Are the Dynamic Linkages Between the Macroeconomy and Asset Prices Time-Varying?

Authors	Massimo Guidolin, and Sadayuki Ono
Working Paper Number	2005-056A
Citable Link	https://doi.org/10.20955/wp.2005.056
Suggested Citation	Guidolin, M., Ono, S.; Are the Dynamic Linkages Between the Macroeconomy and Asset Prices Time-Varying?, Federal Reserve Bank of St. Louis Working Paper 2005-056. URL https://doi.org/10.20955/wp.2005.056

Published In	Journal of Economics and Business
Publisher Link	https://doi.org/10.1016/j.jeconbus.2006.06.009

Are the Dynamic Linkages Between the Macroeconomy and Asset Prices Time-Varying?*

Massimo Guidolin†
Federal Reserve Bank of St. Louis - Research Division

Sadayuki Ono
University of York, UK

Abstract

We estimate a number of multivariate regime switching VAR models on a long monthly US data set for eight variables that include excess stock and bond returns, the real T-bill yield, predictors used in the finance literature (default spread and the dividend yield), and three macroeconomic variables (inflation, real industrial production growth, and a measure of real money growth). Heteroskedasticity may be accounted for by making the covariance matrix a function of the regime. We find evidence of four regimes and of time-varying covariances. The best in-sample fit is provided by a four state model in which the VAR(1) component fails to be regime-dependent. We interpret this as evidence that the dynamic linkages between financial markets and the macroeconomy have been stable over time. We show that the four-state model can be helpful in forecasting applications and to provide one-step ahead predicted Sharpe ratios.

JEL codes: E44, G12, C32, C52.
Keywords: Predictability, Multivariate Regime Switching, Predictive Density Tests, Sharpe ratios.

1. **Introduction**

The possibility that macroeconomic aggregates may predict the evolution of asset prices has been attracting the attention of a wide range of researchers in economics and finance at least since the late 1970s. Against the background of the efficient market hypothesis (EMH) developed in the 1960s and 70s (for which asset prices should follow a random walk or anyway be unpredictable given current information), the existence of statistically detectable predictability patterns has been considered interesting not only for its intrinsic usefulness in asset pricing and portfolio management, but also because a reconciliation between the EMH and the predictive power of macroeconomic variables was perceived as a high-priority research question. Therefore a remarkable bulk of empirical evidence on such predictability relationships linking asset returns and macroeconomic factors has been cumulating, although it is now clear that the EMH may be consistent with predictability.[1]

*Elizabeth La Jeunesse provided excellent research assistance.

†Correspondence to: Research Division, P.O. Box 442, St. Louis, MO 63166, United States. E-mail: Massimo.Guidolin@stls.frb.org; phone: 314-444-8550.

[1] In synthesis, the random walk actually obtains only under special assumptions or after appropriately scaling the asset prices. More generally, the EMH simply implies the existence of a relationship between asset returns and all variables that contain information on the fundamental pricing operator (the stochastic discount factor). Various papers study whether arbitrage pricing theory (APT) can employ macroeconomic variables as risk factors. These studies focus on a *contemporaneous* relation between stock returns and macroeconomic variables. Examples include Chen, Roll, and Ross (1986) and Burmeister and McElroy (1988) among others. Campbell, Lo, and Mackinlay (Chapters 2 and 6, 1997) survey these literatures.

At the same time, a recent literature has investigated whether asset returns may forecast future realized macroeconomic variables, particularly output and inflation (e.g. Stock and Watson (2003)). Ultimately, this is another rationality statement: since financial markets should efficiently aggregate in equilibrium asset prices expectations that relate to future economic conditions (output and relative prices), if the forecasts of market participants are rational one would expect that asset returns could on average predict future macroeconomic conditions, see e.g Fischer and Merton (1984). Of course, it is also true that financial markets routinely provide data in such quantities that it seems also very convenient for professional forecasters and policy makers to investigate whether any useful information may be extracted from financial prices.

Our paper deals with both aspects of the linkages between financial returns and macroeconomic variables and asks whether it is sensible – as it has been routinely done so far – to assume that such dynamic predictability relationships (if any) have been stable over time in the US. In fact, the US economy (as well as the world economy in general) has been changing at such a fast pace that attaching a lot of weight to a prior that such dynamic linkages would have remained stable and unchanging probably requires some careful scrutiny. In particular, the acceleration of the speed of change over the last 30 years, after the two oil shocks, a few experiments concerning the conduct of the monetary policy, and the removal of most of the remaining barriers to the free international flow of goods, services, and capital, may instead justify putting a lot of trust in the opposite prior that somehow the recent experience might be somehow different. Additionally, also stock (with the tech bubble of the 1990s) and bond (with the protracted worldwide decline in long-term interest rates since the early 1980s) markets have been subject to dramatic changes that may lend support to the hypothesis of changing dynamic linkages.

We investigate the hypothesis of time-varying dynamic linkages across financial markets and the macroeconomy in a highly flexible multivariate regime switching framework in which the existence of breaks and of structural change is captured as the system alternating among a number of recurrent regimes with different statistical properties. We stress that the model is flexible but also required by the question we investigate. It is flexible as it does not impose the presence of regimes but uses a number of econometric tools to test whether multiple regimes are needed in order to fit and/or forecast the data. Moreover, flexibility exists in the way regimes enter the model: in particular – when predictability patterns take the form of a vector autoregressive component by which past values of some variables may influence the conditional predictive mean of other variables – it is possible that regime may affect parameters (components) of the model that do not directly affect predictability. On the other hand, using models in such class appears to be a minimal requirement: it is well known that in the presence of non-stationarities, standard linear regression models deliver biased and hence irrelevant estimates.[2] Only by endogenizing the presence of regimes, we can learn in statistically meaningful ways about the issue of predictability.

Our paper gives at least three contributions. First, it estimates a relatively sophisticated range of multivariate $k-$regime VAR models in which heteroskedasticity may be accounted for by making the covariance matrix a function of the regime. The model is applied to an eight-variable vector that includes both stock and bond returns in excess of a T-bill rate, the T-bill yield, typical predictors used in the finance literature

between stock returns and macroeconomic variables. Examples include Chen, Roll, and Ross (1986) and Burmeister and McElroy (1988) among others. Campbell, Lo, and Mackinlay (Chapters 2 and 6, 1997) survey these literatures.

[2]To consider a simple case pertinent to this paper, think of what happens to the estimate of the slope coefficient in a regression model in which the intercept stochastically switches between two alternative values: unless the switching is taken into account, the estimate of the slope coefficient will be biased and inconsistent as the fitted slope will be inevitably higher than the true but unknown one.

(such as the default spread between low- and high-grade bond yields and the dividend yield), and three genuine macroeconomic variables: inflation, real industrial production growth, and a measure of real money growth. Given our objectives, we use the longest available monthly data base for US data, 1926:12 - 2004:12. We find evidence of four regimes and of time-varying covariances, i.e. of heteroskedasticity. The four regimes carry a relatively sensible interpretation as a moderately persistent bull-rebound state, a highly persistent stable state, an expansion, high growth state, and a recession-bear state. The last two regimes have low persistence and hence durations limited to 4-5 months.

Second, we provide evidence that the best in-sample fit to the joint density of the data is provided by a four state model in which the VAR(1) component fails to be regime-dependent. We interpret this as evidence that the dynamic linkages between financial markets and the macroeconomy have been stable over time, which counters a prior of evolving predictability patterns. To our knowledge such evidence of a stable dynamic relationship between financial markets and the US macroeconomy is new.

Third, we document two of the possible uses one could make of our estimation results. We show that the four-state model can be helpful in forecasting applications, in the sense that for many relevant variables (especially the financial ones, stock and bond returns) its recursive, out-of-sample predictive performance turns out to be superior to a simpler (and nested) VAR(1). Additionally, we provide evidence that the one-step ahead predicted Sharpe ratios for both stocks and bonds are much more sensible when evaluated under the four-state model than under a VAR(1). We argue that difference may be crucial in financial applications, such as portfolio management.

The applications of regime switching models in macroeconomics and finance are constantly expanding. Following Hamilton (1989), several papers have proposed to improve the empirical fit of standard, single-equation models for short-term interest rates (e.g. Gray (1996) and Ang and Bekaert (2002)) and stock returns (e.g. Turner, Startz, and Nelson (1989) and Ang and Bekaert (2001)) by allowing for mixtures of given conditional distributions. For instance, Turner, Startz, and Nelson (1989) develop a univariate model with regime shifts in means and variances, showing that mean excess equity returns tend to be low in the high-risk (volatility) period, and viceversa. Allowing for switching in the parameters of an autoregressive conditional heteroskedasticity (ARCH) process, Hamilton and Susmel (1994) report that in-sample performance and out-of-sample forecasts of the regime-switching ARCH are superior to a benchmark single-state GARCH(1,1) specification and that the high-volatility state is likely to occur in recession periods. Guidolin and Timmermann (2005b) extend this class of models to multivariate systems including excess returns on a few equity portfolios as well as bond returns. However, to our knowledge ours is the first paper to undertake a thorough investigation of predictability patterns involving stock and bond returns, along with a rich set of macroeconomic variables.

