IS OPTION-IMPLIED INFORMATION FORWARD-LOOKING? AN APPLICATION IN FORECASTING MARKET REGIMES

Yi Ling Michelle-Joy Low[1]

The University of Melbourne

ABSTRACT

This paper studies the degree to which option implied information is forward-looking. We incorporate information measures based on the risk neutral distribution (RND) of S&P500 options in a Hidden Markov Model (HMM) setting to explore the degree to which RND-based information is useful for forecasting changes in market regime as well as the future distribution of the underlying asset prices. The RND moments are decomposed into their long-term and short-term dynamics, which are considered, in addition to the RND quantiles, measures of potentially forward-looking information. Compared to a benchmark HMM, we find that all RND-based measures contain incremental information for predicting changes in market regime. Furthermore, the long-term dynamics of the RND moments are particularly useful for improving interval forecasts of the future underlying asset price.

[1] Department of Finance, Level 12, 198 Berkeley Street The University of Melbourne 3010 Victoria Australia | Email: y.low2@pgrad.unimelb.edu.au

* For helpful feedback and comments on earlier drafts, I am thankful to Paul Kim and Hayat Khan. I am also especially grateful to Jonathan Dark and John Handley for many invaluable discussions. All remaining errors are my own.

1 INTRODUCTION

Numerous studies have probed the possibility of forecasting asset returns. It is widely believed that regimes exist in market prices; these are frequently interpreted as “bull” and “bear” regimes, where a bull typically characterizes positive trending market prices and relatively low volatility, whereas a bear represents negatively trending market prices with high volatility. There is wide interest among academics and practitioners in extracting these market price trends from the data for *ex-ante* analysis; if a predictive indicator exists which can signal switches over from bull and bear phases of the market, particularly an imminent crisis, investors can time their entry into and exit from the market and profit therefrom.

This setting serves the objective of this paper - are option prices forward-looking about future underlying asset prices? Building on recent work which recommends regime switching models for identifying bull and bear markets, in this paper we condition Hidden Markov Models (HMM) on RND-based measures of information and assess the improvement of their returns forecast performance. Our objective is to query the degree to which the RND contains information about *future* underlying asset prices –is there forward-looking information useful for 1) identifying the state of the market, and/or, 2) predicting the distribution of returns on the underlying asset? This paper augments previous research which has focused largely on the forward-looking information in implied volatility for predicting future volatility, in particular we probe whether there is potentially information contained in the *entire* risk neutral distribution of option prices for the distribution of future asset prices. HMMs allow for the number of states to be determined, and produce forecasts of both the changes in market state and the future market returns distribution. Their application therefore provides visibility into several ways in which RND information may be useful. Secondly and relatedly, we consider a variety of RND-based information measures in our HMM setup, thereby learning about how separate components of the RND may be useful for the prediction of market state and/or returns distribution.

We allow the data to identify the presence of states in market prices and proceed to estimate a large set of HMM models with various specifications for option-implied information. These account differently for the information contained in slow- and fast moving RND dynamics as well as the RND quantiles. Using the recently developed model confidence set (MCS) methodology of Hansen, Lunde and Nasen (2011), the models' forecast performance are jointly compared against a benchmark time homogenous HMM model specified without conditioning information. The results have implications for related economic and financial applications concerned with the prediction of the financial market state and returns over time. We found evidence of forward-looking RND information pertaining to changes in market regime; specifically the tails of the RND as well as its long-range dynamics are useful for predicting changes between market regime. Further, while the RND-based HMM out-of-sample point forecasts of returns are indistinguishable from those of a time homogenous HMM with no conditioning covariates, the former displays significantly improved interval forecasts over the latter, suggesting that option-implied information is forward-looking regarding the future *distribution* of underlying asset returns.

1.1 ARE OPTION PRICES FORWARD-LOOKING?

The question of whether option prices contain forward looking information has been studied often, but remains in debate. Proponents believe option markets do attract informed investors and several reasons have been advanced to support this view (Black, 1975; Manaster and Rendleman, 1982). The leverage effect of options means that there is more to be gained for a given investment in options than directly in the underlying stock, and an investor with important information will likely choose to deal with options. Practical matters also provide reasons for the option market to attract sophisticated informed traders – options trading incurs lower transaction costs relative to trading in stocks; it is easy to short-sell a stock through a long put option or a short call option position; leverage is achieved directly in the option market compared to maintaining margin accounts with tight requirements in the cash market. In theory, Back (1973) develops a model that concludes that trades in option and the underlying asset conveys different types of information, hence the existence of options lead to richer signals transmitted by the market.

The empirical research has studied particularly the ability of option prices to forecast future market movements. Bates (1991) argues that there were early warning signals of the October 1987 market crash. On examining the prices of the S&P 500 futures options, he found that the OTM put options on S&P 500 futures traded unusually higher than OTM calls during the period of the market upsurge prior to the crash. As OTM puts are typically purchased to provide protection against a market decline, this suggests that even as the stock market was rising, the options market manifested expectations of a market decline.

Similarly, Fung (2007) and Birru and Figlewski (2011) found anticipatory power of option prices for imminent crashes. Malz (2001) demonstrated that IVs can be useful for predicting impending market stress in the immediate horizon while Gray, Merton, Bodie (2008) purport that option-implied skewness may "capture some aspects of fire sale risk and counterparty risk between institutions", both of which are critical factors in market crashes.

A number of papers have also explored whether IV indices can serve as leading indicators for the equity market prices. The implied volatility can be viewed as a barometer of investors' fear (Whaley, 2000), where a rising volatility index relates to negative equity returns. Copeland and Copeland (1999) proposed trading strategies based on the VXO index and show that these strategies yielded positive returns. Giot (2005) reasons that a high (low) level of volatility indicates an oversold (overbought) market. His findings substantiate the thesis that regimes of high (low) values VXO are associated with positive (negative) future S&P 100 stock returns for holding periods of up to 60 days. Banerjee et al. (2007) show that a regression of returns of NYSE stock portfolios formed on size, book-to-market and beta on VIX show predictive power, and the evidence is stronger for longer holding period returns (60-day rather than 30-day). Similarly, Diavatopoulos et al. (2009) find that the idiosyncratic IV of stock options has predictive power for future underlying stock returns for up to one-month ahead.

The empirical evidence, however, has not been conclusively in favour of the forward-looking capability of option prices. While Bates (1991) found indications in S&P500 futures option market of early anticipation of the crash of October 1987, he acknowledges that options prices failed to detect the crash in the two months immediately prior to the crash. Particularly for implied volatility, the evidence has been inconclusive as to its forward-looking capacity; early on, Chiras and Manaster (1978) found that once dividends were accounted for, contrary to other studies, implied volatility did not significantly outperform historical standard deviation in forecasting future volatility. Noh, Engle and Kane (1994) developed trading rules based on two competing forecast methods for option prices, namely an implied volatility-based forward-looking method and a GARCH model using historical information. They found that the GARCH forecast model outperformed the implied volatility model in generating trading profits. Canina and Figlewski (1993) found implied volatility to produce poor forecasts of future realized volatility while analogously, White, Becker and Clements (2004) controlled for historical information in their study of the forward-looking information in the VIX – and found that no significant information for future S&P500 volatility. Koopman, Jungbacker and Hol (2005) assessed the performance of historical, implied and realised volatility measures for forecasting daily S&P100 variability, and found that realised volatility models significantly outperformed implied volatility and other models in terms of forecast errors. A related argument has been that option markets incorrectly price the likelihood of future asset price movements [c.f. Ait-Sahalia, Wang and Yared (2000)].

1.2 THE PRESENCE OF REGIME DYNAMICS IN MARKET PRICES

That the debate continues regarding whether options embody forward-looking information bears testament to how forecasting future asset prices is of considerable interest to investors. Previous literature has examined at great length the forecasting capability of option-implied volatility, however, it has yet to be explored extensively the degree to which information in the *entire* risk neutral distribution is forward-looking. Furthermore, in addition to forecasting volatility, it remains to be investigated how option-implied information may be useful in several other ways – namely do options look forward towards market regime switches and changes in returns distributions? According to Shilling (1992), during the period from 1946 to 1991, if investors had stayed out of the market during the 50 weakest months, they could have increased their annual return from 11.2% to 19.0%. Similarly, being absent from the market during the 50 strongest months would have reduced the annual returns to only 4% per annum. Such ex-ante detection of market state changes requires understanding the markets expectation of the future spread of prices.

Many reliable methods of market regime identification have been ex-post in nature; several previous researchers have developed dating algorithms which classify market phases according to a formal set of rules. Bry and Boschan (1971) produced a well-known algorithm for identifying the turning points of business cycles. The Bry-Boschan algorithm was then developed by Pagan and Sossounov (2003) for the analysis of bull/bear states in monthly equity prices. Lunde and Timmermanm (2004) proposed an identification routine based on the location of local maxima and minima in the time series relative to a pre-specified threshold from an earlier identified peak or trough. These methods by Pagan and Sossounov (2003) and Lunde and Timmermann (2004) have been widely applied, but have the drawback that market turning points are identified only ex-post.

As an alternative to ex-post methods which identify market states with complete certainty, recent papers have adopted forward-looking methods to examining bull/bear states. Estimating the probability associated with a given future state provides insight into the likelihood of a change in or perpetuation of current market state, which in turn may be useful for investment and risk management decisions. For this purpose, regime-switching models have been popular choices when dealing with time-series which are expected to be cyclical in nature. The two-state Markov-switching (MS) model of Hamilton (1989) was calibrated to quarterly US GNP growth rates to estimate transition probabilities governing the alternating growth and recession phases. Chauvet and Potter (2000) forwarded a coincident indicator of the market state using a MS specification from stock market returns. Maheu and McCurdy (2000) delivered similar efforts analysing stock market returns, allowing for duration-dependent transition probabilities in their MS model. Guidolin and Timmermann (2005) adopted a three-state regime-switching calibration for identifying bull/bear states in UK stock and bond returns, on which return predictability and portfolio allocation is dependent. The work was then extended in Guidolin and Timmermann (2006), who appended another state to account for nonlinearity in returns attributable to small and big stocks as well as long-term bonds. More recently, Maheu, McCurdy and Song (2012) proposed a four-state MS model for weekly stock returns, which captures the main bull/bear phases as well as within-state short-term reversals. Maheu, McCurdy and Song showed that allowing for both intra- and inter-state dynamics in the model, when incorporated with Bayesian methods for parameter uncertainty, leads to superior ex-ante forecasting of market price directions.

1.3 Using Forward-Looking Information in Forecasting Market Regime Dynamics

We adapt the definitions of market state from the recent literature. Particularly, Maheu and McCurdy (2000) associate bull markets with high returns and low volatility, whereas the converse (low returns and high volatility) is described of bear markets. Guidolin and Timmermann (2005) interprete bull and bear states similarly, but in addition consider a third state whose asset returns and volatilities are closer to their historical averages and tend to take on values between their counterparts in the bull and bear states. Consistent with Guidolin and Timmermann (2005), our model selection criterion determine a three-state HMM to provide suitable description of our data.

As a contribution to the literature, we query several different RND-based measures of option-implied information. Results in earlier work suggest that the RND contains separate information in its slow- and fast-moving dynamics, as well as its tails, which may be forward-looking. Particularly, earlier findings of long-term and short-term dynamics in the RND suggest that the separate RND dynamics contain forward-looking information which may be useful for forecasting changes in the regime of the underlying asset price; meanwhile, we are also interested in the extent to which short-term RND dynamics may be useful for predicting the returns distribution in the short-term. Using long memory methods, we disaggregate the RND moments into their slow- and fast-moving components, which are each considered in separate HMM model specifications. As an alternative, we also consider RND quantiles as potential sources of forward-looking option-implied information, paying attention to potential forward-looking information in the tails.

One-step ahead out of sample forecast results suggest that the RND information is most useful for predicting changes in market regime. Where transition probabilities were conditioned on RND-based measures of information, all considered HMM-RND models significantly outperformed the benchmark HMM model in terms of the number of correctly predicted state transitions. This result remains even for the alternative specification of the HMM-RND models where covariates were specified directly in the state distribution and were not allowed to affect transition probabilities between state.

Results for the out-of-sample returns distribution forecasts show indistinguishable improvements for point forecasts of the future market prices; however the RND contributes significantly to the interval forecasts produced by the HMM-RND models. Overall, HMM-RND models whose transition probabilities or state distribution parameters were conditioned on RND-LT moments were found to produce superior interval forecasts over the other models considered. For certain error metrics considered, the higher order RND moments in levels as well as tail quantiles were also found to produce significantly superior interval forecasts compared to the other considered models.

The rest of this paper is arranged as follows. Section 2 outlines the theory of HMMs. In section 3 we describe our treatment of the RND to obtain measures of option-implied information. Section 4 briefly describes the theory of forecasting and state identification in HMMs, followed by Section 5 which discusses the practical implementation of our out-of-sample forecast and state-prediction analysis with option-implied information. Results of our analysis are presented in Section 6, while Section 7 considers an alternative HMM-RND specification and its implications. Finally, section 8 concludes.

2 BACKGROUND TO HIDDEN MARKOV MODELS

2.1 AN OVERVIEW OF HMMS

Hidden Markov Models (HMM) also known as Markov switching models, are a special class of mixture models which are useful for modelling univariate and multivariate financial time series. Useful references for HMMs include Zucchini and Macdonald (2009), Cappé, Moulines and Rydén (2005) and Ephraim and Merhav (2002).

HMMs are derived from two stochastic processes, one of which is observed, whose observations are dependent on another underlying unobserved process. More specifically, the two processes are:

1. An unobserved Markov process $\{S_t: t = 1,2,\ldots\}$ which determines the state at time t and
2. A process $\{X_t: t = 1,2,\ldots\}$ which is state dependent and produces the observed time series conditional on the state, S_t, at time *t*

The appeal of HMMs therefore lies with allowing for statistical inference regarding some latent unobserved process. It is useful for analysing the stock market since it is believed that the market goes through phases which are, ex ante, unobserved. There are several desirable properties associated with HMMs [cf. MacDonald and Zucchini (1997)]; (i) all moments can be computed, (ii) the computation of the likelihood function is linear in the number of sample observations and is therefore straightforward to compute, (iii) marginal distributions are straightforward to derive, and treatment for omitted observations is simple, (iv) conditional distributions can be derived and out-of-sample forecasts generated from HMMs.

Suppose $\{X_t, t = 1,2,\ldots\}$ is a stochastic process observed at times $= 1,2,\ldots$. The observed data at t are generated conditionally on true unobserved states $\{S_j, j = 1,2,\ldots,m\}$ according to some distributions.

Further, the unobserved states are assumed to be a Markov Chain (MC) that satisfies the Markov property. A simple *m*-state HMM is then defined as

[1] $P(S_t|S_{t-1}, S_{t-2}, \dots) = P(S_t|S_{t-1})$

[2] $P(X_t|X_{t-1}, X_{t-2}, \dots, S_{t-1}, S_{t-2}) = P(X_t|S_t) \qquad t \in \mathrm{N}$

Associated with the MC are transition probabilities that characterize transition from state i to state *j* over time interval *t*:

[3] $\gamma_{ij}(t) = P(S_{r+t} = j|S_r = i) \qquad i, j = 1,2, \dots, m; t, r \in \mathrm{N}$

If these probabilities do not depend on *t*, the MC is said to be homogeneous, otherwise it is non-homogeneous. The elements $\gamma_{ij}(t)$ can be arranged in a matrix to yield the transition probability matrix

[4] $\mathbf{\Gamma}(t) = [\gamma_{ij}(t)]$

For brevity, when *t*=1 the matrix $\mathbf{\Gamma}(\boldsymbol{t})$ is simply expressed as $\mathbf{\Gamma}$.

