant correlation GARCH model", *Applied Financial Economics*, 12; 791-798.

Longin, F., B. Solnik, (1995), " Is the correlation in international equity returns constant: 1960 – 1990?", *Journal of International Money and Finance*, 14: 3-26.

Meyer.J., S. von Cramon-Taubadel (2004) "Asymmetric Price Transmission: A Survey", *Journal of Agricultural Economics* 55(3): 581-611.

Myers, R.J., S.R. Thompson (1989), "Generalized Optimal Hedge Ratio Estimation", *American Journal of Agricultural Economics*, 71: 858-868.

Myers, R.J. (1991), "Estimating time-varying optimal hedge ratios on futures markets", *Journal of Futures Markets*, 11: 39-53.

Pelletier, D.,(2006),"Regime Switching for Dynamic Correlations", *Journal of Econometrics*, 131:445-73

Ramchmand, L., R. Susmel (1998), "Volatility and cross correlation across major stock markets", *Journal of Empirical Finance*, 5:397-416.

Ross, S.R. (1993), *Introduction to Probability Models*, 5th Edition, Academic Press, N.Y.

Schroeder, T.C., (1988). "Price Linkages Between Wholesale & Retail Pork Cuts" *Agribusiness* 4:359-69

Taylor, S., (1986). "Modelling Financial Time Series". Wiley, New York.

Tejeda, H.A., B.K. Goodwin (2009), ""Price Volatility, Nonlinearity and Asymmetric Adjustments in Corn, Soybean and Cattle Markets: Implication of Ethanol-driven Shocks." *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*. St. Louis, MO.

Tse, Y.K., K.C. Tsui (2002), "A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time Varying Correlations", *Journal of Business and Statistics*, 20: 351-362

Appendix 1.1

<u>Expectation Step (as per Hamilton 1990):</u>

Substitution of 'smoothed state' probabilities for the indicator functions in the complete-data log likelihood (from 3. above):

$$
\begin{aligned}
\mathrm{E}\,[\log f\left(\vec{y}_t, \vec{\Delta}_t \middle| \vec{x}_t;\ \theta^{(j-1)}\right)] = \ & \rho^{(j-1)}\ [\ \log f\left(y_1 \middle| \Delta_1 = 1;\ \propto_1^{(j-1)}\right) + \log \rho^{(j-1)}\ ] + \\
& + (1 - \rho^{(j-1)})\ [\ \log f\left(y_1 \middle| \Delta_1 = 0;\ \propto_0^{(j-1)}\right) + \log(1 - \rho^{(j-1)})\ ] \\
& + \textstyle\sum_{t=2}^{T}\ \{ P(\Delta_t = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \log f\left(y_t \middle| \Delta_t = 1;\ \alpha_1^{(j-1)}\right) + \\
& + P(\Delta_t = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \log f\left(y_t \middle| \Delta_t = 0;\ \alpha_0^{(j-1)}\right) + \\
& + P(\Delta_t = 1, \Delta_{t-1} = 1 |\ \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \log(p_t^{11}) + \\
& + P(\Delta_t = 0, \Delta_{t-1} = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \log(1 - p_t^{11}) + \\
& + P(\Delta_t = 1, \Delta_{t-1} = 0 |\ \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \log(1 - p_t^{00}) + \\
& + P(\Delta_t = 0, \Delta_{t-1} = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \log(p_t^{00})\ \}
\end{aligned}
$$

The 'smoothed state' probabilities are obtained from the optimal nonlinear smoother, conditional upon the current 'best guess' of θ, i.e. $\theta^{(j-1)}$.

Appendix 1.2

<u>The Algorithm for calculating the 'smoothed state' probabilities for state *j*, **given** $\theta^{(j-1)}$, $\vec{y}_T$ and $\vec{x}_T$ is</u>:

1.i. Calculate the Conditional densities of y_t, i.e. $f\left(y_t \middle| \Delta_t;\ \alpha_i^{(j-1)}\right)$ by (3.3): a (T x 2) matrix

ii. Calculate Transition Probabilities matrix Π_t as per pg. 19: which is a (T-1) x 4 matrix

2. Calculate 'filtered' joint state probabilities ((T-1) x 4 matrix) by iterating on steps 2a to 2d below, for t = 2,…,T

2.a. Calculate the joint conditional distribution of $(y_t, \Delta_t, \Delta_{t-1})$ (i.e. FOUR #'s) given

$\vec{y}_{t-1}$ and $\vec{x}_{t-1}$:

For t = 2, joint conditional distribution is:

$$f(y_2, \Delta_2, \Delta_1 | y_1, x_1;\ \theta^{(j-1)}) = f\left(y_2 \middle| \Delta_2;\ \alpha^{(j-1)}\right) P(\Delta_2 | \Delta_1, x_1;\ \beta^{(j-1)}) P(\Delta_1)$$

For subsequent time periods t, the joint conditional distribution is:

$$f(y_t, \Delta_t, \Delta_{t-1} | \vec{y}_{t-1}, \vec{x}_{t-1};\ \theta^{(j-1)}) =$$

$$\sum_{\Delta_{t-2}=0}^{1} f(y_t | \Delta_t;\ \alpha^{(j-1)}) P\left(\Delta_t \middle| \Delta_{t-1}, x_{t-1};\ \beta^{(j-1)}\right) P(\Delta_t, \Delta_{t-1} | \vec{y}_{t-1}, \vec{x}_{t-1};\ \theta^{(j-1)})$$

where

$f(y_t | \Delta_t;\ \alpha^{(j-1)})$ *and* $P\left(\Delta_t \middle| \Delta_{t-1}, x_{t-1};\ \beta^{(j-1)}\right)$ are obtained from previous step 1.

and

$P(\Delta_t, \Delta_{t-1} | \vec{y}_{t-1}, \vec{x}_{t-1};\ \theta^{(j-1)})$ is the 'filtered probability' obtained for previous t.

2.b. Calculate conditional likelihood of y_t (ONE #):

(Adds up over All 'states' – in this case two regimes)

$$f\left(y_t \middle| \vec{y}_{t-1}, \vec{x}_{t-1};\ \theta^{(j-1)}\right) = \sum_{\Delta_t=0}^{1} \sum_{\Delta_{t-1}=0}^{1} f(\,(y_t, \Delta_t, \Delta_{t-1} | \vec{y}_{t-1}, \vec{x}_{t-1};\ \theta^{(j-1)})$$

2.c. Calculate the time-t *filtered state* probabilities (FOUR #'s):

$$P(\Delta_t, \Delta_{t-1}|\vec{y}_t, \vec{x}_t;\ \theta^{(j-1)}) = f(y_t, \Delta_t, \Delta_{t-1}|\vec{y}_t, \vec{x}_t;\ \theta^{(j-1)}) \Big/ f(y_t|\vec{y}_{t-1}, \vec{x}_{t-1};\ \theta^{(j-1)})$$

where the numerator is obtained for 2.a. (joint conditional distribution of

$y_t, \Delta_t, \Delta_{t-1}$) and the denominator is the conditional likelihood of y_t , from 2.b.