A literature exists that stresses that the forecasting power of financial variables for key macroeconomic variables is strongly time-varying. For instance, Stock and Watson (2003) report and discuss a bulk of evidence that shows that the US term structure fails to steadily predict output growth. Davis and Fagan (1997) document similar instability in the out-of-sample forecasting performance of yield spreads for nine European countries. Emery (1996) makes a similar point with reference to the instability of predictive relations involving the spread between commercial paper and T-bill yields. However Estrella et al. (2003) have concluded that when there is strong international evidence of forecasting power of the yield spreads for real activity, then the predictive relations also tend to be stable over time. Only a few papers − e.g. Jaditz et al. (1998) − have explored the possibility that carefully specified nonlinear prediction models may reproduce the possible time-variation that many papers have uncovered in the forecasting relations connecting financial variables to output and inflation. Our paper takes a few steps in this direction.

The paper is structured as follows. Section 2 gives a quick literature review that helps focussing on the goals of our exercise. Section 3 gives an introduction to estimation, inference, and forecasting in multivariate regime switching models. Section 4 gives information on the data employed in the paper. Section 5 estimates a range of switching models and proceeds to select the one providing the best fit according to a number of statistical criteria. Parameter estimates and interpretation are provided. The basic finding of no time-variation in the dynamic connections between financial markets and the macroeconomy is presented. Section 6 shows that a four-state model effectively produces (at least in some dimensions) useful out-of-sample forecasts. This validates the possibility that such a model may provide an approximation to the data generating process. Section 7 comments on possible financial applications of our findings. Section 8 concludes. Two appendices detail the technical aspects of the econometric methodology.

2. **Literature Review**

An impressive amount of literature has cumulated that investigates whether US stock and government bond returns are predictable using past values of macroeconomic variables.[3] In fact, using linear regression models, numerous studies have found that a few macroeconomic variables can be found that systematically predict US stock returns. Fama and French (1988) document that the dividend yield forecasts future returns on common stocks. Fama and Schwert (1977) report that real stock returns are negatively related with expected and unexpected components of inflation, which implies that real stocks are not a good hedge for inflation. They also show that industrial production and real GNP growth have forecasting power for financial returns. Cutler, Poterba and Summers (1989) examine the forecasting power of unexpected changes in a number of macroeconomic variables. They found that a positive shock to the rate of growth of industrial production significantly raises returns on a value-weighted NYSE portfolio. Balvers, Cosimano, and Mcdonald (1990) show that industrial production and real GNP predict stock returns with significantly negative coefficients. Several papers have shown that money supply is a key variable that determines fluctuations in stock prices. For instance, Homa and Jaffee (1971) found that the money supply growth rate contains predictive power for quarterly stock returns in the period 1954-1961. Kaul (1987) shows that a negative relationship between real stock returns and inflation in the post-war data may be caused by a counter-cyclical monetary policy. Hardouvelis (1987) examines stock market reactions to announcements on 15 different macroeconomic variables. He finds that monetary news have a significant effect on stock returns in the October 1979 - October 1982 period, when the Federal Reserve followed non-borrowed reserve targets. This type of finding points to the possibility of regime switching in predictability. Campbell (1987) presents evidence that a variety of term structure variables such as two-month and six-month spreads as well as the 1-month T-bill rate, all forecast excess stock returns. Fama and French (1989) confirm this result using data at alternative frequencies and a longer sample period (1927-1987). Similarly, Fama and French (1989) investigate whether default risk is a significant predictor of stock returns using the yield spread between low- and high-grade corporate bonds.

[3]We need to mention that a number of papers have also found that accounting variables that highlight some firm characteristics affect the cross-section of average stock returns. Among many others, we can cite firm size measured by total market capitalization (Banz (1981)), the earnings-price ratio (Basu (1983)), and the ratio of the book value to market value (Rosenberg, Reid, and Lanstein (1985)). Fama and French (1992) conclude that firm size and the book-to-market ratio have the strongest explanatory power for the cross-section of stock returns.

They find that higher spreads predict subsequent increases in stock returns.[4]

The early literature has been generalized in three directions. First, a few papers have tried to extend this evidence to bond markets. Campbell (1987) and Fama and French (1993) show that term and default spreads forecast excess returns on Treasury bills as well as long-term bonds. Second, a literature has extended these results to multivariate, full-information models in which not only macroeconomic factors are allowed to predict future asset returns, but also the opposite may occur. Here results are rather mixed. Using a vector-autoregression moving average (VARMA) approach, James, Koreisha, and Partch (1985) investigate simultaneous relations among stock returns, real activity, money supply, and inflation. Their findings support the notion that stock returns are important predictors of changes in expected inflation and nominal interest rates.[5] Lee (1992) uses a vector-autoregression (VAR) model, finding that an increase in real stock returns forecasts subsequent increases in real activity as measured by the growth rate of industrial production. However, in contrast to the above findings, Canova and De Nicolo (2000) show that US stock returns do not contain significant forecasting power of real activity and inflation, even in open-economy set ups. Using a structural VAR framework characterized by long-run monetary neutrality, Rapach (2001) studies the effects of money supply, aggregate spending and aggregate supply shocks on real stock returns, finding mixed results. Ang and Piazzesi (2003) extend this approach to the bond market and investigate the joint process of bond yields and macroeconomic variables in a VAR in which no-arbitrage restrictions are imposed.

Third, a number of papers have tried to *informally* generalize the evidence on predictability to models in which regimes play a role. For example, Pesaran and Timmermann (1995) introduce a time-varying *choice* of forecasting variables, in which the effective selection is based on a number of alternative model selection criteria (the adjusted R^2, the Akaike, Bayes-Schwartz, and Hannan-Quinn information criteria). They show that the optimal selection of prediction variables significantly changes over time and that only the 1-month T-bill rate is included in the selection over the entire sample. Flannery and Protopadakis (2002) show that macroeconomic announcements concerning inflation and money supply consistently affect the level of stock returns and that market responses on the announcement are time-varying. Employing VAR methods with endogenous break points, Du (2005) offers evidence that the (contemporaneous) correlation between real stock returns and inflation varies over time. He shows that the time-varying correlation is mainly due to changes in monetary policy regimes.

3. **Econometric Methodology**

Suppose that the random vector collecting monthly returns on n different assets and m macroeconomic variables possibly predicting (and predicted by) asset returns follows a $k-$regime Markov switching (MS) $VAR(p)$ process with heteroskedastic component, compactly $MSIAH(k,p)$ (see Krolzig (1997)):

$$\mathbf{y}_t = \boldsymbol{\mu}_{S_t} + \sum_{j=1}^{p} \mathbf{A}_{j,S_t}\mathbf{y}_{t-j} + \boldsymbol{\Sigma}_{S_t}\boldsymbol{\epsilon}_t \qquad (1)$$

[4]There is also an abundant international evidence on linear predictability. Ferson and Harvey (1993) examine the dynamic linkages between 18 international stock market returns and macroeconomic factors. Using data for 12 industrialized countries, Rapach, Wohar, and Rangvid (2005) show that interest rates have the most significant forecasting ability in almost every countries, both in-sample and out-of-sample tests.

[5]Early papers in this literature (e.g. Sims (1980), Estrella and Hardouvelis (1991), and Bernanke and Blinder (1992)) had shown that when short-term interest rates or interest rate spreads were included in VARs for output and inflation, they tended to eliminate the marginal predictive content of the money growth rate.

with $\epsilon_t \tilde{} NID(\mathbf{0}, \mathbf{I}_{n+m})$.[6] S_t is a latent state variable driving all the matrices of parameters appearing in (1). $\boldsymbol{\mu}_{S_t}$ collects the n regime-dependent intercepts, while the $(n+m) \times (n+m)$ matrix $\boldsymbol{\Sigma}_{S_t}$ represents the factor applicable to state S_t in a state-dependent Choleski factorization of the variance covariance matrix of the variables of interest, $\boldsymbol{\Omega}_{S_t}$. Obviously, a non-diagonal $\boldsymbol{\Sigma}_{S_t}$ makes the $n+m$ asset returns and macroeconomic predictors simultaneously cross-correlated, thus capturing *simultaneous* comovements between asset returns and macro factors. Clearly, *dynamic* (lagged) linkages across both different asset markets and between financial markets and macroeconomic influences are captured by the VAR(p) component. In fact, conditionally on the unobservable state S_t, (1) defines a standard Gaussian reduced form VAR(p) model. On the other hand, when $k > 1$, alternative hidden states are possible that will influence both the conditional mean and the volatility/correlation structures characterizing the multivariate process in (1), $S_t = 1, 2, ..., k\ \forall t$. These unobservable states are generated by a discrete-state, homogeneous, irreducible and ergodic first-order Markov chain:[7]

$$\Pr\left(s_t = j | \{s_j\}_{j=1}^{t-1}, \{\mathbf{y}_j\}_{j=1}^{t-1}\right) = \Pr\left(s_t = j | s_{t-1} = i\right) = p_{ij}, \tag{2}$$

where p_{ij} is the generic $[i, j]$ element of the $k \times k$ transition matrix $\mathbf{P}$. Ergodicity implies the existence of a stationary vector of probabilities $\bar{\boldsymbol{\xi}}$ satisfying $\bar{\boldsymbol{\xi}} = \mathbf{P}'\bar{\boldsymbol{\xi}}$. Irreducibility implies that $\bar{\boldsymbol{\xi}} > \mathbf{0}$, meaning that all unobservable states are possible. In practice, $\mathbf{P}$ is unknown and hence $\bar{\boldsymbol{\xi}}$ can be at most estimated given knowledge on $\mathbf{P}$ extracted from the information set $\Im_t = \{\mathbf{y}_j\}_{j=1}^{t}$. For simplicity we will also denote as $\bar{\boldsymbol{\xi}}$ such an estimated vector of ergodic (unconditional) state probabilities.