The state-dependent distributions $P(X_t|S_t)$ may be discrete or continuous. When X_t can only take on discrete values, then

[5] $p_i(x) = P(X_t = x|S_t = i), \qquad i = 1,2, \dots, m$

is the probability mass function of X_t at time t when the MC is at state i; for continuous random variable X_t, $p_i(x)$ is defined analogously as the probability density function.

Given an observed sequence of observations, $\{x_1, x_2, \dots, x_T\}$ assumed to be generated from some *m*-state HMM, the likelihood function can be expressed as

[6] $L_T(\boldsymbol{v}) = \pi \boldsymbol{P}(x_1)\mathbf{\Gamma}\boldsymbol{P}(x_2)\mathbf{\Gamma}\boldsymbol{P}(x_3) \dots \mathbf{\Gamma}\boldsymbol{P}(x_T)\mathbf{1}'$

where $\boldsymbol{v}$ is the set of all model parameters, π is the initial probability distribution of the states, $\mathbf{\Gamma}$ is the transition probability matrix, $\boldsymbol{P}(x_t)$ is a diagonal matrix of conditional probability distribution where the

ith diagonal element is given by [5] and $\mathbf{1}'$ is a column vector of ones. Furthermore, if π is the stationary distribution of the MC, the likelihood can be expressed as

[7] $$L(\boldsymbol{v}) = \pi\boldsymbol{\Gamma}\boldsymbol{P}(x_1)\boldsymbol{\Gamma}\boldsymbol{P}(x_2)\boldsymbol{\Gamma}\boldsymbol{P}(x_3)\ldots\boldsymbol{\Gamma}\boldsymbol{P}(x_T)\mathbf{1}'$$

Maximizing of the likelihood in [6] or [7] can be carried out by direct maximization or by an EM algorithm known as the Baum-Welch algorithm[2].

In applying the HMM to study the stock market, observed asset returns correspond to the state-dependent process,$\{X_t, t = 1,2,\ldots\}$, while the unobserved latent state process relates to the different phases of the equity market. The different phases are commonly interpreted as the bull, bear or some intermediate transition regimes between the bull and bear.

2.2 HMM MODEL SET-UP

We extend the basic homogeneous HMM model above to allow for covariates to affect the transition probabilities. The *msm* package in R for covariates, $\boldsymbol{x}$, to affect transition intensities through a proportional hazard model suggested by Marshall and Jones (1995), where the transition intensity between states i and j at time t, $\mu_{ij}(t)$, is given by

[7] $$\mu_{ij}(t) = \mu_{ij}^0 exp(\boldsymbol{\beta}'\boldsymbol{x}(t))$$

Transition probabilities are calculated assuming that the covariates are piecewise-constant over a given time period.[3]

Another extension to the basic homogeneous model which we considered is to allow for covariates in the state-dependent distributions. Following earlier papers, we assume Gaussian state dependent distributions

---

[2] Details of the EM algorithm are contained in Appendix A1.

[3] From the transition intensities, the transition probability matrix, $\boldsymbol{P}(u) = [p_{rs}(u)]$, where $p_{rs}(u) = \Pr(S_{t+u} = s | S = s)$ can be calculated by taking the matrix exponential of the transition probability matrix, $\boldsymbol{Q}(t) = [\mu_{ij}(t)]$, that is, $\boldsymbol{P}(u) = Exp(u\boldsymbol{Q})$. The matrix exponential is defined by the usual "power series" $Exp(\boldsymbol{Z}) = I + \frac{\boldsymbol{Z}^2}{2!} + \frac{\boldsymbol{Z}^3}{3} + \cdots$ where $\boldsymbol{Z}^k$ is the matrix product, not element-wise scalar multiplication.

with means that depend on covariates. That is, the state-dependent distribution for asset returns, r_t conditional on state s_i is given by

[8] $$r_t | s_i \sim N(\mu_{s_i}, \sigma_{s_i}^2), \qquad i = 1,2,\dots,m$$

[9] $$\mu_{s_i} = \alpha_{s_i} + \boldsymbol{\beta}_{s_i}{}' \boldsymbol{x}_t$$

Option-implied information such as that contained in the RND moments and quantiles may be used as covariates in the above HMM models to test if they improve on the basic homogeneous HMM model in predicting market regimes and return distributions.

The existence of market regimes is well documented in previous research. While bull and bear regimes are commonly acknowledged, some researchers have gone beyond the two-regime market and considered three or four regimes. McCurdy and Song (2012) proposed a four-state regime-switching model which can be interpreted to account for within-state bear rallies and bull corrections in addition to the main bull and bear phases of the financial markets. However, rather than to identify intricate market regime characteristics, our purpose is to investigate the extent of forward-looking information in options. Therefore we propose a more parsimonious and implementable[4] three-state HMM which allows us to pursue our research question. The choice of a three-state model is in line with Guidolin and Timmermann (2005); we similarly interpret the three states to represent bull and bear regimes, with the richer dynamic of a third intermediate transition state which can occur between bull and bear regimes. With the three-state HMM we can assess model performance on returns forecasts with and without option-implied information, when faced with three potential states of the market – the bull, bear and the transitory phases between bull and bear states.

Our choice of a three-state model is supported by the data. We fit 2-, 3- and 4-state HMM with Gaussian state-dependent distributions to daily log returns of SPI index for the period 5/1/1996 - 29/4/2011 and obtained the following results.

[4] Convergence issues due to numerical overflow in HMMs with a larger number of states is a known issue.

Model	S_i	μ_i	σ_i	Log L	AIC	BIC	Mean	Variance	Skewness	Kurtosis
m=2	1 2	-0.00093 0.00079	0.01944 0.00804	11481.74	-22955.48	-22930.57	0.00019	0.01310	-0.16402	5.26338
m=3	1 2 3	-0.00203 0.00004 0.00108	0.02827 0.01200 0.00611	11645.16	-23272.32	-23216.77	0.00022	0.01304	-0.23305	7.74177
m=4	1 2 3 4	-0.01108 -0.00217 0.00024 0.00152	0.00496 0.02850 0.01217 0.00576	11657.48	-23282.98	-23183.33	0.00021	0.01300	-0.25995	7.80331
Obs.							0.00021	0.01309	-0.12580	9.89054

The maximized log-likelihood increases markedly from the 2-state to 3-state HMM, whereas the increase from 3- to 4-state models is only marginal. The model selection criterion AIC clearly indicates the superiority of the 3- and 4-state models over the 2-state model. The likelihood ratio test cannot be used here for model identification because of the presence of unidentified nuisance parameters in the transition probabilities under the null hypothesis. Psaradakis and Spagnolo (2003) suggest using AIC as a heuristic measure to determine the number of regimes. Although the AIC also favours slightly the 4-state over the 3-state model, our main objective is ex ante forecasting for which a parsimonious model is generally preferred. We therefore choose the 3-state as our preferred model. Further, simulated data from the three models produce sample moments which when compared to the moments of the actual observations reveal that the 2-state model underestimates the tail areas of the returns distribution.

3 OPTION-IMPLIED INFORMATION IN THE RND

Our aim in this paper is to enquire if the RND contains forward-looking information that can be exploited to predict stock market regimes as well as returns distribution. We next outline the method employed to extract the RND from options on the S&P 500 index.

3.1 RND MOMENTS EXTRACTION USING THE GENERALIZED LAMBDA DISTRIBUTION

The extraction of the RND follows the theoretical result of Breeden and Litzenberger (1978). The price of a European call option is the discounted risk-neutral expected value[5] of the payoffs:

[10] $$C = e^{-rT} \int_K^\infty (S_T - K) f(S_T)\, dS_T$$

[5] Cox and Ross (1976)

where C is the value of a call option on a stock at time 0 with strike K and expiry T; r is the risk-free rate of interest; S_T denotes the price of the underlying on the expiry date of the option and $f(S_T)$ is the RND for S_T. Equation [10] provides the theoretical underpinning for extracting RND from call option prices. Breeden and Litzenberger (1978) show that the risk neutral distribution implied by European option prices is given by

[11] $$e^{-rT}\frac{\partial^2 C}{\partial K^2} = f(S_T).$$

The practical implementation of RND extraction faces several issues. Firstly, prices are available on a set of discrete traded strikes. To address this, we first apply the method of Shimko (1993). For each day in the sample, Shimko's approach first converts the observed closing European call and put[6] option prices to implied volatility (IV) equivalents using the Black-Scholes formula. A smooth spline function is then fitted to the IV-strike space, from which the fitted function of IVs which correspond to a dense set of strikes is obtained. Finite difference approximation of equation [11] then yields estimated densities at the dense set of strike prices.

Many methods have been proposed in the literature to extract the RND from option prices; thorough surveys of these methods are provided by Jackwerth (1999,2004) and Bliss and Panigirtzoglou (2002). We estimate the RND by drawing a random sample from the estimated discrete density distribution above and fit a Generalized Lambda Distribution[7] to the sample observations. The GLD is very flexible; with four parameters associated with the first four moments of the distribution, it accommodates many distributional shapes [Karian and Dudewicz (2000)], which is especially useful for addressing time-varying RND distributions. The GLD is commonly defined on its quantile function,

[12] $$Q(y) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2}$$

from which its probability density function can be derived

[6] Put option prices are first converted to call price equivalents, "equivalent calls" by put-call-parity. On each day in the sample, we discard put option prices on strikes for which no available equivalent call are traded. For each strike that an equivalent call price overlaps with that of its corresponding actual call, we perform a midpoint interpolation to obtain a single representative call price.

[7] By the well-known statistical result that cdf-transformed values follow a uniform distribution, the GLD parameters are estimated by a grid search over the parameter space such that the transformed sample matches that which is obtained from a uniform distribution.

[13] $$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3 - 1} + \lambda_4 (1-y)^{\lambda_4 - 1}}$$

where $0 \leq y \leq 1$, λ_1 and λ_2 are respectively parameters for location and scale; $\lambda_3$and λ_4 are the shape parameters for skewness and kurtosis respectively. From the extraction procedure described, a daily time series of the four RND moments are obtained.

A second complication is the time-to-maturity effect, in that changes in the uncertainty reflected in the RND decreases in part due to the approach of option expiry[8]. In order to study the factors influencing the uncertainty reflected by the RND, we need to eliminate the "time-to-maturity" effect by holding constant the expiry duration. Traded options, however, are available for a limited number of maturities. Previous studies which have empirically considered constant-maturity RNDs have mostly been in the context of computing Value-at-Risk (dubbed Economic Value-at-Risk[9], E-VaR) which requires well-defined time horizons. Clews et al. (2000) applies a cubic spline to interpolate the Black-Scholes implied volatility surface in delta space at a given time horizon, then constructs the implied volatilities of an hypothetical contract with a desired maturity, from which the RND is derived using the Breeden and Litzenberger (1978) result. We follow Clews et.al (2000) to compute constant 30-day maturity RNDs[10] for each observed day in the sample period.

3.2 MARKET PRICE DATA AND OPTION-IMPLIED INFORMATION

This study uses options data for S&P 500 Index (ticker symbol SPX, hereafter "SPI") obtained from Optionmetrics of WRDS database. The options data contain last trading date, expiry date, bid and ask prices, trading volume and strike prices. Daily data was obtained for the period from 4th January 1996 to 29 April 2011. Following existing literature on calculating option-implied distributions, several filters are imposed on the data prior to extracting the RNDs. First we eliminate all options with zero bid- or ask price and zero trading volume. Second, we incorporate options with the mid-point of the bid and ask prices greater than $0.50 and with the bid-ask spread not more than 30% of the average of bid and ask prices.

[8] Upon expiry, the representative investor rolls over to options with the next available maturity date, thereby increasing uncertainty in the new set of option prices.

[9] Ait-Sahalia and Lo (2000) estimated VaR using the RND for a fixed 126 day horizon, while Panigirtzoglou and Skiadopoulos (2004) computed 14-day E-VaR computations for 1, 3 and 6 months.

[10] Consistent with other authors which have similarly computed 30-day maturity RNDs. European S&P 500 index options are sufficiently liquid such that the interpolation approach of Clews et. al (2000) should reasonable calibrations of 30-day RNDs. The choice of 30-day maturity is also consistent with the methodology used to compute the VIX index.

Thirdly, only options with trading volume greater than 50 contracts are included. These last two filters are designed to avoid microstructure problems caused by far OTM options, large bid-ask spread and thinly traded options. Economic data used in the analyses below are obtained from databases in Datastream and Board of Governors of the Federal Reserve System. Three-month Treasury Bill rates, adjusted to maturity duration of the option, are used as proxy for the risk-free rate.

3.3 Disaggregating the RND into Long-Term and Short-Term components

One innovation we present in this paper is the disaggregation of the RND dynamics into its slow-moving "long term" (LT) dynamics and its fast-moving "short-term" (ST) dynamics. The motivation for this approach stems from several recent works[11] which have found the presence of long- and short-run components in market volatility risk. LT components are said to contain information in investors' expectations due to structural change driven by macroeconomic and/or business cycle factors, whereas ST components are believed to capture transient market skewness risk.

We consider a general stationary and invertible ARFIMA(p, d, q) representation of each RND moment

[14] $$\Phi(L)(1-L)^d x_t = \Theta(L)\epsilon_t,$$

$$\Phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i \ , \ \ \Theta(L) = 1 + \sum_{i=1}^{q} \vartheta_i L^i$$

where x_t denotes each RND moment series considered, p and q respectively denote the lag order of the short-memory autoregressive and moving average components, d is the long memory fractionally differencing parameter, $\Phi(L)$ and $\Theta(L)$ respectively denote the p^{th} and q^{th} order polynomial in the lag operator L, and ϵ_t is Gaussian white noise with variance σ_ϵ^2. Using [14] we can define the fractionally differenced series y_t

[15] $$y_t = (1-L)^d x_t, \qquad -1 < d < 1$$

which can be seen to contain only the short-term dynamics of x_t. It is simple to show that there is a short-memory ARMA(p, q) representation for the fractionally differenced moment, y_t

[11] See Adrian and Rosenberg (2008), Engle and Rangel (2008)

[16] $$\Phi(L)y_t = \Theta(L)\epsilon_t, \quad \text{or}$$

[17] $$\pi(L)y_t = \epsilon_t \quad \text{where } \pi(L) = \frac{\Phi(L)}{\Theta(L)} = 1 - \pi_1 L^1 - \pi_2 L^2 - \cdots, \quad \text{such that}$$

[18] $$y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \cdots + \epsilon_t.$$

Equation [18] therefore gives the autoregressive representation of y_t whose degree of short-memory is determined by the original AR and MA coefficients. Now it can be shown that a binomial expansion of the fractional differencing operator $(1-L)^d$ yields

[19] $$(1-L)^d x_t = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} x_i = y_t, \qquad d > -1$$

[20] $$x_t = (1-L)^{-d} y_t = y_t + \sum_{j=1}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} y_{t-i}, \qquad d < 1$$

Some algebra then shows that the process $z_t = x_t - y_t$, which contains the long-term dynamics of x_t, is given by

[21] $$z_t = x_t - y_t = y_t + \sum_{j=1}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} y_{t-i} - y_t = \sum_{j=1}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} y_{t-i}$$

where $\sum \left(\frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}\right)^2 < \infty$. The infinite order of the summation in equation [21] therefore can be viewed to contain the long-term dynamics in the RND moments. For future reference we refer to z_t, the series measuring long-term RND moments dynamics, as the RND-LT or RND-LT moments.