2.d. These previous FOUR filtered probabilities are used as input for step 2.a. to calculate

the filtered probabilities for the next time period and steps 2.a. – 2.d. are repeated (T-2) times.

3. The calculation of the 'smoothed' joint state probabilities as follows ((T-1) x 6) matrix:

3.a. For t = 2 and given values for (Δ_t, Δ_{t-1}), sequentially calculate the joint probability of

$(\Delta_\tau, \Delta_{\tau-1}, \Delta_t, \Delta_{t-1})$ given $\vec{y}_\tau$ and $\vec{x}_\tau$, for τ = t + 2, t + 3, ……*T:*

$$P(\Delta_\tau, \Delta_{\tau-1}, \Delta_t, \Delta_{t-1}|\vec{y}_\tau, \vec{x}_\tau;\ \theta^{(j-1)}) =$$

$$\frac{\sum_{\Delta_{\tau-2}=0}^{1} f(y_\tau|\Delta_\tau;\ \alpha^{(j-1)}) P\left(\Delta_\tau \middle| \Delta_{\tau-1}, x_{\tau-1};\ \beta^{(j-1)}\right) P(\Delta_{\tau-1}\Delta_{\tau-2}, \Delta_t, \Delta_{t-1}|\vec{y}_{\tau-1}, \vec{x}_{\tau-1};\ \theta^{(j-1)})}{f(y_\tau|\vec{y}_{\tau-1}, \vec{x}_{\tau-1};\ \theta^{(j-1)})}$$

where the first two terms in the numerator are from step 1., the third term in the numerator is from previous computation of step 3.a. , and the denominator by step 2.b.

when τ = t + 2, the numerator's third term is 'initialized' with:

$$P(\Delta_{t+1}, \Delta_t, \Delta_{t-1}|\vec{y}_{t+1}, \vec{x}_{t+1};\ \theta^{(j-1)}) =$$

$$\frac{f(y_{t+1}|\Delta_{t+1}\ ; \alpha^{(j-1)}) P\left(\Delta_{t+1} \middle|\ \Delta_t, x_t;\ \beta^{(j-1)}\right) P(\Delta_t, \Delta_{t-1}|\vec{y}_t, \vec{x}_t;\ \theta^{(j-1)})}{f(y_{t+1}|\vec{y}_t, \vec{x}_t;\ \theta^{(j-1)})}$$

The last term in the numerator is from 2.c.

For each value τ - a (4 x 1) vector of probabilities is produced corresponding to the four valuations of $(\Delta_\tau, \Delta_{\tau-1})$. Hence upon reaching $\tau = T$, we've calculated and saved a (T – 3) x 4 matrix; <u>in which the Last row is used at step 3.b. below</u>.

3.b. Once at $\tau = T$, then the 'smoothed joint state probability for time t' and the chosen valuation of (Δ_t, Δ_{t-1}) is calculated as follows:

$$P(\Delta_t, \Delta_{t-1} | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) = \sum_{\Delta_T=0}^{1} \sum_{\Delta_{T-1}=0}^{1} P(\Delta_T, \Delta_{T-1}, \Delta_t, \Delta_{t-1} | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)})$$

3. c. Steps 3.a. and 3.b. are repeated for all possible time t valuations (Δ_t, Δ_{t-1}) (FOUR in this case), until a smoothed probability has been calculated for *each* of the four possible valuations. Now we have a (1 x 4) vector of 'smoothed joint state probabilities' for (Δ_t, Δ_{t-1}).

3.d. Steps 3.a. – 3.c. are repeated for t = 3, 4, …..T, obtaining a total of T – 1 x 4 smoothed joint state probabilities.

4. Smoothed 'marginal state probabilities' are found by summing over the smoothed joint state probabilities. For example:

$$P\left(\Delta_t = 1 \middle| \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}\right) = P(\Delta_t = 1, \Delta_{t-1} = 1 | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) +$$

$$P(\Delta_t = 1, \Delta_{t-1} = 0 | \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)})$$

These (T – 1 x 6) 'smoothed state probabilities' (FOUR joint and TWO marginal) are used as input for the maximization step.

Appendix 1.3

<u>3.2. Maximization Step:</u>

Once the smoothed probabilities are obtained, the expected complete-data log likelihood given by (3. from pg. 2) is maximized directly with respect to the model parameters.

The first order conditions for the non-linear (logit) transition probabilities parameter vector β, result in a closed form solution for $\beta_0^{(j)}$ and $\beta_1^{(j)}$:

$$\beta_0^{(j)} = \left(\sum_{t=2}^{T} P(s_{t-1} = 0|\vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \frac{\partial \pi_t^{00}(\beta_0)}{\partial \beta_0}\right)^{-1} \text{ x}$$

$$\left(\sum_{t=2}^{T} x_{t-1} \left\{ P\left(s_t = 0, s_{t-1} = 0 \middle| \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}\right) - P\left(s_{t-1} = 0 \middle| \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}\right) \left[\pi_t^{00}\left(\beta_0^{(-1)}\right) - \frac{\partial \pi_t^{00}(\beta_0)}{\partial \beta_0} \beta_0^{(j-1)}\right]\right\}\right)$$

$$\beta_1^{(j)} = \left(\sum_{t=2}^{T} P(s_{t-1} = 1|\vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}) \frac{\partial \pi_t^{11}(\beta_1)}{\partial \beta_1}\right)^{-1} \text{ x}$$

$$\left(\sum_{t=2}^{T} x_{t-1} \left\{ P\left(s_t = 1, s_{t-1} = 1 \middle| \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}\right) - P\left(s_{t-1} = 1 \middle| \vec{y}_T, \vec{x}_T;\ \theta^{(j-1)}\right) \left[\pi_t^{11}\left(\beta_1^{(-1)}\right) - \frac{\partial \pi_t^{11}(\beta_1)}{\partial \beta_1} \beta_1^{(j-1)}\right]\right\}\right)$$