When n and/or m are large, model (1) implies the estimation of a large number of parameters, $k[(n+m) + p(n+m)^2 + (n+m)(n+m+1)/2 + (k-1)]$. For instance, for $k = 2$, $n = 3$, $p = 1$, and $m = 5$ (some of the hyper-parameters characterizing our application), this implies estimation of $2 \times [8 + 8^2 + 4 \times 9 + 1] = 218$ parameters![8] (1) nests a number of simpler models in which either some of parameter matrices are not needed or some of these objects eventually become regime-independent. This simpler model may greatly reduce the number of parameters to be estimated. Among them, we will devote special attention to $MSIH(k, p)$ models,

$$\mathbf{y}_t = \boldsymbol{\mu}_{S_t} + \boldsymbol{\Sigma}_{S_t}\boldsymbol{\epsilon}_t,$$

in which $p = 0$, $MSIA(k, p)$ homoskedastic models,

$$\mathbf{y}_t = \boldsymbol{\mu}_{S_t} + \sum_{j=1}^{p} \mathbf{A}_{j,S_t}\mathbf{y}_{t-j} + \boldsymbol{\Sigma}\boldsymbol{\epsilon}_t,$$

in which the covariance matrix is constant over time, and $MSIH(k, 0) - VAR(p)$ models,

$$\mathbf{y}_t = \boldsymbol{\mu}_{S_t} + \sum_{j=1}^{p} \mathbf{A}_j\mathbf{y}_{t-j} + \boldsymbol{\Sigma}_{S_t}\boldsymbol{\epsilon}_t, \tag{3}$$

[6]Assume the absence of roots outside the unit circle, thus making the process stationary. Ang and Bekaert (2002) have recently shown that formally, it is just sufficient for such a condition to be verified in at least one of the k alternative regimes, to obtain covariance stationarity. Stock and Watson (2003, p. 800) discuss the importance of assessing *marginal* predictive content, which naturally leads to specifying VAR(p) models in which lagged endogeneous variables play a role.

[7]The assumption of a first-order Markov process is not restrictive, since a higher order Markov chain can always be reparameterized as a higher dimensional first-order Markov chain, i.e. substitutability exists between the order of the Markov chain driving S_t and the number of regimes k.

[8]This is the sense in which Marron and Wand (1992) conclude that mixtures of normal distributions provide a flexible family that can be used to approximate many distributions. Mixtures of normals can also be viewed as a nonparametric approach to modeling the return distribution if the number of states, k, is allowed to grow with the sample size.

which are a special case of (1) in which while intercepts and covariance matrix are regime-dependent, the VAR(p) coefficients are not. For instance, model (3) implies the estimation of 'only' $k[(n+m)+(n+m)(n+m+1)/2+(k-1)]+p(n+m)^2$ parameters. For the same configuration mentioned above, this means $2\times[8+4\times 9+1]+8^2 = 154 < 218$. As we will see, this restricted sub-class of models turns out to be important to test the null hypothesis that predictability patterns involving US asset returns and macroeconomic variables are time-varying. Of course, a limit case of (1) is obtained when $k=1$:

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{j=1}^{p} \mathbf{A}_j \mathbf{y}_{t-j} + \boldsymbol{\Sigma}\boldsymbol{\epsilon}_t. \tag{4}$$

This is a standard multivariate Gaussian VAR(p) model, a benchmark in a large portion of the existing empirical macroeconomics and finance literature.

3.1. *Estimation and Inference*

The first step towards estimation and prediction of a MSIAH model is to put the model in state-space form. Collect the information on the time t realization of the Markov chain in a random vector

$$\boldsymbol{\xi}_t = \begin{bmatrix} I(S_t=1) \\ I(S_t=2) \\ \vdots \\ I(S_t=k) \end{bmatrix}$$

where $I(S_t=i)$ is a standard indicator variable. In practice the sample realizations of $\boldsymbol{\xi}_t$ will always consist of unit versors $\mathbf{e}_i$ characterized by a 1 in the i-th position and by zero everywhere else. Another important property is that $E[\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t-1}] = \mathbf{P}'\boldsymbol{\xi}_{t-1}$. The state-space form is composed of two equations:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{X}_t\boldsymbol{\Psi}\left(\boldsymbol{\xi}_t \otimes \boldsymbol{\iota}_{n+m}\right) + \boldsymbol{\Sigma}_M\left(\boldsymbol{\xi}_t \otimes \mathbf{I}_{n+m}\right)\boldsymbol{\epsilon}_t \qquad &\text{(measurement equation)} \\ \boldsymbol{\xi}_{t+1} &= \mathbf{F}\boldsymbol{\xi}_t + \mathbf{u}_{t+1} &\text{(transition equation)} \end{aligned} \tag{5}$$

where $\mathbf{X}_t$ is a $(n+m)\times(n+m)p+1$ vector of predetermined variables with structure $[1\ \mathbf{y}'_{t-1}...\mathbf{y}'_{t-p}]\otimes\boldsymbol{\iota}_n$, $\boldsymbol{\Psi}$ is a $[(n+m)p+1]\times(n+m)k$ matrix collecting the VAR parameters, both means and autoregressive coefficients,

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\mu}_1' & \cdots & \boldsymbol{\mu}_k' \\ \mathbf{A}_{11} & \cdots & \mathbf{A}_{1k} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{p1} & \cdots & \mathbf{A}_{pk} \end{bmatrix},$$

$\boldsymbol{\Sigma}_M$ is a $(n+m)\times(n+m)k$ matrix collecting all the possible k "square root" (Choleski decomposition) factors $[\boldsymbol{\Sigma}_1\ \boldsymbol{\Sigma}_2\ ...\ \boldsymbol{\Sigma}_k]$ such that $\forall t\ \boldsymbol{\Sigma}_M\left(\boldsymbol{\xi}_t\otimes\mathbf{I}_{n+m}\right)\left(\boldsymbol{\xi}_t\otimes\mathbf{I}_{n+m}\right)'\boldsymbol{\Sigma}_M' = \boldsymbol{\Omega}_{S_t}$, the covariance matrix of the asset return innovations $\boldsymbol{\epsilon}_t$. Moreover, $\boldsymbol{\epsilon}_t \sim NID(\mathbf{0},\mathbf{I}_{n+m})$, and in the transition equation $\mathbf{u}_{t+1}$ is a zero-mean discrete random vector that can be shown to be a martingale difference sequence. Also, the elements of $\mathbf{u}_{t+1}$ are uncorrelated with $\boldsymbol{\epsilon}_{t+1}$ as well as $\boldsymbol{\xi}_{t-j}$, $\boldsymbol{\epsilon}_{t-j}$, $\mathbf{y}_{t-j}$, and $\mathbf{X}_{t-j}$ $\forall j\geq 0$. To operationalize the dynamic state-space system (5), assume that the multivariate process (1) started with a random draw from the unconditional probability distribution defined by the vector of state probabilities $\bar{\boldsymbol{\xi}}$. Finally, from the definition of transition probability matrix (2) it follows that since $E[\mathbf{u}_{t+1}|\boldsymbol{\xi}_t]=0$ by assumption,

$$E[\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t] = \mathbf{F}\boldsymbol{\xi}_t$$

implies that $\mathbf{F}$ corresponds to the transposed transition probability matrix $\mathbf{P}'$.[9]