In an earlier working paper, Low (2013) [12], significant effects of structural change in the RND moments were found and and the existence of long-term dynamics using the KPSS (1992) and Lo's rescaled range tests were verified. We proceed to calibrate a long-memory parameter d for each RND moment series which characterizes the LT dynamics in the RND. The range of long memory estimation methods span a large literature[13], but the methods of Maximum Likelihood, Whittle and Robinson (1995b) are often used in financial applications for their computational simplicity and the invariance of their limiting distribution

[12] Details are omitted for brevity but are available upon request.
[13] Robinson (2003) provide detailed treatment on parametric estimation methods; similarly, nonparametric estimation methodology is described in Doukhan, Khezour and Lang (2002)

with respect to d. Table 1 reports the estimated d parameters computed on the sample from 4 January 1996 to 29 April 2011. All calibrated d parameters are indicative of significant memory[14] effects.

3.4 *MEASURES OF OPTION-IMPLIED INFORMATION FOR PREDICTING MARKET REGIMES*

In light of Section 5.3.2, we can proceed with several meas ures of option-implied information considered as covariates for the prediction of bull and bear markets. Having obtained the time-series of RNDs, the following measures of option-implied information are considered for predicting market regimes and returns distribution:

- Long-Term RND dynamics series: Based on long memory methods which are known to be a convenient mathematical measure for occasional breaks in time series, the RND-LT series is a measure of the slow-moving, long-term dynamics in the RND attributable to structural breaks. In periods of high volatility, lone-memory modelling captures a high degree of persistence, which is analogous to regime switching to a high volatility state. The RND-LT has been found to significantly relate to macroeconomic CPI. As a forward-looking measure of the slow-moving changes in the underlying S&P500 index, therefore, the RND-LT series may contain anticipatory information for significant changes in the underlying market regime.
- Short-Term RND series: Following earlier results which indicate different information content in the short-term RND dynamics from its long-term dynamics, we also consider the extent to which RND-ST series may be useful for the prediction of market regime and returns distributions.
- Higher order RND moments: Many of the aforementioned studies in Section 1 have alluded to the potential information in RND volatility and skewness being indicative of future asset prices. In addition to the disaggregated RND moment components, we consider also the entirety of the higher order RND moments as potential forward-looking information sources about future market prices.
- Tail quantiles of the RND: The tails of the RND represent extremes of investor expectations. They may be expected to contain forward-looking information as to the direction of market price

[14] Values of $-0.5 < d < 0.5$ indication stationary long memory; while $d > 0.5$ signals nonstationarity.

changes. We therefore consider RND quantiles, particularly focusing on tail quantiles as potential information measures in this study.

These RND-based variables are included as covariates in the HMM time in-homogenous model specification in [7], [8] and [9]. We therefore have a large number of HMM-RND models, of which each differ by the specification of covariates which capture the effects of option-implied information. In the model assessment, we include in addition to the HMM-TH and HMM-RND model types, returns distribution forecasts generated from a naïve model which forecasts one-step-ahead using each current day's realisations as the next day's predicted values. This addition allows us to gauge simply whether the HMM models in general can produce improved forecasts over a naïve benchmark. The aim of this comparison is not to discover the best performing forecasting model; rather it is to investigate whether stock market regimes exist. If the market goes through phases, we expect the HMM models to outperform the naive model. Table 2 presents a complete list of models we consider in our out-of-sample forecasting exercise and their respective abbreviations.

4 FORECASTING RETURNS AND STATE IDENTIFICATION IN HMMS: THEORY

To gauge the degree to which option-implied information is useful for making ex-ante market regime forecasts, we perform an out-of-sample forecasting exercise using the HMM-RND model. First, the theory underlying out-of-sample forecasting with HMMs is outlined below.

4.1 CONDITIONAL DISTRIBUTIONS

Let $\boldsymbol{R}^{(-t)}$ denote all observations of the daily market returns process at all times apart from time t. That is,

[22] $$\boldsymbol{R}^{(-t)} \equiv (r_1, \dots, r_{t-1}, r_{t+1}, \dots, r_T)$$

From the likelihood function of a HMM, it can be shown (see Zucchini and MacDonald, 2009), for $t = 2,3, \dots, T$ that the conditional distribution

[23] $$P\left(r_t = x \middle| \boldsymbol{R}^{(-t)} = \boldsymbol{r}^{(-t)}\right) = \frac{\pi \boldsymbol{P}(r_1)\boldsymbol{B}_2 \ldots \boldsymbol{B}_{t-1}\boldsymbol{\Gamma}\boldsymbol{P}(x)\boldsymbol{B}_{t+1} \ldots \boldsymbol{B}_T\boldsymbol{1}'}{\pi \boldsymbol{P}(r_1)\boldsymbol{B}_2 \ldots \boldsymbol{B}_{t-1}\boldsymbol{\Gamma}\boldsymbol{B}_{t+1} \ldots \boldsymbol{B}_T\boldsymbol{1}'} \propto \boldsymbol{\alpha}_{t-1}\boldsymbol{\Gamma}\boldsymbol{P}(x)\boldsymbol{\beta}_t'$$

where $\boldsymbol{\alpha}_t = \boldsymbol{\pi}\boldsymbol{P}(r_1)\boldsymbol{B}_2 \ldots \boldsymbol{B}_t$ are referred to as *forward probabilities* and $\boldsymbol{\beta}_t' = \boldsymbol{B}_{t+1} \ldots \boldsymbol{B}_T\boldsymbol{1}'$ are referred to as *backward probabilities* where $\boldsymbol{B}_{t+1} = \boldsymbol{\Gamma}\boldsymbol{P}(r_1)$.

The conditional distribution in [23] is essentially the ratio of two likelihoods of a HMM: the numerator is the likelihood of the observations where the observation r_t is replaced by x, and the denominator is the likelihood when observation r_t is treated as missing. Noting that $\boldsymbol{P}(x)$ is a diagonal matrix, the conditional probabilities in [23] can be expressed as mixtures of the m state-dependent probability distributions, such that

[24] $$P(r_t = x|\boldsymbol{R}^{(-t)} = \boldsymbol{r}^{(-t)}) = \sum_{i=1}^{m} w_i(t)p_i(t)$$

where $w_i(t)$ are the mixing probabilities which are functions of the observations $\boldsymbol{r}^{(-t)}$ as well as of the model parameters. The effects of covariates on this conditional distribution enter through the model parameters incorporated in [24].

5.3.3.2 Forecast distributions

The forecast distribution of a HMM is a particular type of conditional distribution. For discrete observations, the h- step ahead forecast distribution is computed, similar to the conditional distribution in [23], as a ratio of likelihoods:

[25] $$P\left(\boldsymbol{R}_{T+h} = x \middle| \boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\right) = \frac{\pi \boldsymbol{P}(r_1)\boldsymbol{B}_2\boldsymbol{B}_3 \ldots \boldsymbol{B}_T\boldsymbol{\Gamma}^{\boldsymbol{h}}\boldsymbol{P}(r)\boldsymbol{1}'}{\pi \boldsymbol{P}(r_1)\boldsymbol{B}_2\boldsymbol{B}_3 \ldots \boldsymbol{B}_T\boldsymbol{1}'} = \frac{\boldsymbol{\alpha}_T\boldsymbol{\Gamma}^{h}\boldsymbol{P}(r)\boldsymbol{1}'}{\boldsymbol{\alpha}_T\boldsymbol{1}'} = \boldsymbol{\varphi}_T\boldsymbol{\Gamma}^{h}\boldsymbol{P}(r)\boldsymbol{1}'$$

where we denote $\boldsymbol{\varphi}_T = \frac{\boldsymbol{\alpha}_T}{\boldsymbol{\alpha}_T\boldsymbol{1}'}$. Analogous to [24], the forecast distribution can be expressed as a mixture of the state-dependent probability distributions

[26] $$P\left(\boldsymbol{R}_{T+h} = x \middle| \boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\right) = \sum_{i=1}^{m} \xi_i(h)p_i(x)$$

where the weights $\xi_i(h)$ are given by vector multiplication of the elements in $\boldsymbol{\varphi}_T\boldsymbol{\Gamma}^{h}$. Again, option-implied information covariates enter the forecast distributions through the model parameters incorporated in [26].

4.2 DECODING: IDENTIFICATION OF STATES IN A HMM

Of particular importance in assessing bull and bear market model forecast performance is the determination of states within the Markov chain which are most likely to have generated the observed financial time series data. The identification of the state that is most likely to have generated the observable time series at time t is known as *local decoding* [cf. Zucchini and MacDonald (2009)].

The forward and backward probabilities $\boldsymbol{\alpha}_t$ and $\boldsymbol{\beta}_t'$ defined above are related by the following result

[27] $$\alpha_t(i)\beta_t(i) = P\left(\boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}, S_t = i\right)$$

We can then compute the conditional distribution of the market state S_t based on the observed time series for $i = 1,2, \dots, m$ as follows

[28] $$P\left(S_t = i | \boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\right) = \frac{P\left(\boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}, S_t = i\right)}{P\left(\boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\right)} = \frac{\alpha_t(i)\beta_t(i)}{L_T(\boldsymbol{v})}$$

where $L_T(\boldsymbol{v})$ is the likelihood function defined in [6][15]. Given observations for $t \in \{1, \dots, T\}$, we can obtain the distribution of the state S_t. The most likely state, i_t^*, is then given by

[29] $$i_t^* = \arg\max_{i=1,\dots,m} P\left(S_t = i | \boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\right)$$

This approach, commonly known as local decoding, gains its reference from separately obtaining the most probable state for each t which maximises the conditional probability in [28].

4.2.1 *Predicting Market State*

For a given point in time, t, we are also interested in predicting future states of the market, based on the observed financial time series up to time T; that is, we compute the conditional distribution of the state S_t for $t > T$.

[15] Zucchini and Macdonald (2009) suggest a method of scaling the likelihood to address the potential problem of numerical underflow when maximizing the likelihood function.

In the context of financial returns, suppose we observe returns $r_1, \dots, r_T$ up to time T. Then for $t > T$, we have

[30] $$L(\boldsymbol{v})P\big(S_t = i|\boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\big) = \boldsymbol{\alpha}_T\boldsymbol{\Gamma}^{t-T}(,i)$$

where $\boldsymbol{\Gamma}^{t-T}(,i)$ is the ith column of the $\boldsymbol{\Gamma}^{t-T}$ matrix. Rearranging and restating [30] in terms of $t = T + h$, we have the conditional distribution of the state S_t given by

[31] $$P\big(S_{t+h} = i|\boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\big) = \frac{\boldsymbol{\alpha}_T\boldsymbol{\Gamma}^{t-T}(,i)}{L_T(\boldsymbol{v})} = \boldsymbol{\varphi}_T\boldsymbol{\Gamma}^{h}(,i)$$

From [31], option-implied information (again) enters the market state prediction exercise through the transition probability matrix, $\boldsymbol{\Gamma}^h$, and the state dependent distribution parameters in $\boldsymbol{\varphi}_T$. The forecasted state is determined, analogous to equation [29], as the 'most likely' state based on the forecasted state probabilities in [31], and is given by

[32] $$i^*_{t+h} = \arg\max_{i=1,\dots,m} P\big(S_{t+h} = i|\boldsymbol{R}^{(\boldsymbol{T})} = \boldsymbol{r}^{(\boldsymbol{T})}\big).$$

For both the asset returns and market states, the forecasted distributions approach steady states as the forecast horizon increases when $\boldsymbol{\Gamma}$ is stationary.

5 *Using Option-Implied Information for Out-of-Sample Forecasting in HMMs - in Practice*

Our objective is to investigate the extent to which option-implied information is useful for forecasting financial bull and bear markets. Recall that the analysis proceeds in several stages: we first decode the bull and bear market states in the sample. We then estimate two main types of HMMs – a time homogenous HMM (HMM-TH) as a benchmark to be compared against time inhomogenous HMMs (HMM-RNDs) whose state transition properties are specified as functions of covariates defined on RND variables. We compute out-of-sample forecasts of the asset returns and returns distribution using both types HMM models considered and compare their relative forecast performance. Forecast performance is assessed on the forecast returns distributions and the predicted states. Forecast error metrics are computed based on the forecasted returns distributions by each considered model, and the returns realisations. Furthermore, the realised bull and bear market states in the sample are identified by the process of decoding described in Section 4.2. . For each HMM model considered, the predicted state is that state for which model assigns the highest probability of occurring. The competing models are then compared on the basis of their state predictions vis-a-vis the decoded 'actual' states. The predicted states are then compared against the decoded ‘actual’ states for each considered model, and their relative performance assessed.

5.1 Forecasting Bull and Bear Markets using HMMs with Option-Implied Information

From the SPI returns data, we apply the decoding process described in Section 4.2 to identify the market states which occurred over the sample period. A comparative study then enables us to gauge the degree to which option-implied information is useful for bull and bear market regime prediction. Two forms of the HMM are specified: the first is a benchmark time homogenous HMM (HMM-TH) where the transition probabilities remain constant over the sample period. The second form allows for covariates in the HMM (HMM-RNDs). In one instance, covariates are related to the transition probabilities; in another, we allow covariates to affect the parameters of the state-dependent distribution. As mentioned, we consider many RND-based measures of information, specified as the covariates in these models.

We perform a comparative analysis between the HMM models as follows. The last 1000 observations in the sample, from 5 May 2007 to 29 April 2011, are reserved for out-of-sample[16] forecasting, particularly since it contains the period of the Global Finance Crisis which presents a suitable testing ground for which option-implied information may be useful. Both the HMM-TH and HMM-RND models are estimated using sample data over an augmenting window; as we proceed through the forecast period, each

[16] A limitation of the current sample size is the difficulty of consistently estimating the fractional differencing parameter over a reduced portion of the sample. Hence in the forecasting exercise we use RND-LT and RND-ST moments derived from fractional differencing parameters estimated over the *entire* sample, rather than re-estimating a separate fractional differencing on a smaller set of within-sample data for out-of-sample forecasts. In this sense, the current exercise has the drawback in that the RND-LT and RND-ST series are not entirely out-of-sample in nature; to address this, we repeat the out-of-sample forecasting exercise in this paper using RND-LT and RND-ST series derived from alternative estimates of the long memory parameter d and find the results to be consistent with those presented here. This work is available upon request.

subsequent forecast is generated by HMMs estimated on an increasing window size[17], thereby conditioning on increasing information over time.