Which in the case of 2 regimes, specifically becomes:

$$\beta_0^{(j)} = \begin{pmatrix} \beta_{00}^{(j)} \\ \beta_{01}^{(j)} \end{pmatrix} = \begin{pmatrix} \sum_{t=2}^{T} x_{0,t-1} P(s_{t-1} = 0)\pi_{0t}^{00} & \sum_{t=2}^{T} x_{0,t-1} P(s_{t-1} = 0)\pi_{1t}^{00} \\ \sum_{t=2}^{T} x_{1,t-1} P(s_{t-1} = 0)\pi_{0t}^{00} & \sum_{t=2}^{T} x_{1,t-1} P(s_{t-1} = 0)\pi_{1t}^{00} \end{pmatrix}^{-1} \text{ x}$$

$$\begin{pmatrix} \sum_{t=2}^{T} x_{0,t-1}\{P(s_{t-1} = 0, s_{t-1} = 0) - P(s_{t-1} = 0)\left[\pi_t^{00} - \frac{\partial \pi_t^{00}}{\partial \beta_0}\beta_0^{(j-1}\right]\} \\ \sum_{t=2}^{T} x_{1,t-1}\{P(s_{t-1} = 0, s_{t-1} = 0) - P(s_{t-1} = 0)\left[\pi_t^{00} - \frac{\partial \pi_t^{00}}{\partial \beta_0}\beta_0^{(j-1)}\right]\} \end{pmatrix}$$

$$\beta_1^{(j)} = \begin{pmatrix} \beta_{10}^{(j)} \\ \beta_{11}^{(j)} \end{pmatrix} = \begin{pmatrix} \sum_{t=2}^{T} x_{0,t-1} P(s_{t-1} = 1)\pi_{0t}^{11} & \sum_{t=2}^{T} x_{0,t-1} P(s_{t-1} = 1)\pi_{1t}^{11} \\ \sum_{t=2}^{T} x_{1,t-1} P(s_{t-1} = 1)\pi_{0t}^{11} & \sum_{t=2}^{T} x_{1,t-1} P(s_{t-1} = 0)\pi_{1t}^{11} \end{pmatrix}^{-1} \text{ x}$$

$$\begin{pmatrix} \sum_{t=2}^{T} x_{0,t-1}\{P(s_{t-1} = 1, s_{t-1} = 1) - P(s_{t-1} = 1)\left[\pi_t^{11} - \frac{\partial \pi_t^{11}}{\partial \beta_1}\beta_1^{(j-1}\right]\} \\ \sum_{t=2}^{T} x_{1,t-1}\{P(s_{t-1} = 1, s_{t-1} = 1) - P(s_{t-1} = 1)\left[\pi_t^{11} - \frac{\partial \pi_t^{11}}{\partial \beta_1}\beta_1^{(j-1}\right]\} \end{pmatrix}$$

Appendix 2.

			State Dependent Probabilities		Constant Probabilities	
	4 Series Simulated	True Betas for Simulation	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model*	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model
Γ	***correlation Reg.1***	0.85	0.84195	0.02275	0.84352	0.01800
	correlation Reg.1	0.85	0.82031	0.02197	0.82205	0.02523
	correlation Reg.1	0.85	0.8373	0.02855	0.83903	0.03017
	correlation Reg.1	0.85	0.78727	0.02465	0.78879	0.02271
	correlation Reg.1	0.85	0.81216	0.01634	0.8137	0.01143
	correlation Reg.1	0.85	0.81239	0.05112	0.81373	0.04660
	correlation Reg.2	0.255	0.19940	0.02595	0.21938	0.04175
	correlation Reg.2	0.255	0.19427	0.02527	0.21380	0.04096
	correlation Reg.2	0.255	0.19830	0.02614	0.21821	0.04201
	correlation Reg.2	0.255	0.18645	0.02444	0.20515	0.03924
	correlation Reg.2	0.255	0.19234	0.02479	0.21163	0.04013
	correlation Reg.2	0.255	0.19240	0.02732	0.21163	0.04182
λ	***Lambda Transtn***	0.3	0.23683	0.03015	0.26008	0.04919
	LIKELIHOOD		-2835.47		-2845.35	
γ	$\mathbf{b_{11}}$ (for cte Reg.1)	0.8	0.80247	0.24011	0.55558	0.20837
	$\mathbf{b_{21}}$ (for cte Reg.2)	1	1.0364	0.48581	0.44203	0.35982
	$\mathbf{b_{12}}$ (for var Reg.1)	2	1.8113	0.31005	0	0
	$\mathbf{b_{22}}$ (for var Reg.2)	-2	-2.175	0.69446	0	0

*Standard Deviations for Correlations in Regime 2 by Delta Method

			State Dependent Probabilities		Constant Probabilities	
	10 Series Simulated	True Betas for Simulation	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model*	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model
Γ	***correlation Reg.1***	0.85	0.82209	0.04300	0.82261	0.03965
	correlation Reg.1	0.85	0.82399	0.02521	0.82471	0.02581
	correlation Reg.1	0.85	0.81159	0.03736	0.8122	0.03560
	correlation Reg.1	0.85	0.82383	0.03368	0.82446	0.03191
	correlation Reg.1	0.85	0.81733	0.02547	0.81796	0.02304
	correlation Reg.1	0.85	0.8053	0.05510	0.80579	0.05258
	correlation Reg.2	0.255	0.21397	0.02832	0.19349	0.03482
	correlation Reg.2	0.255	0.21447	0.02689	0.19399	0.03418
	correlation Reg.2	0.255	0.21124	0.02746	0.19105	0.03417
	correlation Reg.2	0.255	0.21443	0.02750	0.19393	0.03446
	correlation Reg.2	0.255	0.21273	0.02670	0.19240	0.03380
	correlation Reg.2	0.255	0.20960	0.02924	0.18954	0.03512
λ	***Lambda Transtn***	0.3	0.26028	0.03164	0.23522	0.04079
	LIKELIHOOD		-2861.10		-2864.71	
γ	$\mathbf{b_{11}}$ (for cte Reg.1)	0.8	0.78777	0.17055	0.57303	0.14454
	$\mathbf{b_{21}}$ (for cte Reg.2)	1	0.94704	0.34643	0.41169	0.29068
	$\mathbf{b_{12}}$ (for var Reg.1)	2	1.7775	0.51231	0	0
	$\mathbf{b_{22}}$ (for var Reg.2)	-2	-1.9449	0.67193	0	0