MSIAH models are estimated by maximum likelihood. In particular, estimation and inferences are based on the EM (Expectation-Maximization) algorithm proposed by Dempster et al. (1977) and Hamilton (1989), a filter that allows the iterative calculation of the one-step ahead forecast of the state vector $\boldsymbol{\xi}_{t+1|t}$ given the information set $\Im_t$ and the consequent construction of the log-likelihood function of the data.[10] Appendix A gives a few additional details. Maximization of the log-likelihood function within the M-step is actually made faster by the fact that the first-order conditions defining the MLEs may often be written down in closed form. Appendix B details the general form of such conditions. In particular, notice that the FOCs (16)-(17) all depend on smoothed probabilities $\hat{\boldsymbol{\xi}}_{t|T} \equiv \Pr\left(\boldsymbol{\xi}_t | \Im_T; \boldsymbol{\theta}, \boldsymbol{\rho}\right)$ and therefore they all present a high degree of non-linearity in the parameters $[\boldsymbol{\theta} \ \boldsymbol{\rho}]'$. Therefore the FOCs have to be solved numerically, although convenient iterative methods exist. In fact, the expectation and maximization steps can be used in iterative fashion. Starting with arbitrary initial values $\tilde{\boldsymbol{\theta}}^0$ and $\tilde{\boldsymbol{\rho}}^0$, the expectation step is applied first, thus obtaining a sequence of smoothed probability distributions $\{\hat{\boldsymbol{\xi}}^1_{t|T}\}_{t=1}^T$. Given these smoothed probabilities, (17) is then used to calculate $\hat{\rho}^1$, and (16) to derive $\tilde{\boldsymbol{\theta}}^1$. Based on $\tilde{\boldsymbol{\theta}}^1$ and $\tilde{\boldsymbol{\rho}}^1$, the expectation step can be applied again to find a new sequence of smoothed probability distributions $\{\hat{\boldsymbol{\xi}}^2_{t|T}\}_{t=1}^T$. This starts the second iteration of the algorithm. The algorithm keeps being iterated until convergence, i.e. until $[\tilde{\boldsymbol{\theta}}^l \ \tilde{\boldsymbol{\rho}}^l]' = [\tilde{\boldsymbol{\theta}}^{l-1} \ \tilde{\boldsymbol{\rho}}^{l-1}]'$. Importantly, the likelihood function increases at each step and reaches an approximate maximum in correspondence to convergence (see Baum et al. (1970)).

As for the properties of the resulting ML estimators, under standard regularity conditions (such as identifiability, stability and the fact that the true parameter vector does not fall on the boundaries) Hamilton (1989, 1993) and Leroux (1993) have proven consistency and asymptotic normality of the ML estimator $\tilde{\boldsymbol{\gamma}} = [\tilde{\boldsymbol{\theta}} \ \tilde{\boldsymbol{\rho}}]'$:

$$\sqrt{T}\left(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\right) \xrightarrow{d} N\left(\mathbf{0}, \mathcal{I}_a(\boldsymbol{\gamma})^{-1}\right)$$

where $\mathcal{I}_a(\boldsymbol{\gamma})$ is the asymptotic information matrix, $\mathcal{I}_a(\boldsymbol{\gamma}) \equiv \lim_{T\rightarrow\infty} -T^{-1}E\left[\frac{\partial^2 \ln \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}\right]$. Although other typical choices exist $-$ i.e. either the conditional scores or a numerical evaluation of the second partial derivative of the log-likelihood function with respect to $\tilde{\boldsymbol{\gamma}}$ $-$ in our application we are going to employ a 'sandwich' sample estimator of $\mathcal{I}_a(\boldsymbol{\gamma})$ providing $\widetilde{Var}(\tilde{\boldsymbol{\gamma}})$:

$$\widetilde{Var}(\tilde{\boldsymbol{\gamma}}) = T^{-1}\left[\mathcal{I}_2(\tilde{\boldsymbol{\gamma}})\left(\mathcal{I}_1(\tilde{\boldsymbol{\gamma}})\right)^{-1}\mathcal{I}_2(\tilde{\boldsymbol{\gamma}})\right],$$

where

$$\mathcal{I}_{11}(\tilde{\boldsymbol{\gamma}}) = T^{-1}\sum_{t=1}^T [\mathbf{h}_t(\tilde{\boldsymbol{\gamma}})]\,[\mathbf{h}_t(\tilde{\boldsymbol{\gamma}})]' \qquad \mathbf{h}_t(\tilde{\boldsymbol{\gamma}}) = \frac{\partial \ln p(\mathbf{y}_t|\Im_{t-1};\tilde{\boldsymbol{\gamma}})}{\partial\boldsymbol{\gamma}}$$

and

$$\mathcal{I}_2(\tilde{\boldsymbol{\gamma}}) = -T^{-1}\sum_{t=1}^T\left[\frac{\partial^2 \ln p(\mathbf{y}_t|\Im_{t-1};\tilde{\boldsymbol{\gamma}})}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}\right].$$

[9] In general this dynamic state-space model is neither linear (as the state vector $\boldsymbol{\xi}_t$ also influences the covariance matrix of the process) nor Gaussian, as the innovations driving the transition equation are non-Gaussian random variables.

[10] Some assumptions have to be imposed to guarantee at least the local identifiability of the parameters under estimation. One possibility is to extend results in Leroux (1992) to show that under the assumption of multivariate Gaussian shocks to the measurement equation, MSIAH models will be identifiable up to any arbitrary re-labeling of unobservable states.

As a consequence, and with one important exception, standard inferential procedures are available to test the statistical hypothesis.[11] In particular, call $\phi : \mathcal{R}^q \rightarrow \mathcal{R}^r$ a function that imposes $q - r$ restrictions on the q-dimensional parameter vector $\boldsymbol{\theta}$. We want to test $H_0 : \phi(\boldsymbol{\gamma}) = \mathbf{0}$ vs. $H_1 : \phi(\boldsymbol{\gamma}) \neq \mathbf{0}$ *under the assumption that under both hypotheses the number of regimes k is identical.* Define $\tilde{\boldsymbol{\theta}}_r$ as the restricted estimator, obtained under the null hypothesis. For instance, a test of the hypothesis of homoskedasticity ($H_0 : vech(\boldsymbol{\Sigma}_i) = vech(\boldsymbol{\Sigma}_k)$ $i = 1, 2, ..., k$) implies $r = (k-1)\frac{(n+m)(n+m+1)}{2}$ restrictions and can be formulated as a linear restriction on the matrix $\boldsymbol{\Sigma}_M$. The corresponding Likelihood Ratio (LR) test has asymptotic distribution:

$$LR \equiv 2\left[\ln L(\tilde{\boldsymbol{\theta}}) - \ln L(\tilde{\boldsymbol{\theta}}_r)\right] \xrightarrow{d} \chi_r^2.$$

In the following, we will frequently employ such a test. Finally standard t and F statistics can be calculated using the Wald test. Under asymptotic normality of the unrestricted ML estimator $\tilde{\boldsymbol{\theta}}$, it follows that

$$\sqrt{T}\left[\phi(\tilde{\boldsymbol{\theta}}) - \phi(\boldsymbol{\theta})\right] \xrightarrow{d} N\left(\mathbf{0}, \left.\frac{\partial\phi(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \widetilde{Var}(\tilde{\boldsymbol{\theta}}) \left.\frac{\partial\phi'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}\right) \text{ and}$$

$$Wald \equiv T\phi'(\tilde{\boldsymbol{\theta}})\left[\left.\frac{\partial\phi(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \widetilde{Var}(\tilde{\boldsymbol{\theta}}) \left.\frac{\partial\phi'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}\right]^{-1} \phi(\tilde{\boldsymbol{\theta}}) \xrightarrow{D} \chi_r^2.$$

3.2. *Forecasting*

Under a mean squared prediction error (MSFE) criterion, the required algorithms are relatively simple in spite of the nonlinearity of this class of processes. Let's start from the MSIAH(k,p) process (1). Ignoring for the time being the issue of parameter uncertainty, i.e. the fact that the parameters of the multivariate Markov switching process are unknown and must therefore be estimated, the function minimizing the MSFE is the standard conditional expectation function. For instance, for a one-step ahead forecast we have:

$$E[\mathbf{y}_{t+1}|\Im_t] = \mathbf{X}_{t+1}\hat{\Psi}\left(\hat{\boldsymbol{\xi}}_{t+1|t} \otimes \boldsymbol{\iota}_{n+m}\right)$$

where $\mathbf{X}_{t+1} = [1\ \mathbf{y}_t' ... \mathbf{y}_{t-p+1}'] \otimes \boldsymbol{\iota}_{n+m}$, $\hat{\Psi}$ collects the estimated conditional mean parameters, and $\hat{\boldsymbol{\xi}}_{t+1|t}$ is the one-step ahead, predicted latent state vector to be filtered out of the available information set $\Im_t$ according to transition equation

$$\hat{\boldsymbol{\xi}}_{t+1|t} = \hat{\mathbf{P}}'\hat{\boldsymbol{\xi}}_{t|t},$$

where also the transition matrix $\mathbf{P}$ will have to be estimated. It follows that[12]

$$E[\mathbf{y}_{t+1}|\Im_t] = \mathbf{X}_{t+1}\hat{\Psi}\left(\hat{\mathbf{P}}'\hat{\boldsymbol{\xi}}_{t|t} \otimes \boldsymbol{\iota}_{n+m}\right). \tag{6}$$

[11] The exception concerns the *number of non-zero rows* of the transition matrix $\mathbf{P}$, i.e. the number of regimes k. In this case, even under the assumption of asymptotic normality of the estimator $\tilde{\gamma}$, standard testing procedures suffer from non-standard asymptotic distributions of the likelihood ratio test statistic due to the existence of nuisance parameters under the null hypothesis. We defer the discussion of this important and challenging inferential procedure to Section 5.