The HMM-TH and HMM-RND models are compared on the following basis:

- For the fitted HMM-RNDs, we first study the factors which are significant in affecting the state transition probabilities. For each day in the sample, in-sample one-step-ahead forecasts are computed of the expected daily returns on the SPI, the SPI returns distribution, and the market state.
- **Returns prediction**: Dynamic one-step-ahead point and interval forecasts are computed for returns on the SPI. For assessing point forecasts each day in the forecast period, forecast errors are computed as the difference between the realized return and forecasted expected return, for comparing the HMM-TH against HMM-RND models. For assessing interval forecasts, forecasted confidence intervals for the SPI returns are constructed for each of the HMM-TH and HMM-RND models, based on the forecasted SPI returns distribution each day in the forecast period. Over the entire sample, the percentage of days on which the actual asset return falls outside of the forecasted intervals (interval violation rates) can then be obtained for both models. We then test for significant differences between the forecast interval violation rates for the two models using a ratio test.
- **State prediction**: For each model, the predicted market state can be compared against the decoded market phase on each day in the sample. Each predicted observation in the time series of predicted states is classified as a correct or incorrect market state prediction and represented by an indicator variable. We then test for significant differences between the percentage erroneous predictions made by the two model types to assess if the HMM-RND model significantly outperforms the HMM-TH. In addition, we also jointly compare the competing models on the basis of the absolute errors between forecasted and 'actual' probabilities of the decoded states.

[17] For example, the first (returns distribution and state) day’s forecast in the out-of-sample period is computed using the HMMs estimated on data from 4 January 1996 to 2 May 2007. Moving forward, the second day’s forecast is generated by a HMM estimated on data from 4 January 1996 to 3 May 2007, that is, the estimation window has increased by one day.

5.2 *Assessing Relative Model Performance on Returns Prediction: The Model Confidence Set*

In the context of model comparison, two main considerations arise. Firstly, there are a large number of HMM-RND model types to be assessed against the HMM-TH model, which effectively is a nested version of the former. Secondly, the models may be compared both in terms of their point and interval forecasts, since both forecasts of the expected returns and expected returns distribution are obtained. Point and interval forecast error metrics yield separate information in that while point forecasts errors are informative about the degree of point forecast accuracy, interval forecast error metrics are informative about the confidence with which point forecasts can be considered.

Pertaining to forecasts of returns distribution, we assess the competing models using the model confidence set (MCS) approach of Hansen, Lunde, and Nason (2011), who recently developed this methodology for assessing multiple models for selecting a set of 'best' models. The method has the appeal of being applicable to more general comparison between objects and is not limited to model comparison exercises. The key feature of the MCS is its acknowledgement of information content in the data upon which it is applied; uninformative data metrics therefore produce a larger set of 'best' models than informative metrics. The MCS therefore allows us to compare the forecast performance of the models, 1) jointly, and 2) across different error metrics which can be separately informative about the point and interval forecast performance. Further technical details of the MCS are set out in Appendix A2.

Since the outcome of the MCS depends on the information content of the error metrics employed, we consider a range of standard and alternative error metrics for MCS assessment of our competing models. Standard metrics for point forecasts we apply include the mean absolute error (MAE), mean squared error (MSE). As an alternative, their heteroscedasticity-adjusted counterparts, namely the mean absolute percentage error (MAPE) and proportional mean squared errors (MSE-prop) metrics, are also considered. We also adapt interval distance measures proposed by Yaniv (1997) and more recently, Bratu (2013). Yaniv (1997) defines an accuracy measure known as a normalized error which divides estimation error by the width of the $(1-\alpha)\%$ confidence interval. Here, we adapt from Yaniv (1997) in a metric which scales the mean absolute forecast error by its respective forecast intervals (ABS-SCALED ERR). Along the same vein, Bratu (2013) suggests that distances between the realization and the lower/upper interval limit or the interval centre can be informative about forecast interval accuracy. For a given forecast interval, denote

[33] $$d_1 = realization - lower\ limit,$$

$$d_2 = realization - upper\ limit,$$

$$d_3 = realization - interval\ centre$$

We can then compute an average of the absolute values of d_1, d_2 and d_3 as another measure of interval forecast accuracy (ABSDIST). A lower distance indicates a superior forecast interval. Finally on the same principal, the width of the forecast interval is also considered as a measure of interval forecast performance in the sense that a wider confidence interval gives rise to greater uncertainty (hence less confidence) in the point forecast. In other words, an investor may give preference to a biased forecast which has lower variability. This measure of interval forecast accuracy has semblance of bias and variance trade-off. A biased interval forecast where $d_3 \neq 0$ may be preferred to another when $d_3 = 0$ if the width of the first interval is sufficiently smaller than the second interval to more than compensate for the biasedness.

5. 3 ASSESSING RELATIVE MODEL PERFORMANCE: COVERAGE TESTS AND STATE PREDICTION

In addition of comparing point forecast of out-of-sample returns, we also compare the relative performance of HMM-RND and HMM-TH models by their interval forecasts of the SPI returns and their state predictions[18]. For each model we form 90% confidence interval of out-of-sample forecasts of SPI returns and count the proportion of times that the confidence interval is violated by the actual return. We then test for significance of the difference in proportion of non-violation of each HMM-RND model with the HMM-TH model, employing a chi-squared test of equality of proportion. Table 9 summarizes the results: an asterisk indicates significant difference at the 10% level.

For state prediction, we compare the proportions of correct state predictions by each of the HMM-RND model with the HMM-TH model using the same testing procedure. The results of these tests are also summarized in Table 9.

[18] Note that the metrics in this section have not been applied in the form of the MCS since they are binary indicator measurements of forecast error which in the MCS setting are not sufficiently informative in the context of the MCS, regarding differences in model forecast performance.

6 RESULTS

6.1 THREE-STATE HMM ESTIMATES

Figure 1 displays the time series of SPI returns and prices over the entire sample period with overlaid with indicated regions which identify the decoded states of the market. The decoded states can be interpreted in a similar fashion to that in Guidolin and Timmermann (2005); State 1 is characterised by negative average daily returns (μ_1 =-0.0020344) and the highest volatility (σ_1=0.02826681) of the three states in the model, in line with the interpretation of a bear state. Conversely, State 3 is associated with positive daily returns (μ_3 =0.001084) and the lowest volatility (σ_3 =0.006112) of the three states. State 2 has positive average daily return (μ_2 = 0.000045) which is close to zero, and a mild average volatility (σ_2 = 0.012000) value not larger than that during the State 1 bear, but not less than that during the State 3 bull.

Table 3 displays model estimates of the transition probabilities and average transition probabilities for the three-state time homogenous HMM and time-inhomogeneous HMM-RND respectively. The estimated transition probabilities in both types of models clearly indicate strong persistence within state. Fitted parameter estimates on the log-linear effects of covariate variables on the transition intensities are displayed in Table 4 for selected models. These estimated parameter values on the covariates in the HMM-RND models should be interpreted with care due to multicollinearity arising from correlations between covariate variables; however we may still comment generally on their statistical significance. The results in Table 4 indicate that the higher order moments significantly affect transition intensities between states. Note that the transition intensities define instantaneous transitions from one state to another. The observed data, on the other hand are snapshot of the process at specific times. Hence, it is possible to observe a zero transition intensity between states i and j and yet observe non-zero transition probability because the process could move from state i to some other state k before going to j during the intervening period. Separately examining the fit of long term and short-term higher order moments effects, we find that it is the long term higher order moments effects which appear significant in influencing transition probabilities between states.

6.2 *Out-of-Sample Forecast Results*

6.2.1 MCS Results

The results of the MCS performed on the out-of-sample, one-step-ahead forecast error metrics over the entire forecast period are displayed in Table 5. Tables 6, 7 and 8 analogously display the MCS results on forecast error metrics from sub-segments of the out-of-sample period; Tables 6 and 7 respectively contain results from the first and second halves of the forecast period, while Table 8 contains MCS results collected over the GFC period. Across all considered forecast periods and significance levels, it is clear that the HMM-TH and HMM-RND model types are virtually indistinguishable by metrics that evaluate their point forecasts. The MAE, MSE, MAPE and MSE-PROP point forecast error criterion are uninformative about differences between the models, resulting in all or nearly all models remaining in the MCS, apart from the naïve model. This lack of information in the HMM-RND point forecasts may be attributable to the noise in daily returns which prevents any incremental information in RND moments to significantly improve existing forecasts in a reliable HMM-TH model.

In contrast, the error criteria which account for forecast interval widths provide more interesting results. Compared to the case of point forecast error criterion, the MCS is considerably reduced when the testing is performed using ABS-DIST, ABS-SCALED ERR and CIWIDTH criterion. When considering the full forecast period, the LT-MEANVOL model is the only model remaining in the MCS, hence being identified as the best model in the set. Examining the results for the forecast sub-periods, Tables 6, 7 and 8 show that LT-MEANVOL model is consistently in the MCS. Other models types which are found in the MCS include LT-VOL, LT-VOLKURT as well as the Q01, Q025, Q75 and Q95 models. We note that the RND-TH model has been dropped from the MCS in all cases of the interval forecast error-based criteria. The results therefore suggest that the HMM-RND models which incorporate the long-term effects of higher order moments or tail quantiles significantly improve SPI returns interval forecasts over the HMM-TH benchmark case. This result is consistent with the intuition that information about the low-frequency dynamics in investors' expectations (as captured by the LT moments) as well as information in the RND tails (as captured in tail quantiles) should provide improved confidence about next period's distribution of returns.

6.2.2 Prediction Interval Coverage Ratio Test Results

The first three columns in Table 9 contain the results of 90% prediction interval violations for all the considered models. Column 1 ("nvio") reflects number of violations recorded over the entire out-of-sample forecast period, while columns 2 ("GFC) and 3 ("post-GFC") reflect the number of violations recorded in the first and second halves of the forecast period respectively. The overall unsatisfactory results of 15% (which exceeds the ideal of 10% for 90% forecast intervals) violations on average across models is substantially driven by the effects of the GFC period. Around the GFC period, the HMM models struggled on the whole to produce prediction intervals which accurately contained the corresponding level of violations.

While the total number of prediction interval violations over the entire sample period is found to be statistically indistinguishable between the HMM-TH and some HMM-RND cases, the HMM-RND does significantly outperform the HMM-TH for several covariate combinations including VOL and LT-MEANVOL. Turning to sub-period forecasts, once again, VOL, LT-MEANVOL and some extreme quantiles - Q025, Q95, Q01-025-975-99 exhibit significantly improved violation rates in the first half of the forecast period than the HMM-TH model. In contrast, the violation rates during the second half of the forecast period are found to be indistinguishable between the HMM-TH and HMM-RND models. Consistent with the MCS results, these results suggest that LT-MEANVOL and tail quantiles may contain forward-looking information which is particularly useful during turbulent market periods. .

6.2.3 Results of State Transition Forecasts

Columns 4 to 12 of Table 9 summarise the state transition predictions results across all considered models. Header labels in Table 9 "snfm" denote the number of decoded states n where $n = 1,2,3$ which were predicted by a model as state $m, m = 1,2,3$. The columns where $n = m$ therefore represent correctly predicted market regimes. Asterisks denote significant difference of state prediction produced by a HMM-RND compared to the HMM-TH model benchmark. The results indicate clearly that the RND contains significant information for predicting state changes. The number of correctly predicted states (s1f1, s2f2, s3f3) by nearly all HMM-RND models is significantly larger than for the HMM-TH model; the HMM-TH model tends to underestimate the number of incidence of State 1, correctly predicting only 20% of the 189 decoded instances of State 1. One particularly notable error of HMM-TH model is forecasting 15% of the State 3 as State 1. These two are the extreme states and a high percentage of error in misclassification indicates poor performance of the homogeneous model in state prediction. In contrast, considering the HMM-RND model there are only two occurrences of misclassifying State 3 as State 1. Similarly, the incidences of misclassifying State 1 as State 3 are relatively rare for HMM-RND models compared to the HMM-TH model. Of the 1000 out-of-sample state forecasts, the RND-TH model produces 42.6% correctly forecasted state transitions. The HMM-RND models, on the other hand, average 81.2% correctly forecast state transitions, which is nearly a 100% improvement over the HMM-TH benchmark. The results in this section show that when the state transition probabilities are related to option-implied information the HMM models improve markedly in predicting changes in market regimes. These results support the hypothesis that the RND contains forward-looking information about market states.

Table 10 shows the results of the MCS analysis of MAE between the predicted and 'actual' probabilities of the decoded states. One clear pattern that emerges from the results is that tail quantiles of the RND contain useful information for predicting change in market regime. For the full out-of-sample period, as well as the various sub-periods, including the period of the GFC, HMM-RND models with tail quantiles - 1%, 95%,97.5% and 99%, remain in the model confidence set. More models remain in the model confidence set during the period of the GFC compared to the other sub-periods, which reflects the greater uncertainty of the market during the GFC, and therefore renders the MAE of predicted state probabilities less informative in distinguishing the models.

7 ALTERNATIVE HMM-RND MODEL SPECIFICATION

The analysis in Section 5 is repeated on a variation of the HMM-RND model types; namely, as an alternative to covariate-dependent transition probabilities between states, we allow the specification of the state-dependent distribution to depend on covariates instead. The general HMM-RND then takes the form given by equations [1], [2], [8] and [9] above.

7.1.1 MCS results

Analogous to Section 6.2.1, the results of the MCS performed on the out-of-sample forecast error metrics over the entire forecast period are displayed in Table 11. Tables 12 and 13 respectively contain results from the first and second halves of the forecast period, while Table 14 contains MCS results collected over the GFC period. For the point forecast error metrics, the earlier result in Section 6.2.1 is repeated in that all HMM models are generally indistinguishable, where the MCS is left containing most (in the case of MAE and MSE criterion), if not all (MAPE and MSE-PROP) considered models. This is true of the full forecast period and its sub-periods.

For error metrics based on prediction interval width, similar to the MCS results in Section 6.2.1, the MCS drops the HMM-TH model, retaining a reduced set of HMM-RND models. In contrast with earlier results however, the MCS in these results contains HMM-RND models with different covariates. This is consistent with our having allowed covariates to directly influence the state distribution parameters instead of indirectly through the between-state transition probabilities. For the full forecast period, ABSDIST and CIWIDTH error criterion result in the selection of HMM-RND with Q01-Q025-Q05 as covariates. The ABS-SCALED ERR criterion selects models with VOLSKEW, SKEWKURT, VOLSKEWKURT, LT-VOLKURT, LT-SKEWKURT and Q10 covariates. In both the first half of the forecast period and GFC sub-period, the HMM-RND model with ST-VOLKURT covariates emerges frequently in the MCS; however, during the second half of the forecast period, the prediction interval width error criterion select HMM-RND models with the following covariate combinations: VOLSKEWKURT, LT-KURT, LT-MEANVOL, LT-VOLKURT, Q10 and Q01-025-05.

These results are consistent with earlier results in that the RND moments contain significant information for interval forecasting of future SPI returns, while such information is not perceived for point forecasts of future returns. In addition, the results suggest that different components of the RND embody information specific to different sub-periods. During the segments of the forecast period which covered the period of the GFC, incorporating RND-ST volatility and kurtosis as covariates in the HMM was found to produce the best interval forecasts; in contrast, post GFC, the HMM models with RND-LT higher order moments covariates were among the best performing models for interval forecasts.