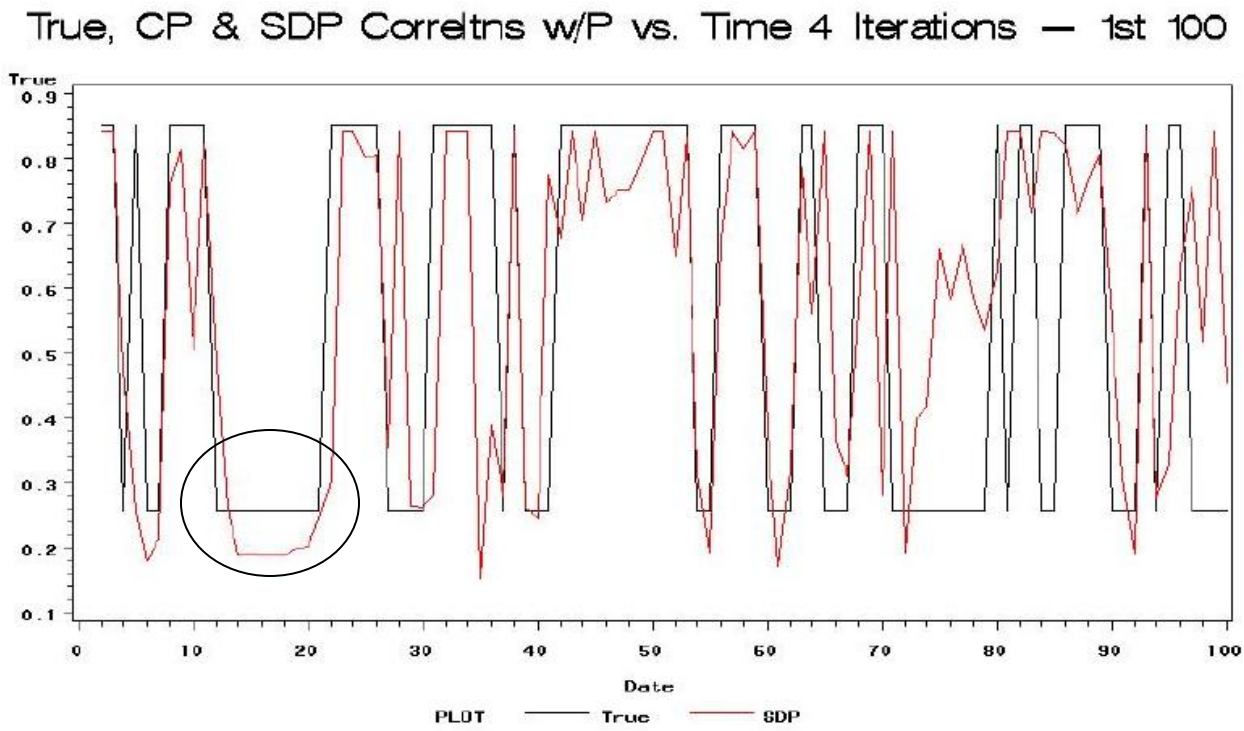
*Standard Deviations for Correlations in Regime 2 by Delta Method

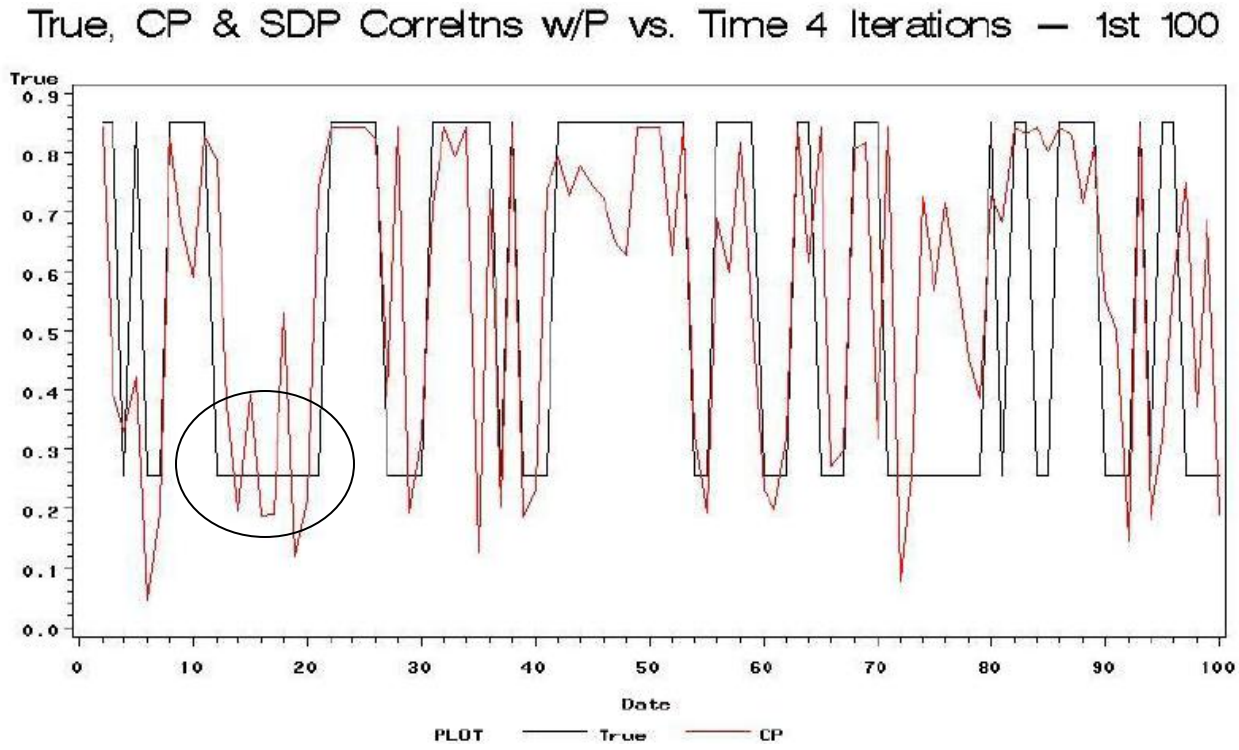
			State Dependent Probabilities		Constant Probabilities	
	15 Series Simulated	True Betas for Simulation	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model*	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model
Γ	***correlation Reg.1***	0.85	0.80159	0.03571	0.80324	0.03231
	correlation Reg.1	0.85	0.82191	0.03606	0.82382	0.03734
	correlation Reg.1	0.85	0.8231	0.04408	0.82493	0.04392
	correlation Reg.1	0.85	0.81552	0.03394	0.81746	0.03645
	correlation Reg.1	0.85	0.81236	0.02986	0.81408	0.02706
	correlation Reg.1	0.85	0.82311	0.04134	0.8251	0.04407
	correlation Reg.2	0.255	0.18986	0.03442	0.19184	0.03852
	correlation Reg.2	0.255	0.19468	0.03527	0.19675	0.03972
	correlation Reg.2	0.255	0.19496	0.03582	0.19702	0.04016
	correlation Reg.2	0.255	0.19316	0.03489	0.19523	0.03939
	correlation Reg.2	0.255	0.19242	0.03455	0.19443	0.03880
	correlation Reg.2	0.255	0.19496	0.03564	0.19706	0.04017
λ	***Lambda Transtn***	0.3	0.23686	0.04163	0.23883	0.04699
	LIKELIHOOD		-2908.17		-2919.19	
γ	$\mathbf{b_{11}}$ (for cte Reg.1)	0.8	0.90659	0.25991	0.58582	0.1834
	$\mathbf{b_{21}}$ (for cte Reg.2)	1	1.015	0.42433	0.39757	0.26283
	$\mathbf{b_{12}}$ (for var Reg.1)	2	2.0372	0.60088	0	0
	$\mathbf{b_{22}}$ (for var Reg.2)	-2	-2.1888	0.76046	0	0