[12] For $h > 1$-steps ahead forecasts the task is much more challenging as: (1) $\mathbf{X}_{t+h}$ is unknown and must be predicted itself; (2) $E[\mathbf{X}_{t+T}|\Im_t]$ involves sequences of predictions $\{E[\mathbf{y}_{t+1}|\Im_t], ..., E[\mathbf{y}_{t+T-1}|\Im_{t+T-2}]\}$ and as such $\{\hat{\boldsymbol{\xi}}_{t+1|t}, ..., \hat{\boldsymbol{\xi}}_{t+T-1|t}\}$ which are likely to impress patterns of cross-correlation to the unconditional values of the parameters to be used, because of the presence of regime switching.

4. **The Data**

We use monthly data for the longest available period, 1926:12 - 2004:12, for a total of 937 observations per each time series. Financial returns data are from the Center for Research on Security Prices (CRSP) at the University of Chicago. In particular, we employ in this paper data on the three most important segments of the US financial market: stocks (value-weighted stock returns for the NYSE, NASDAQ, and the AMEX exchanges, for the periods in which related data are available), bonds (a CRSP index of 10-year to maturity long-term US government bonds), and money market instruments (30-day Treasury bills, again taken from the CRSP monthly interest rates database).[13]

Additionally, we employ 5 predictor variables which are either directly identifiable with important macroeconomic aggregates, or that at least have been identified with / associated to general business cycle conditions in previous research. In the first group we have the CPI inflation rate (seasonally adjusted), the rate of growth of industrial production (seasonally adjusted), and the rate of growth of a measure of adjusted monetary base. These three series are available at FRED® II, the data collection at the Federal Reserve Bank of St. Louis. The practice of seasonally adjusting the data in real time experiments (see section 7) assumes that market participants are effectively able to 'see through' the veil of time series variation purely caused by seasonal factors. In the latter group we have two variables. The first is the dividend yield, once made available by CRSP, and defined as aggregate dividends on the value-weighted CRSP portfolio of stocks over the previous twelve month period divided by the current stock price. The second is the default spread, defined as the differential yield on Moody's Bbb (low rating) and Aaa (high rating) seasoned corporate bond securities with similar maturities. These two variables have played a key role both in the seminal work by Fama and French (1988, 1989) on stock and bond returns predictability, as well as in the recent literature on optimal asset allocation under predictable asset returns, see e.g. Campbell and Viceira (1999) and Brandt (1999).

In our empirical analysis we use the following transformed variables. Given their crucial role in financial decisions, we study *real* stock and T-bill returns, defined as the difference between nominal, realized monthly returns and the inflation rate. For similar reasons, given the important literature on term spreads in the US yield curve, we use the long-short bond term spread (which is a notion of term premium) defined as the difference between the CRSP long-term bond and 1-month T-bill returns. Finally, also IP growth and monetary base growth are measured in real terms, by deducting from nominal rates of growth the realized IP inflation rate.

Tables 1 and 2 report a few basic summary statistics concerning the series under investigation. Mean values are consistent with commonly known facts: for instance, the mean excess stock return is 0.65% per month, i.e. 7.80% per year, which represents a typical value in the equity premium literature, with an annualized volatility of 19.1%; the mean term premium is 0.14% per month, i.e. 1.68% per year, a moderate but plausible average slope of the US term structure; the average annualized real T-bill rate is 0.60% which, summed to a mean annualized inflation rate of 3.12%, delivers a mean annualized nominal short-term interest rate of 3.72%, once more in line with the typical values reported in the asset pricing literature. Both real money and industrial production growth are positive on average, 0.48 and 2.52 percent in annualized terms, respectively. All the series display evident departures from a (marginal) Gaussian distribution, which would

[13]The bond returns data are completed by using the Ibbotson-Sinquifeld data over the period 1926-1946, whose criterion of construction are anyway perfectly homogeneous to the ones employed by CRSP for the early part of the sample.

imply zero skewness (i.e. a symmetric distribution) and a kurtosis coefficient of 3. On the opposite, both excess stock returns and all macroeconomic variables are characterized by huge kurtosis values (in excess of 10), an indication of distributions with tails considerably fatter than a normal. The dividend yield has only moderate kurtosis, but it is also skewed to the right (which is to be expected, since the dividend yield cannot be negative by construction). Even in the case of excess bond returns, a formal Jarque-Bera test of marginal normal distribution rejects with a 0.000 p-value. Finally, for all series but one (excess bond returns) there is evidence of strong and statistically significant first-order serial correlation, as evidenced by Portmanteau Ljung-Box statistics (of order 4) in excess of the 1% critical value under a $\chi^2_{(4)}$. Similarly, there is evidence of volatility clustering (heteroskedasticity), as all Ljung-Box statistics (of order 4) applied to squared values of the variables are highly significant. Table 2 reports the correlation coefficients between pairs of series. Although many coefficients are statistically different from zero, the largest correlations are between the dividend yield and the default spread (positive) and between real 1-month T-bill rate and the inflation rate (negative, imperfect reaction of short-term rates to inflation). Figure 1 completes the picture showing the series. To allow a Reader to familiarize with those long data series, we also superimpose as gray intervals the periods that have been officially dated as recessions by the NBER. Visibly, most series tend to be particularly volatile in the first part of the sample and especially during the Great Depression, 1930-33.

5. **Empirical Results**

5.1. *Model Selection*

The first important task in any empirical analysis is the selection of an appropriate econometric model to represent the dynamic linkages between asset returns and macroeconomic forces. Since testing the hypothesis that the predictability patterns involving macroeconomic and financial variables has been stable over time revolves around achieving a best possible specification of a sufficiently rich model such as (1), we make an extensive effort. We estimate a large number of variants of (1) and use six alternative criteria to gauge the correct specification of the candidate models. In the following we separately describe each of these criteria and their results. Tables 3-4 and Figures 2-5 show related results.

The first selection criterion applied in Table 3, third and fourth columns, concerns the appropriate number of regimes k in model (1).[14] In particular, we would like to test whether the null of a single-state model ($k = 1$) can be rejected in favor of $k > 1$. As already stressed, when $k = 1$, (1) reduces to a simpler Gaussian VAR(p) model. As discussed in Garcia (1998), testing for the number of states in a regime switching framework may be tricky. Given some $k \geq 2$, the problem is that under any number of regimes smaller than k there are a few parameters of the unrestricted model $-$ some (or all) elements of the transition probability matrix associated to the rows that correspond to "disappearing states", plus corresponding parameters in the conditional mean and/or heteroskedasticity functions — that can take any values without influencing the maximized likelihood function. We say that these parameters become a nuisance to the estimation. The result is that the presence of nuisance parameters gives the likelihood surface so many degrees of freedom that computationally one can never reject the null that the non-zero values of those parameters were purely

[14]In the table, the switching models are classified as MSIAH(k, p), where I, A and H refer to state dependence in the intercept, vector autoregressive terms and heteroskedasticity. p is the autoregressive order. Models in the class MSIH(k, 0)-$VAR(p)$ have regime switching in the intercept but not in the VAR coefficients. A MSI(1,0) is a simple multivariate Gaussian IID model; a MSIA(1,p) is a Gaussian VAR(p) model.

due to sampling variation. Additionally, likelihood ratio (LR) tests in the presence of nuisance parameters imply that even asymptotically the LR statistic fails to have a standard chi-square distribution with number of degrees of freedom equal to the number of restrictions imposed.