We have considered HMM models with covariates that affect the transition probabilities and mean of state-dependent distributions separately. Our aim is to investigate whether different aspects of the RND contain forward-looking information for predicting changes in market states and mean of state-dependent returns distribution. Generally, our results suggest that the long-term moments and tail quantiles of RND are useful for predicting state transitions, while the short-term moments and extreme quantiles affect state-dependent returns distribution.

7.1.2 Prediction Interval Coverage Ratio Test Results

The coverage ratio tests results are displayed in Table 15 analogously to the results in Section 6.2.2. We note several key features – most obviously, the coverage ratios under this HMM model set up significantly deteriorate relative to the former case where covariates were included in the state transition probabilities, but not directly in the state distribution parameter specification. This is particularly true of violation rates in the first half of the forecast period which contained the GFC window. During this sub-period, many HMM-RND models, particularly with RND-ST covariates, performed worse than the RND-TH benchmark, indicating that RND information when specified in the state distribution acts to deteriorate return distribution forecasts during this turbulent time. The second half of the forecast period produced more similar results to Section 6.2.2; a few RND-LT covariates were found to produce lower violation rates than the RND-TH model, though insignificantly so. Overall during this second half of the forecast period, HMM-RND models were found to be mostly indistinguishable from the HMM-TH benchmark.

The suggestion from results overall would be that generally the RND-LT moments contain significant information which can improve forecasts during normal market conditions, particularly when this information is incorporated in the transition probabilities of the model. However, from the generally unsavoury forecast performance from the other HMM-RND models with covariates specified in the state distribution - we gather that the value of option-implied information is less reliable during times of market crises for forecasting the underlying asset returns distributions.

7.1.3 Results of State Transition Forecasts

Similar to Section 6.1.3, Table 15 reports a summary of state transition predictions results across all considered models, with significant indicators analogous to Table 9. The results here mirror earlier results in Section 6.1.3 in suggesting that RND contains significant information for predicting state changes. In this set of results, the HMM-RND models average 60.8% correctly forecasted state transitions, still substantially improved over the HMM-TH benchmark of 42.6%, although not as favourable as the case of HMM-RND models with transition probability covariates. This is to be expected, since in this model set up, we have allowed covariates to influence the state distribution parameters directly, not but the transition probabilities between states. In light of the results in Section 7.1.2, we conclude that RND information is less obviously useful in the context of predicting future returns distribution as compared to forecasting changes in market state.

Model confidence set analysis of MAE between predicted and actual probabilities of decoded states is presented in Table 16. Although the covariates enter the HMM models through the means of the state-dependent distributions, rather than through the transition probabilities, they are still informative in distinguishing the models' performance in state prediction. For the full out of sample period, as well as the various sub-periods, not more than two models remain in the model confidence set. Once again, extreme quantiles - 1%, 2.5% and 5%, and tail areas - kurtosis, of the RND are informative.

8 CONCLUSION

In this paper we probed the degree of forward looking information in the RND using Hidden Markov Models. Previous literature recommends the use of HMMs for forecasting market regime changes and asset price distributions. We proffer that if option prices truly contain forward-looking information about the future prices of the underlying asset, the incorporation of RND-based information into the specification of HMMs for predicting market price regime should significantly improve market return forecasts over a benchmark model.

Following earlier literature in a three-state HMM set up, we consider a large range of HMM specifications, incorporating different types of RND-based information to study the potential information specific to different components of the RND. Namely, we disaggregated the RND moments into long- and short-run components using long memory methods, and separately examined their influence on the transition probabilities and state distributions when specified in the HMM model for S&P 500 index returns. We also considered as covariates a number of quantiles from the RND to see if the quintile distributions may be informative for forecasting future market regime changes and returns distributions. Out-of-sample forecasts were generated from each considered model, and a range of their error metrics were assessed to gauge the degree to which RND-based information can significantly improve out-of-sample forecasts.

The results indicate that the RND contains information most significantly for forecasting changes in market regime. All HMM-RND model specifications significantly outperform the benchmark HMM-TH in correctly forecasting the state of the market one-step-ahead. While the models under consideration are mostly indistinguishable when assessed on point forecast error metrics, the HMM-RND models are instead found to significantly improve one-step-ahead interval forecasts. In particular, the HMM-RND models with covariates based on the RND-LT moments appear to produce the most significant improvements in the entire out-of-sample forecast period. Examining sub-segments of the out-of-sample forecast period, we find that the inclusion of RND-LT moments in the HMM specification improves forecast performance during the period of the GFC, relative to the benchmark HMM-TH model.

This paper's results augment previous research which did not consider information in the entire risk neutral distribution of option prices in investigating their forward-looking potential. Our contribution highlights the need to recognise the separate sources of information which can be retrieved from options prices and provides a practical approach through long memory methods. To further assess the economic value of the information content embodied in the RND, future work entails testing the profitability of trading strategies which time market regimes for investment according to on the results in this paper.

TABLES AND FIGURES

Figure 1: Decoded states of the S&P 500 Index, January 1996 to December 2010

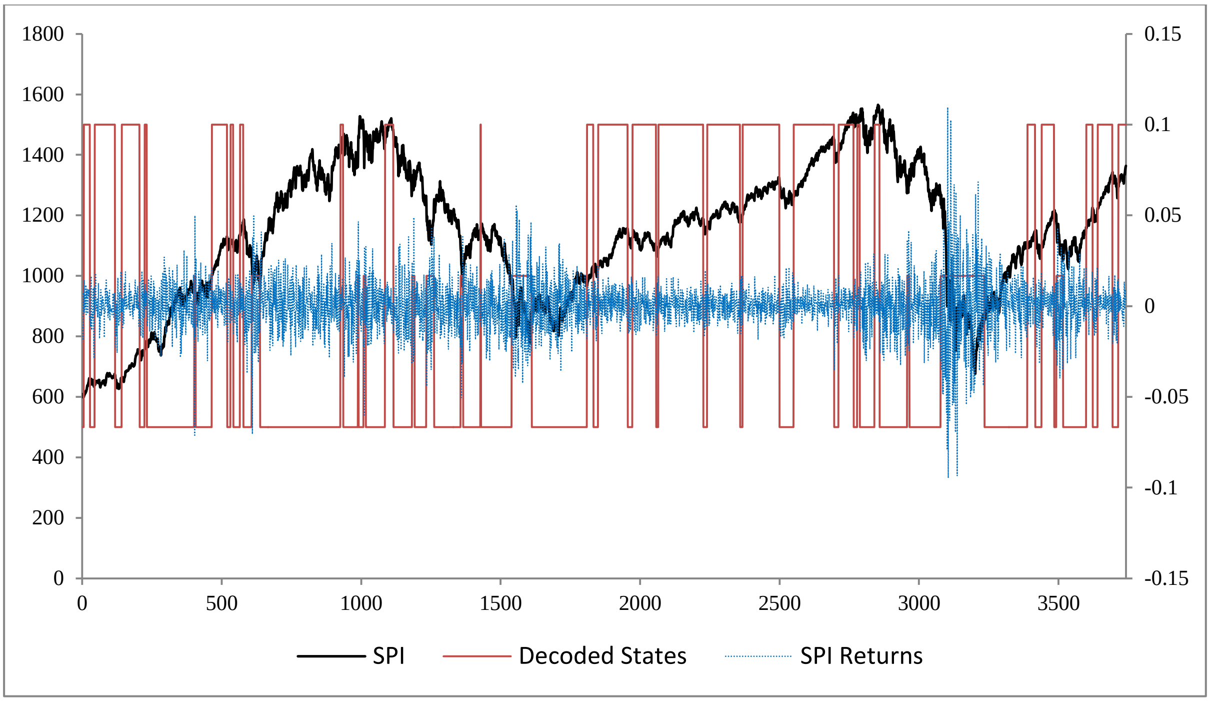

TABLE 1: TESTS FOR LONG MEMORY AND ESTIMATED FRACTIONAL DIFFERENCE PARAMETERS

	Mean	StdDev	Skew	Kurt
Lo's R/S	2.0883 **	3.8236 ***	2.8497 ***	3.1973 ***
KPSS	0.2568	1.0432 ***	2.1059 ***	2.0122 ***
Robinson's d (1994)	0.3776	0.4831	0.4598	0.4505
Robinson's d (1995)	0.5840	0.5396	0.3303	0.2936
Whittle d	0.1260	0.3761	0.4999	0.3797
MLE d	0.4994	0.4995	0.4111	0.3774

Table 1: The first two rows display test statistics from the Lo (1991)'s modified Rescaled range test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test respectively. Significance codes: ***, ** and * denote significant long-range dependence at the 1%, 5% and 10% levels respectively. The remaining rows displays estimates of the fractional differencing parameter, d.

TABLE 2: LIST OF HMM MODELS CONSIDERED IN THE OUT-OF-SAMPLE FORECAST ANALYSIS

KEY TO MODEL ABBREVIATIONS	
RND-TH	Time-homogenous HMM
MEAN	Time- inhomogeneous HMM, RND mean covariate
VOL	Time- inhomogeneous HMM, RND volatility covariate
SKEW	Time- inhomogeneous HMM, RND skewness covariate
KURT	Time- inhomogeneous HMM, RND kurtosis covariate
MEANVOL	Time- inhomogeneous HMM, RND mean and volatility covariates
VOLSKEW	Time- inhomogeneous HMM, RND volatility and skewness covariates
VOLKURT	Time- inhomogeneous HMM, RND volatility and kurtosis covariates
SKEWKURT	Time- inhomogeneous HMM, RND skewness and kurtosis covariates
VOLSKEWKURT	Time- inhomogeneous HMM, RND volatility, skewness and kurtosis covariates
LT-MEAN	Time- inhomogeneous HMM, RND-LT mean covariate
LT-VOL	Time- inhomogeneous HMM, RND-LT volatility covariate
LT-SKEW	Time- inhomogeneous HMM, RND-LT skewness covariate
LT-KURT	Time- inhomogeneous HMM, RND-LT kurtosis covariates
LT-MEANVOL	Time- inhomogeneous HMM, RND-LT mean and RND LT-volaility covariates
LT-VOLSKEW	Time- inhomogeneous HMM, RND-LT volatility and RND LT-skewness covariates
LT-VOLKURT	Time- inhomogeneous HMM, RND-LT volatility and RND LT-kurtosis covariates
LT-SKEWKURT	Time- inhomogeneous HMM, RND-LT skewness and RND LT-kurtosis covariates
LT-VOLSKEWKURT	Time- inhomogeneous HMM, RND-LT volatility, RND-LT skewness and RND LT-kurtosis covariates
ST-MEAN	Time- inhomogeneous HMM, RND-ST mean covariate
ST-VOL	Time- inhomogeneous HMM, RND-ST volatility covariate
ST-SKEW	Time- inhomogeneous HMM, RND-ST skewness covariate
ST-KURT	Time- inhomogeneous HMM, RND-ST kurtosis covariates
ST-MEANVOL	Time- inhomogeneous HMM, RND-ST mean and RND ST-volaility covariates
ST-VOLSKEW	Time- inhomogeneous HMM, RND-ST volatility and RND ST-skewness covariates
ST-VOLKURT	Time- inhomogeneous HMM, RND-ST volatility and RND ST-kurtosis covariates
ST-SKEWKURT	Time- inhomogeneous HMM, RND-ST skewness and RND ST-kurtosis covariates
ST-VOLSKEWKURT	Time- inhomogeneous HMM, RND-ST volatility, RND-ST skewness and RND ST-kurtosis covariates
Q01	Time- inhomogeneous HMM, 1% RND quantile covariate
Q025	Time- inhomogeneous HMM, 2.5% RND quantile covariate
Q05	Time- inhomogeneous HMM, 5% RND quantile covariate
Q10	Time- inhomogeneous HMM, 10% RND quantile covariate
Q25	Time- inhomogeneous HMM, 25% RND quantile covariate
Q40	Time- inhomogeneous HMM, 40% RND quantile covariate
Q50	Time- inhomogeneous HMM, 50% RND quantile covariate
Q60	Time- inhomogeneous HMM, 60% RND quantile covariate
Q75	Time- inhomogeneous HMM, 75% RND quantile covariate

Q85	Time- inhomogeneous HMM, 85% RND quantile covariate
Q90	Time- inhomogeneous HMM, 90% RND quantile covariate
Q95	Time- inhomogeneous HMM, 95% RND quantile covariate
Q975	Time- inhomogeneous HMM, 97.5% RND quantile covariate
Q99	Time- inhomogeneous HMM, 99% RND quantile covariate
Q01-Q99	Time- inhomogeneous HMM, 1% and 99% RND quantile covariates
Q025-Q975	Time- inhomogeneous HMM, 2.5% and 975% RND quantile covariates
Q01-Q025-Q05	Time- inhomogeneous HMM, 1%, 2.5% and 5% RND quantile covariates
Q95-Q975-Q99	Time- inhomogeneous HMM, 95%, 97.5% and 99% RND quantile covariates
Q01-Q025-Q975-Q99	Time- inhomogeneous HMM, 1%, 2.5%, 97.5% and 99% RND quantile covariates

TABLE 3: ESTIMATED TRANSITION PROBABILITIES AND AVERAGE TRANSITION PROBABILITIES FOR THE THREE-STATE TIME- HOMOGENOUS HMM AND TIME-INHOMOGENEOUS HMM-RNDS