*Standard Deviations for Correlations in Regime 2 by Delta Method

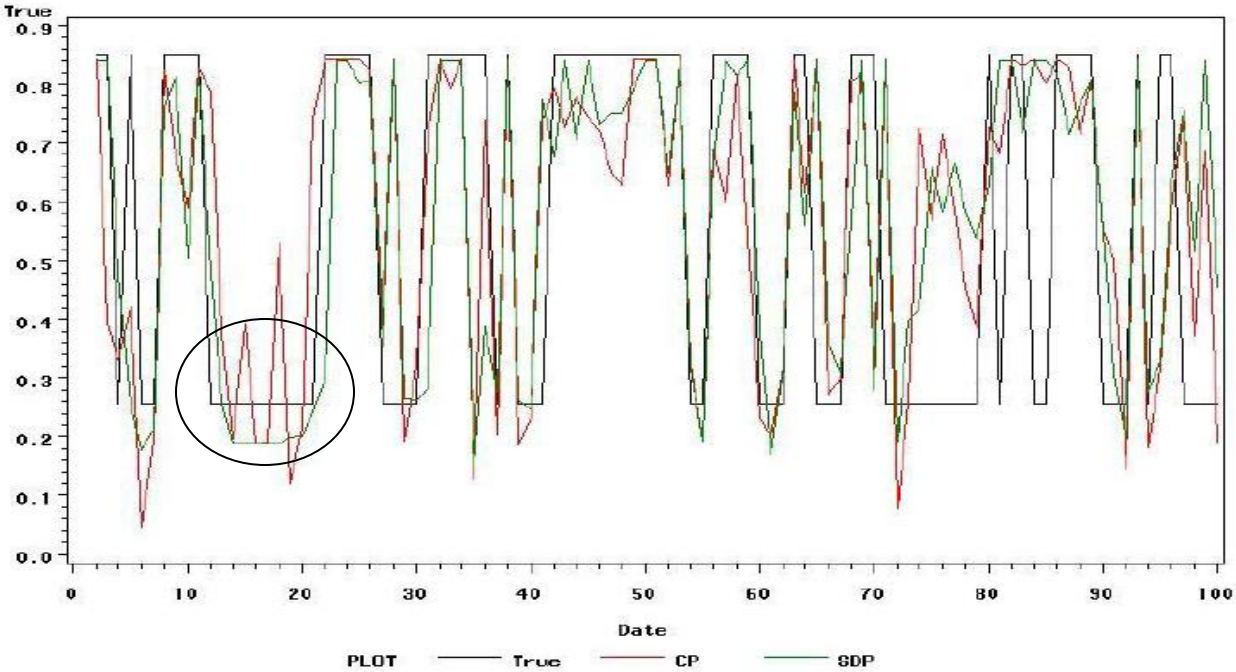
	Un_Restricted_Model		State Dependent Probabilities		Constant Probabilities	
	4 Series Simulated	True Betas for Simulation	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model*	Mean of Betas obtained from model	Stnd Devtn. of Betas obtained from model
$\hat{\Gamma}_1$	***correlation Reg.1***	0.85	0.843547	0.01126	0.863472	0.01364
	correlation Reg.1	0.85	0.831235	0.01577	0.858808	0.01812
	correlation Reg.1	0.85	0.839049	0.01946	0.873877	0.02353
	correlation Reg.1	0.85	0.837465	0.01184	0.835509	0.03578
	correlation Reg.1	0.85	0.839907	0.00886	0.84368	0.00964
	correlation Reg.1	0.85	0.83229	0.01561	0.841237	0.01951
$\hat{\Gamma}_2$	***correlation Reg.2***	0.255	0.181037	0.04439	0.24073	0.07972
	correlation Reg.2	0.255	0.269242	0.04885	0.242506	0.03026
	correlation Reg.2	0.255	0.237202	0.06632	0.212283	0.04423
	correlation Reg.2	0.255	0.251787	0.05697	0.279669	0.05431
	correlation Reg.2	0.255	0.277317	0.03697	0.285705	0.02399
	correlation Reg.2	0.255	0.294502	0.01751	0.272715	0.04962
	LIKELIHOOD		-2349.00		-2407.00	
γ	$\mathbf{b_{11}}$ (for cte Reg.1)	0.8	0.99047	0.07021	0.428275	0.12847
	$\mathbf{b_{21}}$ (for cte Reg.2)	1	0.849528	0.11754	0.521075	0.08040
	$\mathbf{b_{12}}$ (for var Reg.1)	2	2.58407	0.426541	0	0
	$\mathbf{b_{22}}$ (for var Reg.2)	-2	-2.3205	0.537097	0	0

Appendix 3.