Since we would like to employ the LR principle to test the null of $k = 1$, we adopt two different strategies to deal with nuisance parameter issues. Davies (1977) circumvents the problem of estimating the nuisance parameters under the alternative hypothesis and derives instead an upper bound for the significance level of the LR test under nuisance parameters:[15]

$$\Pr\left(LR > x\right) \leq \Pr\left(\chi_1^2 > x\right) + \sqrt{2x}\exp\left(-\frac{x}{2}\right)\left[\Gamma\left(\frac{1}{2}\right)\right]^{-1}.$$

where $\Gamma(\cdot)$ is the standard gamma function. Therefore the fourth column of Table 3 systematically tests the null of $k = 1$ against $k > 1$ (the exact number of regimes varies with the different models) and reports p-values calculated under Davies' upper bound. Obviously, even adjusting the presence of nuisance parameters, the evidence against specifying traditional single-state models is overwhelming: the smallest LR statistic takes a value of 139, which is clearly above any conceivable critical value regardless of number of restrictions imposed. A related test is proposed by Wolfe (1971) and applied by Turner, Startz, and Nelson (1989). The modified LR test is:

$$LR^{Wolf} = -\frac{2}{T}(T-3)\left[\ln L(\tilde{\boldsymbol{\gamma}}) - \ln L(\tilde{\boldsymbol{\gamma}}_r)\right] \xrightarrow{d} \chi_r^2$$

where $\tilde{\boldsymbol{\gamma}}_r$ is obtained under the null of simple multivariate normality, T is the sample size, and $r = k(k-1)$ since in the absence of regime switching there are $k(k-1)$ which cannot be estimated. We also calculate this special adjusted LR tests (not reported to save space) and find anyway at least triple digit values, which again points to overwhelming rejections of the null of one regime only. This gives a first, crucial implication: the data propose strong evidence of time-variation in the coefficients characterizing models apt to capture the dynamic linkages between financial and macroeconomic variables in the US. Notice that however this does not yet imply that patterns of predictability can be effectively treated as time-varying.

Once we establish that $k \geq 2$ is appropriate, this only rules out models of type $MSI(1,p)$, i.e. the first few rows of Table 3 only. We therefore proceed to select an appropriate model within the more general regime switching class $MSIAH(k,p)$ with $k \geq 2$. As in a few other applied papers on regime switching models (e.g. Sola and Driffill (1994) and Guidolin and Timmermann (2005b)), to this purpose we employ a battery of information criteria, i.e. the Akaike (AIC), Bayes-Schwartz (BIC), and Hannan-Quinn (H-Q) criteria. These criteria are supposed to trade-off in-sample fit with prediction accuracy and rely on the principle that a correctly specified model should not only provide an accurate in-sample fit of the data at hand, but also prove useful to precisely forecast out-of-sample. In practice, information criteria identify the ex-ante potential out-of-sample by penalizing models with a large number of parameters. A well-performing model ought to minimize each of the information criteria. The range of models estimated in Table 3 is wide and spans models with $k = 2, 3, 4$, $p = 1, 2$, and with and without a regime-dependent covariance matrix. When possible, also models like (3) are estimated, since they are relatively parsimonious as well as

[15] Hansen (1992) proposes to compute the likelihood as a function of the unknown and non-estimable nuisance parameters so that the asymptotic distribution is generated in each case numerically from a grid of regime-dependent nuisance parameters. The test statistic becomes then $LW_T \leq \sup_{\boldsymbol{\rho}} LW_T(\boldsymbol{\rho})$, where the right hand side converges in distribution to a function of a Brownian bridge. In most of the cases a closed form expression cannot be found and the bound must be calculated by simulation and becomes data-dependent. This task is clearly overwhelming for multivariate regime switching models and will not be pursued here.

economically interesting, implying the contemporaneous presence of regimes (in intercepts and covariances) along with dynamic linkages which are constant over time.[16] Columns 5-7 of Table 3 show that some tension exists among different criteria. The AIC is minimized by a richly parameterized MSIAH(4,1) model in which 444 parameters have to be estimated. We notice that although the MLE estimation could be carried out, issues may exist with a model that implies a saturation ratio (i.e. the number of available observations per estimated parameter) of only 16.9.[17] However, this is less than surprising as the AIC is generally known to select large models in nonlinear frameworks (see e.g. Fenton and Gallant (1996)). Next, the H-Q seems to be undecided between a relatively parsimonious MSIH(4,0)-VAR(1) model (with saturation ratio of almost 30) and a richer MSIAH(3,1) (saturation ratio of 23). Notice that these two models imply a different number of regimes, 3 vs. 4. So, if on the one hand it seems obvious that regime switching matters, the precise number of states required seems to be debatable. Finally, the BIC selects once more a relatively tight MSIH(4,0)-VAR(1) model.

The last column of Table 3 shows the outcomes of standard LR tests within classes of models characterized by the same number of regimes, i.e. for which nuisance parameter problems do not exist so that standard asymptotic results apply. The column should be read as testing the null that augmenting a smaller model by a certain feature − either increasing p to the higher integer (A) or making the covariance matrix regime-dependent (H) − does not significantly increase the maximized log-likelihood. For instance, in the row of the MSIA(3,1) model we read: 'A: 0.000' to imply that going from a MSIH(3,0) to a MSIAH(3,1) the log-likelihood increases by more than one would impute to random chance; 'H: 0.000' to imply that the log-likelihood increase caused by a move from a MSIA(3,1) to a MSIAH(3,1) is highly significant. As previously observed in other nonlinear estimation contexts (see e.g. Gallant and Tauchen (1997)), LR tests tend to be not very selective, as they fail to trade-off in-sample fit for parsimony. On any account, we find evidence that relatively rich regime switching models are required to fit the data at hand. In particular, the hypothesis of no need of at least a VAR(1) component as well as of regime-specific covariance matrices is always rejected using the LR test.

All in all, we are left with two plausible and competing candidate models. The first one is a four-regime MSIH(4,0)-VAR(1) model that is directly selected by both the H-Q and the parsimonious BIC criterion. The second is a three-regime MSIAH(3,1) model that obtains a good 'score' in a H-Q metric. We elect not to pursue estimation of the richer MSIAH(4,1) (selected by the AIC) because of the high probability of it being over-parameterized (its saturation ratio is almost half the MSIH(4,0)-VAR(1)). Notice at this point that these two models are structurally different both in a statistical and in an economic sense:

- MSIAH(3,1) is obviously a three-regime model while MSIH(4,0)-VAR(1) is a four-regime model. It is clearly important to understand how many different regimes can be reliably singled out in the dynamic law of motions of US asset returns and a number of fundamental macroeconomic variables. Although this is first of all a difference in statistical properties, it is clear that economic implications may differ too.

- There is an additional and deeper difference: MSIAH(3,1) relies on regime switching VAR(1) coeffi-

[16] On the other hand, it appears unadvisable to try and estimate models with $k \geq 5$ since the number of parameters quickly grows to levels that make either estimation uncertainty overwhelming or that cause the MLE-EM routines to fail. For instance, a MSIAH(5,1) model implies 560 paramters, i.e. only 13.4 observations per parameter.

[17] A commonly applied rule of thumb proposes that nonlinear estimation results based on saturation ratios inferior to approximately 20 ought to be taken with a high dose of caution.

cients, i.e. in this model the dynamic linkages between financial markets and the macroeconomy are strictly regime-dependent and therefore time-varying. This means that both the ways in which the current macroeconomic stance predicts subsequent asset prices and in which asset returns may possibly forecast future economic conditions may have changed with time as the US economy has evolved over the 8 decades spanned by our sample. On the contrary, MSIH(4,0)-VAR(1) implies constant VAR(1) coefficients and hence time homogeneous predictability among financial and macroeconomic variables.

Since both differences appear crucial both under an econometric and an economic perspective, Section 5.2 uses additional, powerful tools to select the best model between MSIAH(3,1) and MSIH(4,0)-VAR(1).

5.2. *Density Specification Tests*

As we have discussed in Section 3, regime switching model consists of flexible mixture models which − if the number of regimes k is expanded with the sample size − may be thought of as providing a seminonparametric approximation of the process followed by the joint conditional density of the data, see Marron and Wand (1992). In this framework, it has become customary to require that regime switching models provide a correct specification for the entire conditional distribution of the variables at hand. In particular, notice that any decision maker will want to use a relatively sophisticated multivariate, multi-state model only if it provides a useful approximation not only of a few conditional moments of interest, but for the entire nonlinear dynamic process generating the data. Fortunately, the seminal work of Diebold et al. (1998) has spurred increasing interest in specification tests based on the h-step ahead accuracy of fit of a model for the underlying density. These tests are based on the probability integral transform or z-score. This is the probability of observing a value smaller than or equal to the realization of returns, $\tilde{\mathbf{y}}_{t+1}$, under the null that the model is correctly specified. Under a k-regime mixture of normals, this is given by

$$\begin{aligned}
\Pr\left(\mathbf{y}_{t+1} \leq \tilde{\mathbf{y}}_{t+1} | \Im_t\right) &= \sum_{i=1}^{k} \Pr\left(\mathbf{y}_{t+1} \leq \tilde{\mathbf{y}}_{t+1} | \Im_t, S_{t+1} = i\right) \Pr(S_{t+1} = i | \Im_t) \\
&= \sum_{i=1}^{k} \Phi_{m+n}\left(\boldsymbol{\Sigma}_i^{-1}\left[\mathbf{y}_{t+1} - \boldsymbol{\mu}_i - \sum_{j=1}^{p} \mathbf{A}_{j,i}\mathbf{y}_{t+1-j}\right]\right) \Pr(S_{t+1} = i | \Im_t) \\
&\equiv z_{t+1} \in \mathcal{R}, \qquad (7)
\end{aligned}$$

where $\Phi_{m+n}(\cdot)$ is the standard $(m+n)-$variate normal cdf. As already stressed by Rosenblatt (1952), if the model is correctly specified, z_{t+1} should be independently and identically distributed (IID) and uniform on the interval $[0, 1]$. The uniform requirement relates to the fact that deviations between realized values and projected (fitted) ones should be conditionally normal and as such describe a uniform distribution once it 'filtered through' an appropriate Gaussian cdf. The IID condition reflects the fact that if the model is correctly specified, errors ought to be unpredictable and fail to show any detectable structure. Tests are performed in sample, i.e. comparing predicted (fitted) values with data.[18]