S1F1	S1F2	S1F3	S2F1	S2F2	S2F3	S3F1	S3F2	S3F3
0.9656	0.0344	0.0000	0.0063	0.9794	0.0144	0.0000	0.0211	0.9789
0.8311	0.1025	0.0664	0.0171	0.9823	0.0006	0.0177	0.0011	0.9813
0.3402	0.3768	0.2829	0.0094	0.9756	0.0150	0.1942	0.0195	0.7863
0.7343	0.1642	0.1015	0.0161	0.9827	0.0013	0.0149	0.0030	0.9821
0.6690	0.2359	0.0951	0.0147	0.9825	0.0028	0.0143	0.0043	0.9814
0.3316	0.3741	0.2942	0.0087	0.9807	0.0106	0.0148	0.0093	0.9759
0.5702	0.2900	0.1398	0.0191	0.9799	0.0009	0.0151	0.0052	0.9797
0.3411	0.4365	0.2224	0.0098	0.9837	0.0065	0.1975	0.0171	0.7854
0.6680	0.2441	0.0879	0.0149	0.9810	0.0041	0.0141	0.0058	0.9801
0.6588	0.1230	0.2183	0.0137	0.9826	0.0037	0.0147	0.0028	0.9826
0.8271	0.1091	0.0637	0.0171	0.9823	0.0006	0.0174	0.0010	0.9816
0.3657	0.2971	0.3372	0.0151	0.9819	0.0030	0.2206	0.0207	0.7587
0.7565	0.1473	0.0962	0.0159	0.9827	0.0014	0.0147	0.0033	0.9819
0.6084	0.1396	0.2520	0.0140	0.9813	0.0047	0.0157	0.0074	0.9769
0.4186	0.2245	0.3569	0.0241	0.9736	0.0022	0.0095	0.0046	0.9859
0.5465	0.3875	0.0660	0.0134	0.9855	0.0011	0.0163	0.0054	0.9783
0.7032	0.1470	0.1498	0.0149	0.9824	0.0026	0.0169	0.0041	0.9790
0.6694	0.2136	0.1170	0.0147	0.9822	0.0031	0.0112	0.0055	0.9833
0.6308	0.2634	0.1057	0.0152	0.9808	0.0040	0.0150	0.0073	0.9776
0.8398	0.0924	0.0678	0.0168	0.9825	0.0007	0.0192	0.0009	0.9799
0.8198	0.0876	0.0926	0.0164	0.9826	0.0010	0.0185	0.0009	0.9806
0.8019	0.1346	0.0636	0.0168	0.9825	0.0006	0.0171	0.0013	0.9816
0.7867	0.1117	0.1017	0.0159	0.9834	0.0007	0.0189	0.0014	0.9797
0.8412	0.0886	0.0702	0.0168	0.9825	0.0007	0.0181	0.0009	0.9811
0.7926	0.1202	0.0872	0.0167	0.9824	0.0009	0.0195	0.0015	0.9790
0.7858	0.1244	0.0898	0.0165	0.9827	0.0008	0.0177	0.0013	0.9811
0.7945	0.1048	0.1007	0.0166	0.9824	0.0010	0.0183	0.0012	0.9805
0.8111	0.1241	0.0647	0.0179	0.9815	0.0007	0.0186	0.0014	0.9800
0.4181	0.2451	0.3368	0.0109	0.9800	0.0091	0.1211	0.0075	0.8715
0.3601	0.2282	0.4117	0.0094	0.9816	0.0090	0.0406	0.0084	0.9510
0.3301	0.1103	0.5596	0.0055	0.9770	0.0175	0.0086	0.0060	0.9855
0.3581	0.5342	0.1077	0.0164	0.9777	0.0059	0.0455	0.0083	0.9463
0.7527	0.1782	0.0691	0.0161	0.9832	0.0007	0.0291	0.0018	0.9691
0.8350	0.1024	0.0625	0.0168	0.9826	0.0006	0.0192	0.0010	0.9799
0.8152	0.0917	0.0931	0.0165	0.9826	0.0009	0.0199	0.0010	0.9791
0.4801	0.2246	0.2953	0.0122	0.9830	0.0048	0.1160	0.0075	0.8766
0.3342	0.2767	0.3891	0.0123	0.9831	0.0046	0.2273	0.0248	0.7479

0.2747	0.1573	0.5680	0.0029	0.9733	0.0238	0.2772	0.0516	0.6712
0.2747	0.1573	0.5680	0.0029	0.9733	0.0238	0.2772	0.0516	0.6712
0.3307	0.4312	0.2380	0.0086	0.9834	0.0080	0.0149	0.0091	0.9760
0.4137	0.2287	0.3576	0.0131	0.9827	0.0042	0.1585	0.0138	0.8277
0.3632	0.2290	0.4078	0.0085	0.9824	0.0091	0.0233	0.0084	0.9683
0.3900	0.3030	0.3069	0.0118	0.9817	0.0065	0.1889	0.0156	0.7956
0.3997	0.3998	0.2004	0.0175	0.9778	0.0047	0.1889	0.0181	0.7930
0.3878	0.2686	0.3436	0.0105	0.9818	0.0077	0.0963	0.0078	0.8959
0.3401	0.4297	0.2302	0.0125	0.9809	0.0066	0.2067	0.0196	0.7737
0.3304	0.1877	0.4819	0.0061	0.9761	0.0177	0.0142	0.0098	0.9761

Table 3: This table displays model estimates of the transition probabilities and average transition probabilities for the three-state time homogenous HMM and time-inhomogeneous HMM-RND respectively.

TABLE 4: ESTIMATED LOG-LINEAR EFFECTS OF COVARIATES IN SELECTED HMM-RND MODELS

VOLSKEWKURT					LT-VOLSKEWKURT				
	State	1	2	3		State	1	2	3
VOL	1	0.00	-3.18	0.00	LT-VOL	1	0.00	-4.87	0.00
	2	10.96 **	0.00	-29.39 **		2	-11.36	0.00	-4.41
	3	0.00	4.55	0.00		3	0.00	3.75	0.00
SKEW	1	0.00	0.26	0.00	LT-SKEW	1	0.00	0.90	0.00
	2	-0.04	0.00	1.72 **		2	3.44 **	0.00	3.48 **
	3	0.00	0.68	0.00		3	0.00	-0.61	0.00
KURT	1	0.00	0.03	0.00	LT-KURT	1	0.00	0.15	0.00
	2	0.15 **	0.00	0.30 **		2	0.66 **	0.00	0.60 **
	3	0.00	0.09	0.00		3	0.00	-0.02	0.00

ST-VOLSKEWKURT					Q01-Q025-Q975-Q99				
	State	1	2	3		State	1	2	3
ST-VOL	1	0.00	4.56	0.00	Q01	1	0.00	10.10	0.00
	2	-0.74	0.00	-19.71		2	-18.02	0.00	-12.54
	3	0.00	2.72	0.00		3	0.00	23.50	0.00
ST-SKEW	1	0.00	0.50	0.00	Q025	1	0.00	-9.89	0.00
	2	-2.04 **	0.00	0.40		2	41.21	0.00	89.72 **
	3	0.00	-0.41	0.00		3	0.00	-2.14	0.00
ST-KURT	1	0.00	0.02	0.00	Q975	1	0.00	22.17	0.00
	2	-0.11	0.00	0.04		2	-52.10 **	0.00	-39.00
	3	0.00	-0.05	0.00		3	0.00	-44.77	0.00
					Q99	1	0.00	-36.02	0.00
						2	12.88	0.00	-1.85
						3	0.00	33.67	0.00

Table 4: This table contains the fitted parameter estimates on the log-linear effects of covariate variables on the transition intensities for selected HMM-RND models, namely with VOLSKEWKURT, LT-VOLSKEWKURT, ST-VOLSKEWKURT, and Q01-Q025-Q975-Q99 covariates. Significance stars "**" denote significance at the 95% level.

TABLE 5: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH TRANSITION PROBABILITY COVARIATES, FULL FORECAST PERIOD 5 MAY 2007 TO 29 APRIL 2011

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
1%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	ALL Models	ALL Models	LT-MEANVOL	LT-MEANVOL	LT-MEANVOL
5%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	ALL Models	ALL Models	LT-MEANVOL	LT-MEANVOL	LT-MEANVOL
10%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	ALL Models	ALL Models	LT-MEANVOL	LT-MEANVOL	LT-MEANVOL

Table 5: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the entire forecast period, where HMM-RND models were specified with covariate-dependent transition probabilities. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 6: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH TRANSITION PROBABILITY COVARIATES, FIRST HALF OF FORECAST PERIOD 5 MAY 2007 TO 5 MAY 2009

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
1%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	ALL Models	ALL Models	LT-MEANVOL, Q025, Q75, Q95, Q01-Q025-Q975-Q99	VOL, VOLSKEW, LT-MEANVOL, Q025, Q75, Q95, Q025-Q975, Q01-Q025-Q05, Q01-Q025-Q975-Q99	LT-MEANVOL, Q025, Q75, Q85, Q95, Q01-Q025-Q975-Q99
5%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	ALL Models	ALL Models	LT-MEANVOL, Q75, Q95	LT-MEANVOL, Q025, Q75, Q95	LT-MEANVOL, Q025, Q75, Q85, Q95, Q01-Q025-Q975-Q99
10%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	ALL Models	ALL EXCEPT Naïve, LT-VOL, LT-VOLKURT, LT_SKEWKURT, LT-VOLSKEWKURT, Q01, Q025, Q75, Q85, Q025-Q975, Q01-Q025-Q05, Q95-Q975-Q99	LT-MEANVOL, Q75	LT-MEANVOL, Q025, Q75, Q95	LT-MEANVOL, Q025, Q75, Q85, Q95

Table 6: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the first half of the forecast period, where HMM-RND models were specified with covariate-dependent transition probabilities. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 7: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH TRANSITION PROBABILITY COVARIATES, SECOND HALF OF FORECAST PERIOD 5 MAY 2009 TO 29 APRIL 2009

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
1%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	All Models	All Models	LT-VOL, LT-MEANVOL	MEANVOL, LT-VOL, LT-MEANVOL, LT-VOLKURT, LT-SKEWKURT, Q01	LT-VOL, LT-MEANVOL
5%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	All Models	All Models	LT-VOL, LT-MEANVOL	MEANVOL, LT-VOL, LT-MEANVOL, LT-VOLKURT, Q01	LT-VOL, LT-MEANVOL
10%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	All Models	All Models	LT-VOL, LT-MEANVOL	MEANVOL, LT-VOL, LT-MEANVOL, LT-VOLKURT, Q01	LT-VOL, LT-MEANVOL

Table 7: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the second half of the forecast period, where HMM-RND models were specified with covariate-dependent transition probabilities. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 8: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH TRANSITION PROBABILITY COVARIATES, GLOBAL FINANCIAL CRISIS SUB-PERIOD 2 JANUARY 2008 TO 31 DECEMBER 2008

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
1%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	All Models	All Models	LT-MEANVOL	VOL, MEANVOL, VOLSKEW, LT-MEANVOL, Q025, Q05, Q10, Q75, Q025-Q975, Q01-Q025-Q05	VOL, LT_MEANVOL
5%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	All Models	All Models	LT-MEANVOL	VOL, MEANVOL, VOLSKEW, LT-MEANVOL, Q025, Q05, Q10, Q025-Q975, Q01-Q025-Q05	VOL, LT_MEANVOL
10%	ALL EXCEPT Naïve	ALL EXCEPT Naïve	All Models	All Models	LT-MEANVOL	VOL, MEANVOL, VOLSKEW, LT-MEANVOL, Q025, Q025-Q975, Q01-Q025-Q05	LT_MEANVOL

Table 8: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the 2008 GFC portion of the forecast period, where HMM-RND models were specified with covariate-dependent transition probabilities. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 9: PREDICTION INTERVAL COVERAGE AND STATE TRANSITION FORECASTS, HMM-RND MODELS WITH TRANSITION PROBABILITY COVARIATES

	nnvio		GFC		Post GFC	s1f1	s2f2	s3f3	Total Correct		s1f2	s1f3	s2f1	s2f3	s3f1	s3f2
RND-TH	151		97		54	38	255	133	426		14	137	128	201	34	60
Mean	145		89		56	124	523	153	800	***	61	4	5	56	0	74
Vol	123	*	74	*	49	156	517	153	826	***	33	0	18	49	0	74
Skew	139		81		58	130	542	151	823	***	55	4	6	36	0	76
Kurt	141		88		53	126	533	170	829	***	63	0	4	47	0	57
MeanVol	129		83		46	139	524	152	815	***	50	0	7	53	0	75
VolSkew	139		77		62	139	528	162	829	***	50	0	6	50	0	65
VolKurt	148		89		59	134	520	167	821	***	53	2	7	57	0	60
SkewKurt	144		91		53	130	523	164	817	***	53	6	8	53	0	63
VolSkewKurt	144		89		55	126	514	170	810	***	57	6	10	60	0	57
LT_Mean	147		90		57	123	525	153	801	***	64	2	6	53	0	74
LT_Vol	136		89		47	127	501	153	781	***	62	0	32	51	0	74
LT_Skew	143		83		60	126	539	156	821	***	59	4	5	40	0	71
LT_Kurt	137		80		57	148	534	145	827	***	36	5	10	40	0	82
LT_MeanVol	119	**	73	*	46	154	507	151	812	***	35	0	36	41	0	76
LT_VolSkew	137		81		56	152	502	166	820	***	34	3	11	71	0	61
LT_VolKurt	148		96		52	127	505	156	788	***	62	0	8	71	0	71
LT_SkewKurt	146		98		48	135	491	171	797	***	51	3	14	79	0	56
LT_VolSkewKurt	151		96		55	105	510	163	778	***	75	9	5	69	0	64
ST_Mean	150		94		56	116	518	154	788	***	72	1	7	59	0	73
ST_Vol	145		88		57	112	520	156	788	***	73	4	8	56	0	71
ST_Skew	145		89		56	117	528	152	797	***	66	6	5	51	0	75
ST_Kurt	149		95		54	112	528	146	786	***	75	2	2	54	0	81
ST_MeanVol	146		91		55	121	518	155	794	***	65	3	9	57	1	71
ST_VolSkew	143		88		55	113	531	146	790	***	71	5	7	46	0	81
ST_VolKurt	142		87		55	114	534	150	798	***	72	3	4	46	1	76
ST_SkewKurt	152		96		56	101	514	157	772	***	75	13	8	62	0	70
ST_VolSkewKurt	149		93		56	117	517	154	788	***	69	3	7	60	0	73
Q01	130		82		48	140	521	151	812	***	49	0	7	56	0	76
Q025	127		74	*	53	155	514	161	830	***	34	0	18	52	0	66
Q05	137		80		57	132	523	169	824	***	57	0	6	55	0	58
Q10	138		81		57	140	513	167	820	***	49	0	11	60	0	60
Q25	147		88		59	123	532	152	807	***	62	4	5	47	0	75
Q40	148		92		56	123	519	151	793	***	63	3	6	59	0	76
Q50	147		91		56	119	521	157	797	***	68	2	7	56	0	70
Q60	141		87		54	123	528	156	807	***	63	3	6	50	0	71
Q75	127		77		50	159	498	172	829	***	30	0	21	65	0	55
Q85	134		77		57	153	508	174	835	***	36	0	11	65	0	53
Q90	143		88		55	125	515	172	812	***	57	7	13	56	0	55
Q95	135		74	*	61	153	516	169	838	***	36	0	12	56	0	58
Q975	136		79		57	151	515	172	838	***	38	0	12	57	0	55
Q99	141		82		59	150	526	167	843	***	39	0	5	53	0	60
Q01-99	138		78		60	148	519	171	838	***	41	0	10	55	0	56
Q025-975	136		76		60	151	518	173	842	***	38	0	11	55	0	54
Q01-025-05	130		76		54	153	518	158	829	***	36	0	10	56	0	69
Q95-975-99	134		78		56	150	513	167	830	***	38	1	14	57	0	60
Q01-025-975-99	133		73	*	60	151	515	173	839	***	38	0	13	56	0	54

Table 9: This table contains the results of 90% prediction interval violations for all the considered models, as well as the summary of the state transition predictions results from HMM-TH and HMM-RND models, where HMM-RND models are specified with covariate-dependent transition probabilities. Column 1 ("nvio") reflects number of violations recorded over the entire out-of-sample forecast period, while columns 2 ("GFC) and 3 ("post-GFC") reflect the number of violations recorded in the first and second halves of the forecast period respectively. The header terms "s*n*f*m*" in the remaining columns denote the number of decoded states *n* where *n*=1,2,3 which were predicted by a model as state *m,m*=1,2,3. Significance code: '*' 10%, '**' 5%, '***' 1% significance levels.