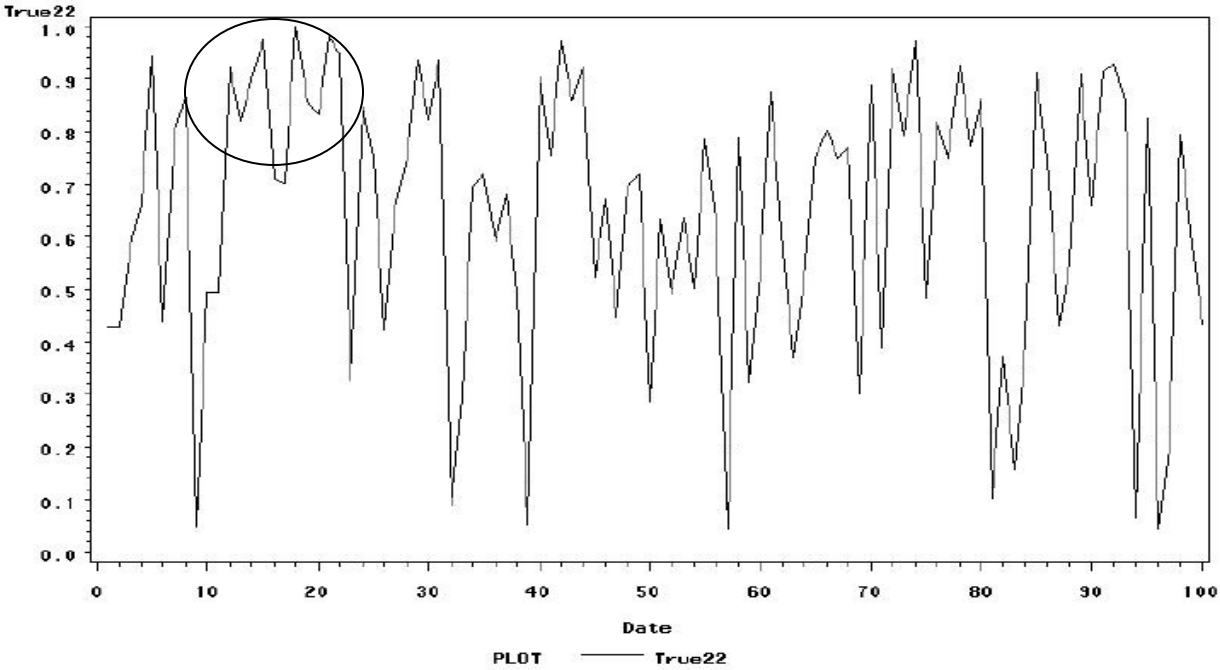

True, CP & SDP Correltns w/P vs. Time 4 Iterations — 1st 100
True
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10
20
30
40
50
60
70
80
90
100
Date
PLOT
True
SDP


True, CP & SDP Correltns w/P vs. Time 4 Iterations — 1st 100
True
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40
50
60
70
80
90
100
Date
PLOT
True
CP

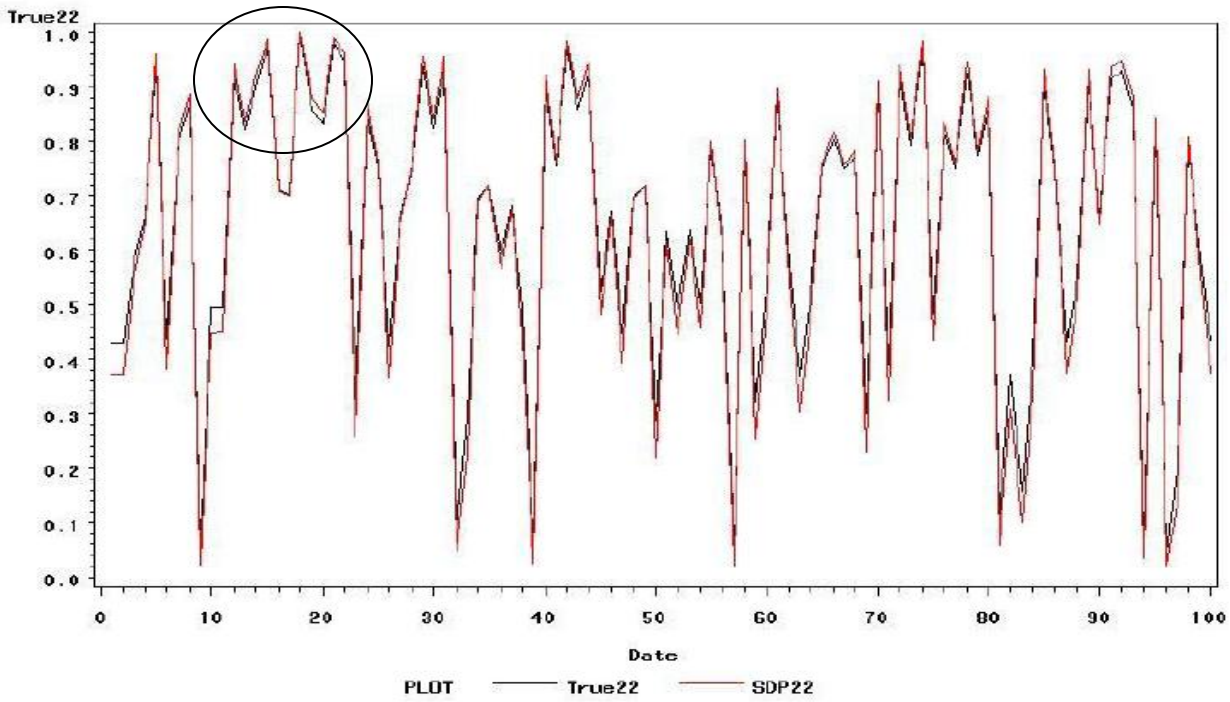
True, CP & SDP Correltns w/P vs. Time 4 Iterations — 1st 100



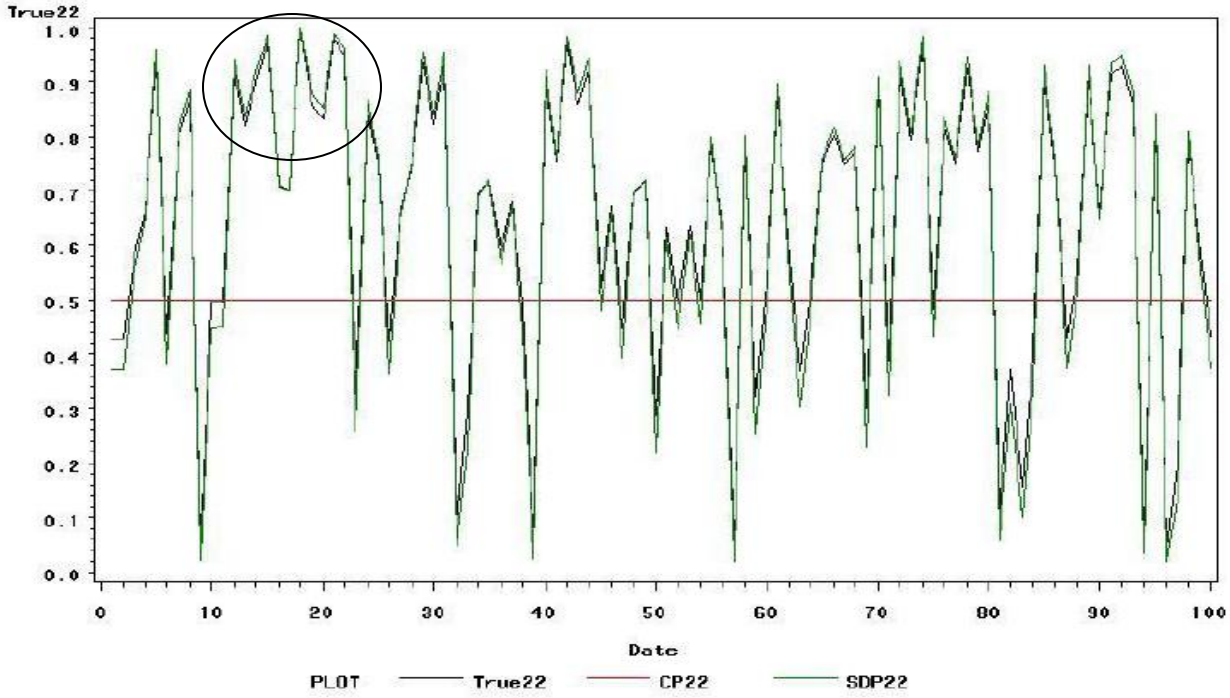
True Probs for P(St =2 ¦ St−1 =2) in 4 ltrtns — 1st 100