Unfortunately, testing whether a distribution is uniform is not a simple task, as test statistics popular in the statistics literature often rely on the IID-ness of the series, which is at stake here as well. Therefore Berkowitz (2001) has recently proposed a likelihood-ratio test that inverts Φ to get a transformed z-score,

$$z^*_{t+1} \equiv \Phi^{-1}(z_{t+1}),$$

[18] Section 6 performs a few genuine out-of-sample tests.

which essentially transforms the z-score back into a bell-shaped random variable. Provided that the model is correctly specified, z^* should be IID and normally distributed (IIN(0, 1)). We follow Berkowitz (2001) and use a likelihood ratio test that focuses on a few salient moments of the return distribution. Suppose the log-likelihood function is evaluated under the null that $z^*_{t+1} \sim IIN(0,1)$:

$$L_{IIN(0,1)} \equiv -\frac{T}{2}\ln(2\pi) - \sum_{t=1}^{T}\frac{(z_t^*)^2}{2}.$$

Under the alternative of a misspecified model, the log-likelihood function incorporates deviations from the null, $z^*_{t+1} \sim IIN(0,1)$:

$$z^*_{t+1} = \alpha + \sum_{j=1}^{q}\sum_{i=1}^{l}\beta_{ji}(z^*_{t+1-i})^j + \sigma u_{t+1}, \tag{8}$$

where $u_{t+1} \sim IIN(0,1)$. The null of a correct return model implies $q \times l + 2$ restrictions – i.e., $\alpha = \beta_{ji} = 0$ ($j = 1, ..., q$ and $i = 1, ..., l$) and $\sigma = 1$ – in equation (8). Let $L(\hat{\alpha},\ \{\hat{\beta}_{ji}\}_{j=1}^{q}{}_{i=1}^{l},\ \hat{\sigma})$ be the maximized log-likelihood obtained from (8). To test that the null model (some version of (1)) is correctly specified, we can then use the following test statistic:

$$LR_{ql+2} \equiv -2\left[L_{IIN(0,1)} - L(\hat{\alpha}, \{\hat{\beta}_{ji}\}_{j=1}^{q}{}_{i=1}^{l}, \hat{\sigma})\right] \xrightarrow{d} \chi^2_{ql+2}.$$

In addition to the standard Jarque-Bera test that considers skewness and kurtosis in the z-scores to detect non normalities in z^*_{t+1}, it is customary (see e.g. Guidolin and Timmermann (2004)) to present three likelihood ratio tests, namely a test of zero-mean and unit variance ($q = l = 0$), a test of lack of serial correlation in the z-scores ($q = 1$ and $l = 1$) and a test that further restricts their squared values to be serially uncorrelated in order to test for omitted volatility dynamics ($q = 2$ and $l = 2$). Notice that a rejection of the null of normal transformed z-scores has the same meaning as rejecting the null of a uniform distribution for the raw z-scores, i.e. the model fails in generating a density with the appropriate shape. A rejection of the zero-mean, unit variance restriction points to specific problems in the location and scale of the density underlying the model. A rejection of the restriction that $\{z^*_{t+1}\}$ is IID (i.e. the presence of serial correlation in levels of squares) points to dynamic misspecifications. In particular:

- When $\{z^*_{t+1}\}$ appears serially correlated, then it is likely that the model has omitted some source of persistence in the series under investigation.
- When $\{z^*_{t+1}\}$ displays correlation in squares (i.e. volatility clustering), it is possible that the model has misspecified some sources of heteroskedasticity, i.e. of persistence in either volatilities or correlations.

Sometimes density specification tests are applied not to the vector $\mathbf{y}_{t+1}$, but to each of its individual components:

$$\begin{aligned}
\Pr\left(y^h_{t+1} \leq \tilde{y}^h_{t+1}|\Im_t\right) &= \sum_{i=1}^{k}\Pr\left(y^h_{t+1} \leq \tilde{y}^h_{t+1}|\Im_t, S_{t+1} = i\right)\Pr(S_{t+1} = i|\Im_t) \\
&= \sum_{i=1}^{k}\Phi\left(\sigma^{-1}_{h,i}\left[y^h_{t+1} - \mu_{h,i} - \mathbf{e}'_h\sum_{j=1}^{p}\mathbf{A}_{j,i}\mathbf{y}_{t+1-j}\right]\right)\Pr(S_{t+1} = i|\Im_t) \\
&\equiv z^h_{t+1} \qquad h = 1, ..., m+n,
\end{aligned}$$

where $\sigma_{h,i}$ is the volatility of variable h in state i, and $\mathbf{e}_h$ is a vector with a one in position h and zeros elsewhere. The reason for this choice is that – especially when the dimension of $\mathbf{y}_{t+1}$ is high (like $m+n=8$) – rejections from the z-scores based on the generalized multivariate residuals in (7) may often provide limited information on which variables are responsible for the rejection, i.e. of the dimensions over which the regime switching model is failing. This is also our choice in this paper.

Table 4 reports Berkowitz-style, transformed z-score tests for three models: a benchmark, linear Gaussian VAR(1) which can be taken to correspond to the bulk of the existing literature on the linear predictability of financial and macroeconomic variables; of course, the MSIAH(3,1) and MSIH(4,0)-VAR(1) model, respectively. Strikingly, a simple yet popular VAR(1) is resoundingly rejected by *all* tests and for *all* variables, financial and macroeconomic ones. Rejections tend to be harsh: the highest VAR(1) p-value appearing in the table is 0.001, i.e. there is actually a very thin chance that the data might have been generated by a simple linear Gaussian homoskedastic model. In fact, the rejections are so strong that it becomes even difficult to understand in which direction one should be moving to amend the VAR(1) model to improve the in-sample performance.

The picture improves, albeit not drastically, when a MSIAH(3,1) model is estimated. For most tests and variables, the LR test statistics decline by a factor between 30 and 200% when we move from a single- to a multi-state model. The exceptions are few (the Jarque-Bera tests for excess bond returns, the inflation rate, and adjusted monetary base growth). However all (but one, for the real T-bill rate and when testing the zero-mean unit-variance properties) of the related p-values remain highly significant, indicating strong rejection of the null of correct specification of the three-state model in which the predictability patterns are time-varying. This means that specifying $k=3$ and allowing the VAR(1) coefficients to change with the regime produces a density which is structurally different from the density that has generated the data.

Figures 2 and 3 provide further evidence on the sources of misspecifications within a MSIAH(3,1) model. Figure 2 shows the empirical distribution of $\{z^*_{t+1}\}$ for each of the eight variables (continuous lines) and compares it with a corresponding normal variate with matched mean and variance (dotted lines). The failures of the model are obvious for most of the variables, with the exceptions of T-bill short-term real yields and (possibly at least) the default spread. In many cases, the empirical score distributions are either leptokurtic (too much mass at the center and in the tails, e.g. excess stock returns, CPI inflation, and IP and monetary base real growth rates) or even multi-modal (excess bond returns and the dividend yield). Figure 3 provides a different visual viewpoint by displaying quantile-quantile (q-q) plots for each of the eight variables. Notice that if $\{z^*_{t+1}\}$ is $N(0,1)$, the q-q should approximately look like a 45-degree straight line in the q-q planes. This seems to happen only for real T-bill yields. On the other hand, a few plots assume an S-shape, i.e. the slope is too high at the center of the distribution (i.e. more mass is put under the distribution of $\{z^*_{t+1}\}$ than under a $N(0,1)$) and too flat for intermediate values in the support (where mass is missing vs. the Gaussian case). This is the case of excess stock and bond returns, inflation, real IP and monetary growth rates. In two other cases – dividend yield and the default spread – the q-q plots are simply flatter than a 45 degrees line, a sign that too much mass is simply moved to the tails of the corresponding distributions.