TABLE 10: MCS ANALYSIS OF MAE BETWEEN THE PREDICTED AND 'ACTUAL' PROBABILITIES OF THE DECODED STATE, HMM-RND MODELS WITH TRANSITION PROBABILITY COVARIATES

	Fracdiff	A_Whittle	Robinson
1%	LT_KURT, Q975, Q99, Q01-Q99, Q95-Q975-Q99	Q975, Q99, Q01-Q99, Q95-Q975-Q99	Q975, Q99, Q01-Q99, Q95-Q975-Q99
5%	LT_KURT, Q975, Q99, Q01-Q99, Q95-Q975-Q99	Q975, Q99, Q01-Q99, Q95-Q975-Q99	Q975, Q99, Q01-Q99, Q95-Q975-Q99
10%	LT_KURT, Q975, Q99, Q01-Q99, Q95-Q975-Q99	Q975, Q99, Q01-Q99, Q95-Q975-Q99	Q975, Q99, Q01-Q99, Q95-Q975-Q99

Table 10: This table report the results of the MCS performed using on the mean absolute forecast errors over the 2008 over the full out-of-sample forecast period, where HMM-RND models were specified with covariate-dependent transition probabilities. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 11: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH STATE DISTRIBUTION MEAN COVARIATES, FULL FORECAST PERIOD 5 MAY 2007 TO 29 APRIL 2011

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
1%	ALL EXCEPT Naïve, ST-VOL, ST-MEANVOL, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q95, Q975, Q01-Q025-Q05	ALL EXCEPT ST-MEANVOL, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q95, Q01-Q025-Q05	All Models	All Models	Q01-Q025-Q05	VOLSKEW, SKEWKURT, VOLSKEWKURT, LT-VOLKURT, LT-SKEWKURT, Q10, Q40, Q50	Q01-Q025-Q05
5%	HMM-TH, MEAN, SKEW, KURT, VOLKURT, SKEWKURT, LT-MEAN, LT-VOL, LT-SKEW, LT-VOLKURT, ST-MEAN, Q01, Q025, Q10, Q25, Q40, Q50, Q60	HMM-TH, MEAN, SKEW, KURT, MEANVOL, VOLSKEW, VOLKURT, SKEWKURT, LT-MEAN, LT-VOL, LT-SKEW, LT-KURT, LT-VOLKURT, ST-MEAN, Q01, Q25, Q40, Q50, Q60	All Models	ALL EXCEPT Naïve, ST-VOL, Q01-Q025-Q05	Q01-Q025-Q05	VOLSKEW, SKEWKURT, VOLSKEWKURT, LT-VOLKURT, LT-SKEWKURT, Q10, Q50	Q01-Q025-Q05
10%	HMM-TH, MEAN, SKEW, KURT, VOLKURT, SKEWKURT, LT-MEAN, LT-VOL, LT-SKEW, LT-VOLKURT, ST-MEAN, Q01, Q025, Q10, Q25, Q40, Q50, Q60	HMM-TH, MEAN, SKEW, KURT, MEANVOL, VOLSKEW, VOLKURT, SKEWKURT, LT-MEAN, LT-VOL, LT-SKEW, LT-KURT, LT-VOLKURT, ST-MEAN, Q01, Q25, Q40, Q50, Q60	All Models	HMM-TH, MEAN, SKEW, MEANVOL, VOLKURT, SKEWKURT, VOLSKEWKURT, LT-MEAN, LT-SKEW, ST-MEAN, Q40, Q50, Q60	Q01-Q025-Q05	VOLSKEW, SKEWKURT, VOLSKEWKURT, LT-VOLKURT, LT-SKEWKURT, Q10	Q01-Q025-Q05

Table 11: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the entire forecast period, where HMM-RND models were specified with covariate-dependent state distribution means. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 12: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH STATE DISTRIBUTION MEAN COVARIATES, FIRST HALF OF FORECAST PERIOD 5 MAY 2007 TO 5 MAY 2009

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
1%	ALL EXCEPT Naïve, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q01-Q025-Q05	ALL EXCEPT Naïve, ST-VOLSKEWKURT, Q01-Q025-Q05	All Models	All Models	ST-VOLKURT, Q01-Q025-Q05	VOLSKEW, SKEWKURT, VOLSKEWKURT, LT-SKEWKURT, ST-VOLKURT, Q40, Q50, Q60, Q95, Q01-Q025-Q05	ST-VOLKURT, Q95
5%	RND-TH, MEAN, VOL, SKEW, KURT, MEANVOL, VOLSKEW, VOLKURT, SKEWKURT, LT-MEAN, LT-VOL, LT-SKEW, LT-VOLKURT, ST-MEAN, Q01, Q025, Q10, Q25, Q40, Q50, Q60	RND-TH, MEAN, VOL, SKEW, KURT, MEANVOL, VOLSKEW, VOLKURT, SKEWKURT, LT-MEAN, LT-VOL, LT-SKEW, LT-KURT, LT-VOLSKEW, LT-VOLKURT, LT-VOLSKEWKURT, ST-MEAN, ST-SKEWKURT, Q01, Q025, Q25, Q40, Q50, Q60	All Models	All Models	ST-VOLKURT	VOLSKEW, SKEWKURT, VOLSKEWKURT, LT-SKEWKURT, ST-VOLKURT, Q40, Q50, Q60, Q95, Q01-Q025-Q05	ST-VOLKURT
10%	RND-TH, MEAN, VOL, SKEW, KURT, MEANVOL, VOLSKEW, VOLKURT, SKEWKURT, LT-MEAN, LT-VOL, LT-SKEW, LT-VOLKURT, ST-MEAN, Q01, Q025, Q10, Q25, Q40, Q50, Q60	RND-TH, MEAN, SKEW, KURT, SKEWKURT, LT-MEAN, LT-SKEW	All Models	All Models	ST-VOLKURT	VOLSKEW, SKEWKURT, VOLSKEWKURT, LT-SKEWKURT, ST-VOLKURT, Q40, Q50, Q60, Q95, Q01-Q025-Q05	ST-VOLKURT

Table 12: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the first half of the forecast period, where HMM-RND models were specified with covariate-dependent state distribution means. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 13: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH STATE DISTRIBUTION MEAN COVARIATES, SECOND HALF OF FORECAST PERIOD 5 MAY 2009 TO 29 APRIL 2009

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
0%	ALL EXCEPT Naïve, ST-VOL, ST-MEANVOL, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q01-Q025-Q05	ALL EXCEPT Naïve, ST-VOLSKEW, ST-VOLSKEWKURT, Q01-Q025-Q05	All Models	All Models	Q01-025-05	VOLSKEWKURT, LT-KURT, LT-MEANVOL, LT-VOLKURT, Q10	Q01-025-05
5%	ALL EXCEPT Naïve, ST-VOL, ST-MEANVOL, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q01-Q025-Q05	ALL EXCEPT Naïve, ST-VOL, ST-MEANVOL, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q85, Q90, Q95, Q01-Q025-Q05, Q01-Q025-Q975-Q99	All Models	All Models	Q01-025-05	VOLSKEWKURT, LT-KURT, LT-MEANVOL, LT-VOLKURT, Q10	Q01-025-05
10%	ALL EXCEPT Naïve, ST-VOL, ST-MEANVOL, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q01-Q025-Q05	ALL EXCEPT Naïve, ST-VOL, ST-MEANVOL, ST-VOLSKEW, ST-VOLKURT, ST-VOLSKEWKURT, Q85, Q90, Q95, Q01-Q025-Q05, Q01-Q025-Q975-Q99	All Models	All Models	Q01-025-05	VOLSKEWKURT, LT-KURT, LT-MEANVOL, LT-VOLKURT, Q10	Q01-025-05

Table 13: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the second half of the forecast period, where HMM-RND models were specified with covariate-dependent state distribution means. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 14: MODEL CONFIDENCE SET PERFORMED ON OUT-OF-SAMPLE FORECAST METRICS, HMM-RND MODELS WITH STATE DISTRIBUTION MEAN COVARIATES, GLOBAL FINANCIAL CRISIS SUB-PERIOD 2 JANUARY 2008 TO 31 DECEMBER 2008

	MAE	MSE	MAPE	MSE-PROP	ABS-DIST	ABS-SCALED ERR	CI-WIDTH
1%	ALL EXCEPT Naïve, Q01-Q025-Q05	ALL EXCEPT Naïve, Q01-Q025-Q05	All Models	All Models	ST-VOLKURT, Q01-Q025-Q05	ST-VOLKURT	ST-VOL, ST-VOLKURT, Q01-Q025-Q05
5%	ALL EXCEPT Naïve, ST-VOLSKEW, ST-VOLSKEWKURT, Q01-Q025-Q05	ALL EXCEPT Naïve, Q01-Q025-Q05	All Models	All Models	Q01-025-05	ST-VOLKURT	ST-VOL, ST-VOLKURT, Q01-Q025-Q05
10%	ALL EXCEPT Naïve, ST-VOLSKEW, ST-VOLSKEWKURT, Q01-Q025-Q05	ALL EXCEPT Naïve, Q01-Q025-Q05	All Models	All Models	Q01-025-05	ST-VOLKURT	ST-VOL, ST-VOLKURT, Q01-Q025-Q05

Table 14: This table report the results of the MCS performed using on the out-of-sample forecast error metrics over the 2008 GFC segment of the forecast period, where HMM-RND models were specified with covariate-dependent state distribution means. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

TABLE 15: PREDICTION INTERVAL COVERAGE AND STATE TRANSITION FORECASTS, HMM-RND MODELS WITH STATE DISTRIBUTION MEAN COVARIATES, FULL OUT-OF-SAMPLE PERIOD

	nnvio		GFC		Post GFC		s1f1	s2f2	s3f3	Total correct		s1f2	s1f3	s2f1	s2f3	s3f1	s3f2
RND-TH	151		97		54		38	255	133	426		14	137	128	201	34	60
Mean	145		100		45		98	321	190	609	***	91	0	101	162	1	36
Vol	154		104		50		144	231	208	583	***	45	0	117	236	1	18
Skew	161		107		54		87	304	215	606	***	102	0	52	228	1	11
Kurt	147		100		47		130	281	216	627	***	57	2	84	219	0	11
MeanVol	161		107		54		19	406	209	634	***	169	1	16	162	0	18
VolSkew	142		90		52		121	259	214	594	***	64	4	64	261	0	13
VolKurt	148		95		53		116	306	216	638	***	71	2	25	253	0	11
SkewKurt	132		87		45		160	305	216	681	***	27	2	76	203	0	11
VolSkewKurt	143		98		45		92	236	211	539	***	89	8	98	250	0	16
LT_Mean	145		100		45		98	321	190	609	***	91	0	101	162	1	36
LT_Vol	147		97		50		156	328	211	695	***	31	2	91	165	2	14
LT_Skew	150		94		56		120	284	211	615	***	67	2	91	209	2	14
LT_Kurt	147		104		43		166	279	202	647	***	22	1	107	198	0	25
LT_MeanVol	166		112		54		92	333	219	644	***	97	0	52	199	1	7
LT_VolSkew	151		98		53		105	291	210	606	***	84	0	135	158	1	16
LT_VolKurt	137		99		38		155	252	202	609	***	32	2	115	217	1	24
LT_SkewKurt	137		89		48		126	314	213	653	***	63	0	77	193	0	14
LT_VolSkewKurt	136		98		38		162	265	216	643	***	27	0	115	204	0	11
ST_Mean	161		107		54		113	334	214	661	***	74	2	33	217	0	13
ST_Vol	177		108		69		74	229	207	510	***	110	5	33	322	2	18
ST_Skew	154		103		51		97	311	209	617	***	91	1	54	219	0	18
ST_Kurt	153		98		55		140	279	212	631	***	47	2	128	177	0	15
ST_MeanVol	210	***	121	*	89	***	70	258	217	545	***	109	10	3	323	0	10
ST_VolSkew	230	***	145	***	85	***	15	216	212	443		172	2	127	241	11	4
ST_VolKurt	174		99		75	*	27	168	196	391		126	36	94	322	0	31
ST_SkewKurt	144		94		50		166	333	203	702	***	22	1	71	180	0	24
ST_VolSkewKurt	199	***	120		79	**	51	326	184	561	***	100	38	0	258	0	43
Q01	140		93		47		154	354	181	689	***	34	1	126	104	0	46
Q025	152		99		53		110	285	202	597	***	78	1	124	175	2	23
Q05	160		109		51		82	363	196	641	***	105	2	61	160	1	30
Q10	140		94		46		146	405	173	724	***	39	4	64	115	0	54
Q25	149		99		50		89	312	199	600	***	97	3	117	155	0	28
Q40	144		91		53		121	277	209	607	***	67	1	158	149	1	17
Q50	147		94		53		113	330	210	653	***	76	0	76	178	1	16
Q60	148		96		52		148	206	214	568	***	40	1	165	213	2	11
Q75	164		107		57		80	349	208	637	***	107	2	85	150	0	19
Q85	154		101		53		72	308	216	596	***	100	17	81	195	1	10
Q90	155		102		53		91	203	209	503	***	89	9	130	251	4	14
Q95	158		103		55		125	251	218	594	***	63	1	126	207	0	9
Q975	168		114		54		12	255	212	479	**	176	1	110	219	3	12
Q99	167		109		58		70	348	217	635	***	118	1	14	222	0	10
Q01-99	156		102		54		105	364	206	675	***	82	2	17	203	0	21
Q025-975	169		111		58		21	309	202	532	***	165	3	86	189	0	25
Q01-025-05	181	*	92		89	***	128	451	29	608	***	55	6	69	64	79	119
Q95-975-99	160		109		51		123	381	204	708	***	66	0	75	128	1	22
Q01-025-975-99	154		100		54		75	353	201	629	***	110	4	49	182	0	26

Table 15: This table contains the results of 90% prediction interval violations for all the considered models, as well as the summary of the state transition predictions results from HMM-TH and HMM-RND models, where HMM-RND models are specified with covariate-dependent state distribution means. Column 1 ("nvio") reflects number of violations recorded over the entire out-of-sample forecast period, while columns 2 ("GFC) and 3 ("post-GFC") reflect the number of violations recorded in the first and second halves of the forecast period respectively. The header terms "s*n*f*m*" in the remaining columns denote the number of decoded states *n* where *n*=1,2,3 which were predicted by a model as state *m,m*=1,2,3. Significance code: '*' 10%, '**' 5%, '***' 1% significance levels.

TABLE 16: MCS ANALYSIS OF MAE BETWEEN THE PREDICTED AND 'ACTUAL' PROBABILITIES OF THE DECODED STATE, HMM-RND MODELS WITH DISTRIBUTION MEAN COVARIATES

	Fracdiff	A_Whittle	Robinson
1%	LT-KURT, ST-VOLSKEWKURT	LT-VOLKURT, ST-KURT, ST-VOLKURT, ST-VOLSKEWKURT, Q975	LT-KURT, LT-VOLSKEWKURT, ST-KURT, ST-VOLKURT, ST-VOLSKEWKURT, Q975,Q99
5%	LT-KURT	ST-VOLKURT, ST-VOLSKEWKURT, Q975	LT-KURT, LT-VOLSKEWKURT, ST-KURT, ST-VOLKURT, ST-VOLSKEWKURT, Q975,Q99
10%	LT-KURT	ST-VOLKURT, ST-VOLSKEWKURT, Q975	LT-KURT, LT-VOLSKEWKURT, ST-KURT, ST-VOLKURT, ST-VOLSKEWKURT, Q975

Table 16: This table report the results of the MCS performed using on the mean absolute forecast errors over the 2008 over the full out-of-sample forecast period, where HMM-RND models were specified with covariate-dependent state distribution means. The null hypothesis is that models contained in the resultant MCS are indistinguishable based on a considered error metric, at the 1%, 5% and 10% levels of significance.