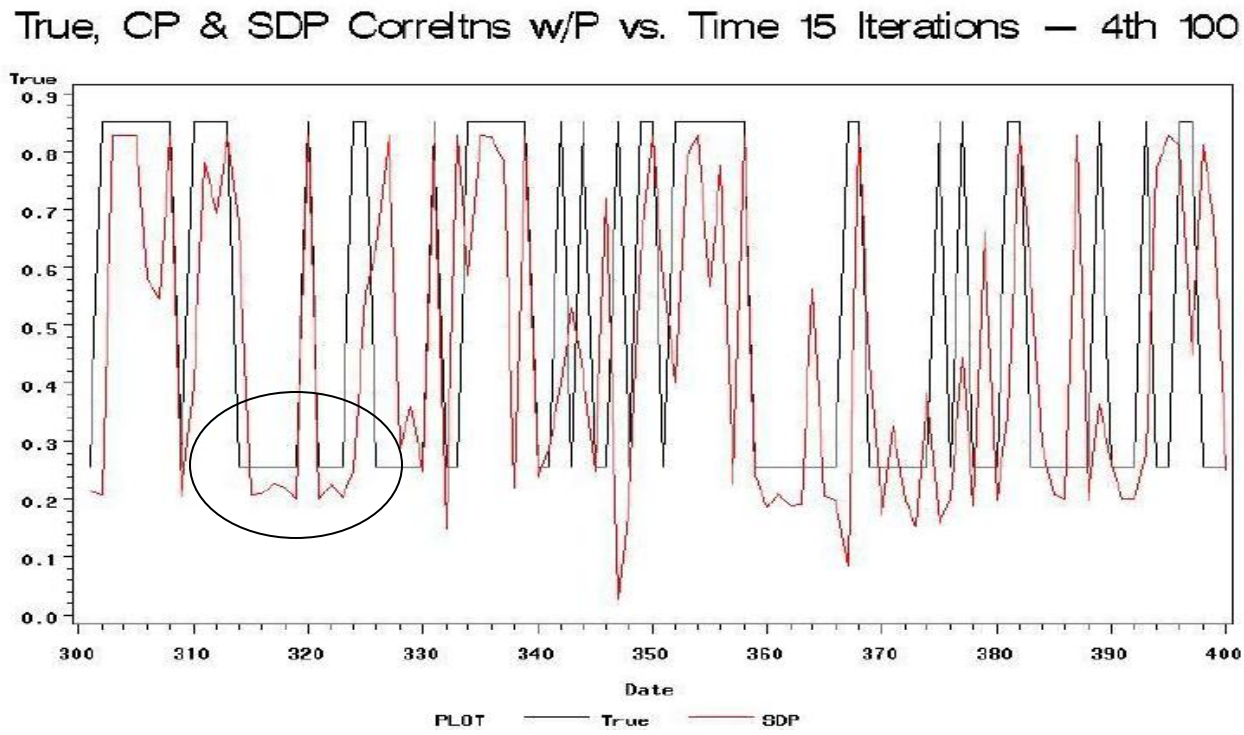


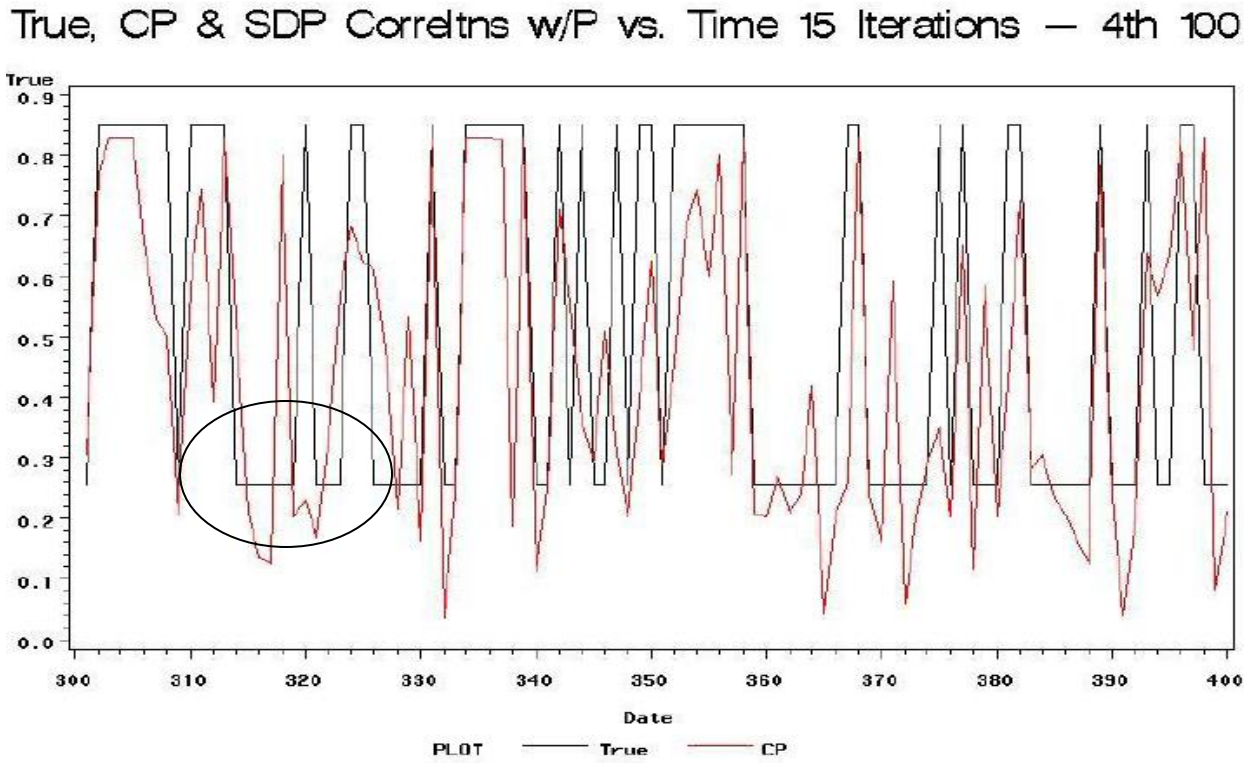
True Probs & SDP for P(St =2 ¦ St−1 =2) in 4 Itrtns — 1st 100



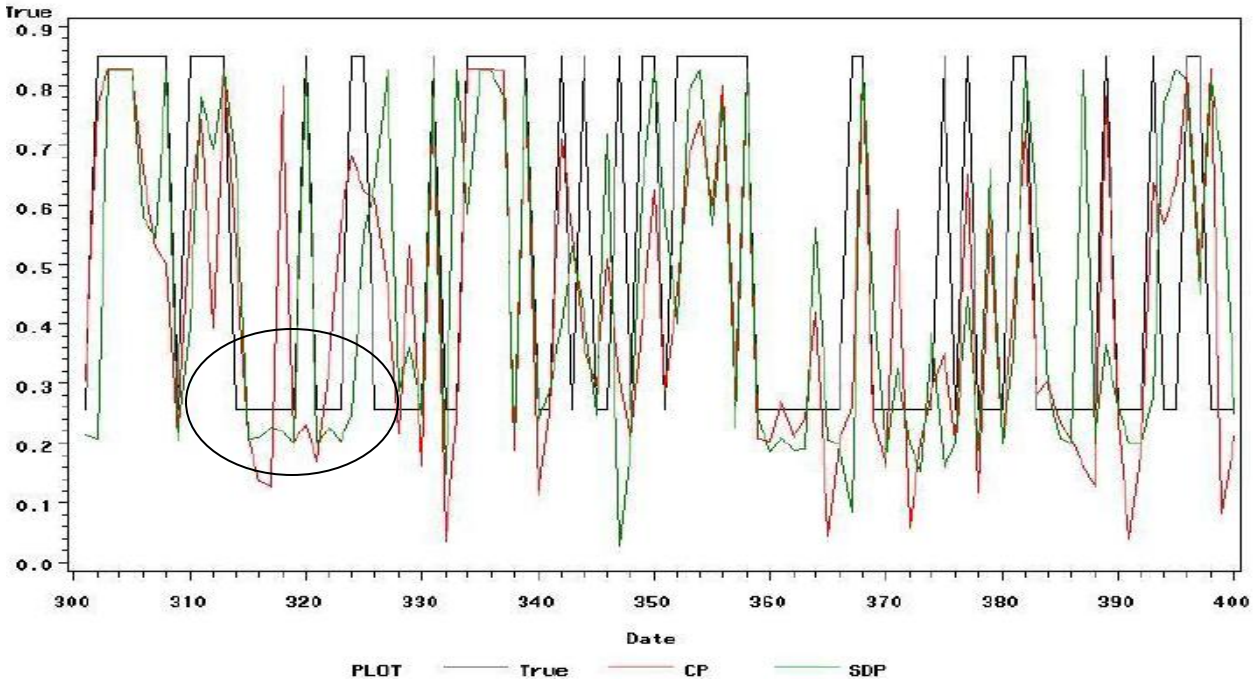
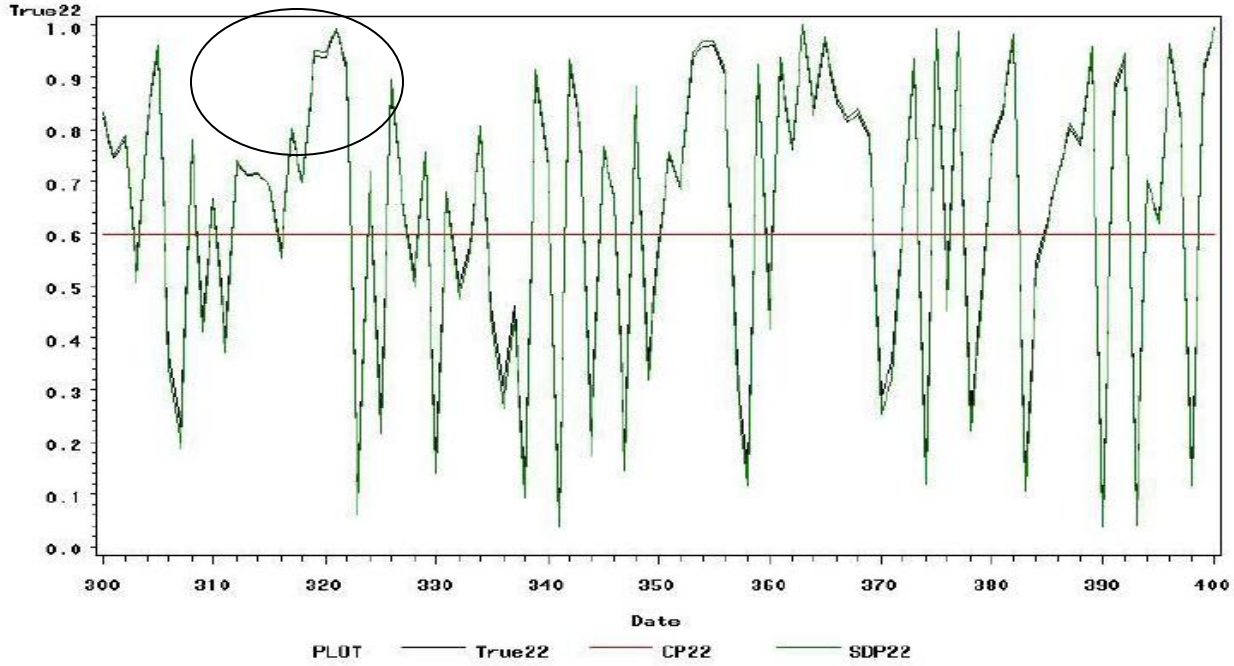
True Probs & SDP & Cte for P(St =2 ¦ St−1 =2) in 4 Itrtns — 1st 100




True, CP & SDP Correltns w/P vs. Time 15 Iterations — 4th 100
True
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
300
310
320
330
340
350
360
370
380
390
400
Date
PLOT
True
SDP


True, CP & SDP Correltns w/P vs. Time 15 Iterations — 4th 100
True
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
300
310
320
330
340
350
360
370
380
390
400
Date
PLOT
True
CP

True, CP & SDP Correltns w/P vs. Time 15 Iterations — 4th 100



True Probs & SDP & Cte for P(St =2 | St−1 =2) in 15 Itrtns — 4th 100