REFERENCES

Adrian, T., & Rosenberg, J. (2008). Stock Returns and Volatility: Pricing the Short‐Run and Long‐Run Components of Market Risk. The Journal of Finance, 63(6), 2997-3030.

Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. The Journal of Finance, 53(2), 499-547.

Aït-Sahalia, Y., Wang, Y., Yared, F. (2001). Do option markets correctly price the probabilities of movement of the underlying asset?, Journal of Econometrics", vol.102, no.1, pp.67-110.

Banerjee, P. S., Doran, J. S., & Peterson, D. R. (2007). Implied volatility and future portfolio returns, Journal of Banking & Finance, 31(10), 3183-3199.

Bates, D. S. (1991). The crash of'87: Was it expected? The evidence from options markets. The Journal of Finance, 46(3), 1009-1044.

Baum, L. E., Petrie, T., Soules, G. & Weiss, N. (1970), A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, Annals of Mathematical Statistics 41, 164–171.

Birru, J., & Figlewski, S. (2012). Anatomy of a Meltdown: The Risk Neutral Density for the S&P 500 in the Fall of 2008. Journal of Financial Markets, 15(2), 151-180.

Black, F. (1975). Fact and fantasy in the use of options. Financial Analysts Journal, 36-72.

Bliss, R. R., & Panigirtzoglou, N. (2002). Testing the stability of implied probability density functions. Journal of Banking & Finance, 26(2), 381-422.

Bratu, M. (2013). The Assessment And Improvement Of The Accuracy For The Forecast Intervals. In Working Papers of Macroeconomic Modelling Seminar (No. 132602). Institute for Economic Forecasting.

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. Journal of business, 621-651.

Bry, G., & Boschan, C. (1971). Front matter to" Cyclical Analysis of Time Series: Selected Procedures and Computer Programs". In Cyclical Analysis of Time Series: Selected Procedures and Computer Programs (pp. 13-2). UMI.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. Review of Financial studies, 6(3), 659-681.

Cappé, O., Moulines, E., & Rydén, T. (2005). Inference in hidden Markov models. Springer.

Chauvet, M., & Potter, S. (2001). Recent changes in the US business cycle. The Manchester School, 69(5), 481-508.

Chiras, D. P., & Manaster, S. (1978). The information content of option prices and a test of market efficiency. Journal of Financial Economics, 6(2), 213-234.

Clews, R., Panigirtzoglou, N., Proudman, J., (2000). Recent developments in extracting information from option markets, Bank of England Quarterly Bulletin, 40, 50–60.

Copeland, M., Copeland, T., (1999). Market timing: Style and size rotation using the VIX'. Financial Analysts Journal", 55, 73–81.

Cox, J. C., & Ross, S. A. (1976). The valuation of options for alternative stochastic processes. Journal of financial economics, 3(1), 145-166.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977), Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society. Series B. Methodological 39(1), 1-38. With discussion.

Diavatopoulos, D., Doran, J. S., Peterson, D. R. (2008). The information content in implied idiosyncratic volatility and the cross-section of stock returns: Evidence from the option markets, Journal of Futures Markets", 28(11), 1013-1039.

Doukhan, P., Khezour, A., & Lang, G. (2003). Nonparametric Estimation for Long-Range Dependent Sequences. Theory and Applications of Long-Range Dependence, 303.

Ephraim, Y., & Merhav, N. (2002). Hidden markov processes. Information Theory, IEEE Transactions on, 48(6), 1518-1569.

Engle, R. F., & Rangel, J. G. (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. Review of Financial Studies, 21(3), 1187-1222.

Fung, J. K. (2007). The information content of option implied volatility surrounding the 1997 Hong Kong stock market crash. Journal of Futures Markets, 27(6), 555-574.

Giot, P. (2005). Relationships between implied volatility indexes and stock index returns, The Journal of Portfolio Management", 31(3), pp.92-100.

Gray, D., Merton, R., & Bodie, Z. (2008, September). A Contingent Claims Analysis of the Subprime Credit Crisis of 2007-2008. In CREDIT 2008 Conference on Liquidity and Credit Risk, Venice (pp. 22-23).

Guidolin, M., & Timmermann, A. (2005). Economic Implications of Bull and Bear Regimes in UK Stock and Bond Returns*. The Economic Journal, 115(500), 111-143.

Guidolin, M., & Timmermann, A. (2006). An econometric model of nonlinear dynamics in the joint distribution of stock and bond returns. Journal of Applied Econometrics, 21(1), 1-22.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica: Journal of the Econometric Society, 357-384.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. Econometrica, 79(2), 453-497.

Jackwerth, J.C. (1999). Option Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review, Journal of Derivatives, 7(2), 66-82.

Jackwerth, J.C. 2004. Option-implied risk-neutral distributions and risk aversion, Charlotteville: Research Foundation of AIMR.

Karian, Z., & Dudewicz, E. (2000). Fitting statistical distributions to data: The generalised lambda distribution and the generalised bootstrap methods. Boca Raton, USA.

Koopman, S. J., Jungbacker, B., & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. Journal of Empirical Finance, 12(3), 445-475.

Liporace, L. A. (1982), Maximum likelihood estimation for multivariate observations of Markov sources, Institute of Electrical and Electronics Engineers. Transactions on Information Theory 28(5), 729–734.

Lunde, A., & Timmermann, A. (2004). Duration dependence in stock prices. Journal of Business & Economic Statistics, 22(3).

Maheu, J. M., & McCurdy, T. H. (2000). Identifying bull and bear markets in stock returns. Journal of Business & Economic Statistics, 18(1), 100-112.

Maheu, J. M., McCurdy, T. H., & Song, Y. (2012). Components of bull and bear markets: bull corrections and bear rallies. Journal of Business & Economic Statistics, 30(3), 391-403.

Malz, A. M. (2001). Financial crises, implied volatility and stress testing. Riskmetrics group, New York.

Manaster, S., & Rendleman, R. J. (1982). Option prices as predictors of equilibrium stock prices. The Journal of Finance, 37(4), 1043-1057.

Noh, J., Engle, R. F., & Kane, A. (1994). Forecasting volatility and option prices of the S&P 500 index. The Journal of Derivatives, 2(1), 17-30.

Pagan, A. R., & Sossounov, K. A. (2003). A simple framework for analysing bull and bear markets. Journal of Applied Econometrics, 18(1), 23-46.

Panigirtzoglou, N., & Skiadopoulos, G. (2004). A new approach to modeling the dynamics of implied distributions: Theory and evidence from the S&P 500 options. Journal of Banking & Finance, 28(7), 1499-1520.

Psaradakis, Z., & Spagnolo, N. (2003). On the determination of the number of regimes in Markov-switching autoregressive models. Journal of Time Series Analysis, 24(2), 237-252.

Rabiner, L. (1989), A tutorial on hidden markov models and selected applications in speech recognition, Institute of Electrical and Electronics Engineers. Transactions on Information Theory 77(2), 257–284.

Robinson, P. M. (Ed.). (2003). Time series with long memory. Oxford University Press.

Shilling, J.D. (1992). Market timing: Better than a buy-and-hold strategy, Financial Analysts Journal, 48, 46-50.

Shimko, D. (1993). Bounds of probability. Risk, 6(4), 33-37.

Whaley, R. E. (2000). The investor fear gauge. The Journal of Portfolio Management, 26(3), 12-17.

White, S. I., Becker, R., & Clements, A. E. (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts?

Wu, C.-F. J. (1983), On the convergence properties of the EM algorithm, The Annals of Statistics 11(1), 95–103.

Yaniv, I., & Foster, D. P. (1997). Precision and accuracy of judgmental estimation. Journal of behavioral decision making, 10(1), 21-32.

Zucchini, W., & MacDonald, I. L. (2009). Hidden Markov models for time series: an introduction using R. CRC Press.

APPENDIX A

A1 The EM Algorithm

The maximum likelihood estimation of HMMs is commonly performed using the expectation maximization (EM) algorithm (Baum et al. 1970, Dempster et al. 1977).

The HMM likelihood function can be computed recursively by

[A1.1] $L_T = \boldsymbol{\alpha}_T \mathbf{1}'$, where

[A1.2] $\boldsymbol{\alpha}_1 = \boldsymbol{\pi}\boldsymbol{P}(x_1)$ and $\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1}\boldsymbol{\Gamma}\boldsymbol{P}(x_t),\ t = 2,3,\dots,T.$

The EM algorithm, also known as the Baum-Welch algorithm is a popular approach. A desirable property of the EM approach is that is can be applied for estimating HMMs with homogenous, but not necessarily stationary Markov chains. There has been extensive literature on the EM algorithm and its use in HMMs; comprehensive coverage of this topic can be found in pieces such as Baum et al. (1970), Dempster et al. (1977), Rabiner (1989), Liporace (1982), Wu (1983). We refer to Zucchini and MacDonald (2009) as the detailed reference on which this section draws.

The EM algorithm applies an iterative approach while accommodating missing data. The method can be generally described 1) the **E**xpectation stage: Calculate the conditional expectations of the functions of the missing data based on current observations and current estimate of the model parameters, $\boldsymbol{\vartheta}$, and, 2) the **M**aximization stage: Maximize the log-likelihood function of the parameters of interest, $\boldsymbol{\vartheta}$, where the functions of missing data have been replaced by their conditional expectations.

Repeat the two steps until a pre-specified convergence criterion is achieved. In the context of estimating HMMs, the EM algorithm can be fleshed out as follows. Consider a stochastic sequence of states, $s_1, s_2, \dots s_m$ which follow a Markov chain. Then we can define

[A1.3] $u_j(t) = 1$ if and only if $s_t = j$, $t = 1,2,\dots,T$

[A1.4] $v_{jk}(t) = 1$ if and only if $s_{t-1} = j$ and $s_t = k$, $t = 1,2,\dots,T.$

In the case of a HMM, based on the observations $x_1, x_2, \dots, x_T$ with missing data $s_1, s_2, \dots, s_T$, we can compute a Q-function representing the log-likelihood of the parameters $\boldsymbol{\vartheta}$

[A1.5] $$log\left[P\left(\boldsymbol{x}^{(\boldsymbol{T})}, \boldsymbol{s}^{(\boldsymbol{T})}\right)\right] = log\left(\pi_{s_1} \prod_{t=2}^{T} \gamma_{s_{t-1},s_t} \prod_{t=1}^{T} p_{s_t}(x_t)\right)$$

$$= log\pi_{s_1} + \sum_{t=2}^{T} log\gamma_{s_{t-1},s_t} + \sum_{t=1}^{T} logp_{s_t}(x_t)$$

$$= \sum_{j=1}^{m} u_j(1) log\pi_j + \sum_{j=1}^{m} \sum_{k=1}^{m} (\sum_{t=2}^{T} v_{jk}(t)) log\,\gamma_{jk}$$

$$+ \sum_{j=1}^{m} \sum_{t=1}^{T} u_{jk}(t)\, logp_j(x_t).$$

where $\boldsymbol{x}^{(\boldsymbol{T})}$ and $\boldsymbol{s}^{(\boldsymbol{T})}$ represent all past values, $\{x_1, .., x_T\}$ and $\{s_1, .., s_T\}$ respectively.

The EM algorithm specifically for estimating HMM therefore proceeds as follows

1. **E:** The unobserved variables $v_{jk}(t)$ and $u_j(t)$ are replaced by their conditional expectations based on the observed data $\boldsymbol{x}^{(\boldsymbol{T})}$ and current parameter estimates, therefore $\hat{u}_j(t) = P\left(S_t = j \middle| \boldsymbol{x}^{(\boldsymbol{T})}\right)$ and $\hat{v}_{jk}(t) = P\left(S_{t-1} = j, S_t = k \middle| \boldsymbol{x}^{(\boldsymbol{T})}\right)$. Stationarity of the underlying Markov chain is not necessarily assumed.
2. **M:** Following step 1, then maximise [5.22] with respect to the parameters of the state dependent distribution, the initial distribution, $\boldsymbol{\pi}$, and the transition probability matrix, $\boldsymbol{\Gamma}$.

Some complexity may arise with the estimation of the parameters of the state dependent distribution in the M step depending on the nature of the assumed distribution; closed form solutions are available for certain distributions (such as normal and Poisson), however numerical methods can be applied for other distributional types such as the negative binomial. We follow recent work (c.f. Maheu, McCurdy and Song (2012) which has demonstrated that Gaussian state-dependent distributions can adequately capture the observable market price dynamics.

A2 The Model Confidence Set

Formally, let M_0 denote a group of models (where each model is indexed $i = 1,2, \dots, m_0$) from which the best set of models M_* is to be determined according to some pre-specified criterion. In our context, such criterion would be some error function based on the point and interval forecast error metrics, $g(.)$.

Between any pair of models i and j, the variable $d_{ij,t} = g(e_{i,t}) - g(e_{j,t}) \; \forall \; i, j \in M_0$ is used to assess their relative performance. Then the set of optimum models is determined where

[A2.1] $$M_* \equiv \{i \in M_0 : E(d_{ij,t}) \leq 0 \; \forall \; j \in M_0\}.$$

The MCS procedure begins with the entire set of considered models and sequentially performs significance tests, ejecting the models which are deemed to be significantly inferior to the other models in M_0. In each iteration of the testing, the following hypotheses are tested

[A2.2] $$H_{0,M} : E(d_{ij,t}) = 0 \; \forall \; i, j \in M \qquad \text{vs.} \qquad H_{1,M} : E(d_{ij,t}) \neq 0 \; for \; some \;\; i, j \in M$$

where $M \subset M_0$ is the set of remaining models for M_*, and testing is performed at some pre-specified level of significance, α. When the null hypothesis $H_{0,M}$ is rejected on an iteration of testing, a model is eliminated from the set of considered models. The remaining set of models is then subject to the hypothesis testing in [A2.2] again. This exercise is repeated until $H_{0,M}$ is no longer rejected; the resultant set of models $\widehat{M}_{*,1-\alpha}$ is then known as the model confidence set.

The implementation of the MCS method is in two parts, involving a test of *equivalence*, δ_M and a rule of elimination, e_M. The former test permits the testing of $H_{0,M}$, while the latter test determines the model to be dropped from M. The equivalence test applies the range statistic which is given by

[A2.3] $$T_M \equiv \max_{i,j \in M} |t_{ij}|, \qquad \text{where} \quad t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{var(\bar{d}_{ij})}}, \;\; \bar{d}_{ij} = \frac{1}{T} \sum_{t=1}^{T} d_{ij,t}.$$

The elimination rule e_M accordingly selects to drop the model i for which $t_{e_M,j} = T_M$ for some $j \in M$. Under the null, the asymptotic distribution of the test statistic T_M is nonstandard, hence the distribution of the test statistic is bootstrapped as is consistent with Hansen, Lunde, and Nason (2011). A comment in order regarding our application of the MCS is that it involves the comparison of nested models, which Hansen, Lunde, and Nason (2011) caution can lead to variability in the size of the tests therein. As such, we carry out the MCS methodology for several α% values namely the 1%, 5%, 10% and 20% significance levels.