On the contrary, the improvement is strong and significant when we fit a four-state model. The p-values associated with the various tests generally increase and out of 32 combinations among tests/variables, we have that the null of no misspecification fails to be rejected in 15 cases, with p-values exceeding 0.05. Of the remaining 17 tests, in 7 the p-values are between 0.01 and 0.05, i.e. the rejection is rather mild. However, the (marginal) conditional density of real T-bill yields and the default spread remains hard to capture using a regime switching model: for these two variables (which generate 7 of the 10 highly significant rejections) there are signs of consistent departures from normality, of serial correlation in the scores, and of volatility clustering. The bright side is that for 3 variables – remarkably all of the financial variables, including the dividend yield – the tests give evidence of correct specification, with only some weak hesitations caused by the potential presence of additional volatility clustering not simply accommodated by regime switching covariance matrices in excess stock returns, and by deviations of the shape of the bond and dividend yield scores from normality.[19] Moreover, the improvement vs. the three-state model is clear: in only one case the LR statistic increases when the number of regimes is increased (the test for correct location and scale of the distribution of the transformed z-scores) and the variation in the corresponding p-value is moderate. Figures 4 and 5 visualize the marked improvement. Whilst Figure 4 depicts the existence of some residual deviations for the default spread, with the exception of this case Figure 5 presents almost perfect q-q plots, i.e. roughly aligned around a 45 degrees line.[20]

Finally, we apply density specification tests also at the multivariate level. We obtain a Jarque-Bera statistic of 6.45 (p-value of 0.040) and

$$LR_2 = 5.56 \quad \text{p-value: } 0.062$$
$$LR_3 = 7.44 \quad \text{p-value: } 0.059$$
$$LR_6 = 13.04 \quad \text{p-value: } 0.042.$$

Even if some issues remain concerning the overall shape of the distribution of the scores and possibly omitted heteroskedasticity (both natural since the marginals of two variables are not completely accounted for), this is considerable evidence in favor of four states over three. Therefore in the following we estimate and comment on the MSIH(4,0)-VAR(1) model. Ultimately, what decides of the actual usefulness and precision of a model is its ability to forecast future outcomes, see Section 6.

5.3. *A Four-State Model*

Before commenting on the nature of the estimated MSIH(4,0)-VAR(1) model, it is crucial to stress what the selection of this model over a three-state MSIAH(3,1) means for the main thesis of this paper. The four-state model in Table 5 implies that although multiple regimes are required to approximate the joint conditional density of financial returns and macroeconomic variables in the US, there is also no evidence of time-variation in the structure of the predictability patterns linking financial markets and the economy at large. This means that although both expected returns and overall monetary and economic conditions (as captured by inflation, real industrial production growth, and possibly the default and term spread) have been subjective to recurring structural breaks, the dynamic linkages between financial prices and such monetary and economic conditions have been remarkably stable over time. This is quite a remarkable and novel result: in spite of the complete flexibility of model (1) in terms of making the vector-autoregressive coefficients a function of the underlying latent regime, such an hypothesis is rejected by the BIC, put in some doubt by the H-Q, and again strongly rejected by the density specification tests in Section 5.2. In plain terms, this means

[19] The good results for financial variables confirm Ang and Bekaert's (2002) and Guidolin and Timmermann's (2004) findings that regime switching models provide an excellent description of the dynamic behavior of asset returns.

[20] In Figure 4 minor problems persist for the dividend yield, which seems to show two modes. However, Table 4 makes it clear that dynamic misspecifications are the most relevant cause of rejection of the null of correct specification for the dividend yield.

that there does not seem to be a useful (and statistically sound) way to differentiate between the measured response of stock prices to inflation news or of IP growth to movements in the bond term spread (just to cite two among the many interesting links) during the 1929 crash, the rapid growth of the post-WWII period, and the booming economy of the 1990s.[21]

Table 5 presents parameter estimates. Panel A reports estimates of a benchmark, single-state VAR(1) model. Panel B shows MLE-EM estimates of the four-state model. Panel A shows that in a standard VAR(1) many (if not the majority!) of the estimated coefficients are in fact non-significant. The implications for predictability of financial returns are rather interesting: apart from a weak own serial correlation (coefficient is 0.10), excess stock returns are essentially unpredictable using any of the macro instruments entertained in the paper. A minor exception is the real rate of growth of the monetary base (coefficient 0.13), as in Homa and Jaffee (1971). The same is true for excess bond returns, which can be just (weakly) predicted from past excess stock returns (coefficient -0.02). Much more predictability characterizes real short-term interest rates, which are (as expected) highly persistent (coefficient 0.58) and can also be predicted off past default spread (coefficient 1.28) and real IP growth (-0.06). Finally, there is again limited evidence of past asset returns predicting inflation (from past real T-bill yields) and real growth (from both excess stock returns, with a positive coefficient), as in Lee (1992) or Plosser and Rouwenhorst (1994).

However, only limited confidence should be attributed to these results for three reasons. First, we know from Section 5.1 that single-state VAR(1) models are rejected even when account is taken of nuisance parameter issues. If there are multiple regimes in the data, we can expect that all estimates obtained from single-regime models might be biased and therefore irrelevant. Second, notice that even the significant VAR coefficients in Panel A of Table 5 are often rather small. For instance, a one-standard deviation increase in rate of growth of the real adjusted monetary base would translate (ceteris paribus, i.e. assuming that such a shock could be identified in the economy without other contemporaneous effects) into a 0.29% additional excess stock return and a 0.09% increase in real IP growth, both rather negligible in economic terms. Even a one standard deviation increase in the default spread would cause a decline in subsequent inflation of 0.08% only. Third, the fit provided by a VAR(1) model is way less-than-perfect. In Table 6 we calculate unconditional means from the VAR(1) using the standard results that under stationarity

$$E[\mathbf{y}_{t+1}] = (\mathbf{I}_{n+m} - \mathbf{A})^{-1}\boldsymbol{\mu},$$

as well as period-specific means of the type

$$E_{\tau_0 \to \tau_1}[\mathbf{y}_{t+1}] = \frac{1}{(\tau_1 - \tau_0 + 1)} \sum_{t=\tau_0}^{\tau_1} \left[\boldsymbol{\mu} + \sum_{j=1}^{p} \mathbf{A}_j \mathbf{y}_{t-j} \right],$$

where $\{\mathbf{y}_{t-j}\}_{t=\tau_0}^{\tau_1}$ consists of the values actually realized over the sample $[\tau_0, \tau_1]$. For convenience we isolate four sub-periods, 1926-1946, 1947-1966, 1967-1986, and 1987-2004. When τ_0 = 1926:12 and τ_1 = 2004:12, the mean corresponds to the full sample period. Comparing panels A (data) and B (VAR(1)) of Table 6, it is clear that $E_{26\to 04}[\mathbf{y}_{t+1}]$ under the VAR does match the data sample means but: (i) this does not apply to unconditional means, i.e. over the long run the VAR(1) forecasts values for the variables that are often

[21] The only related finding we are aware of is Estrella et al.'s (2003) conclusion that on a cross-section of countries, when the term structure forecasts real activity, the relationship seems to be stable over time. We specialize our investigation to the US, but extend the analysis to a number of financial and macroeconomic variables.

radically different from those observed on average in-sample;[22] (ii) there are several sub-samples in which the observed means are radically different than the $E_{\tau_0 \to \tau_1}[\mathbf{y}_{t+1}]$ implied by the VAR. For instance, the VAR implies too high excess stock returns in 1926-1946 and too low stock returns in 1987-2004, a potentially costly error in decision-making terms. Similarly, the real IP growth is grossly over-estimated over the first part of the sample while the opposite occurs with reference to the last 18 years. In a sense, a VAR(1) model presents a 'rosy' picture of the Great Depression and misses altogether the stable period of phenomenal growth and of bull market of the 1990s. This is not completely surprising as one of the roles of regime switching is to accommodate within the mixture extremely bad and good periods as separate states.

Things greatly improve under a well-specified regime switching model. Table 5, panel B, starts by showing that the fraction of parameters in the conditional mean function that get precisely estimated substantially grows when multiple states are allowed. For instance, most of the intercepts are now significant or highly significant. However the most visible changes concern indeed the amount (and in some cases, the structure) of the predictability patterns implied by the model. On one hand, modeling regimes erases all traces of own- and cross-serial correlation involving excess asset returns. This is unsurprizing as structural breaks (regimes) are well known to artificially inflate the degree of persistence of series. On the other hand, excess stock and bond returns become now highly predictable using lagged values of three variables: real T-bill rates (which forecast lower excess returns, since the real short-term rate enters the discount rate in asset pricing models) as in Campbell (1987), the default spread (which forecasts higher excess returns, as a reward to increased risk premia, as in Fama and French (1989)), and the inflation rate (which forecasts lower future excess returns, presumably as a consequence of the recessions that need to be induced to bring inflation under control) as in Fama and Schwert (1977). Excess stock returns are also predicted by past IP growth (as in Cutler, Poterba and Summers (1989)), although the economic effect is almost negligible; similarly, real T-bill yields are partially predictable from past default spreads (even if a one standard deviation in the spread only increases T-bill real rates by 0.02%).

We also obtain evidence of predictability of macro variables, especially real IP growth which is not only persistent, but also forecasted by past T-bill yields and default spreads with significant coefficients, similarly to Bernanke (1983) and Stock and Watson (1989). CPI inflation remains predictable from past interest rates and the default spread, besides being highly persistent. Interestingly, the ability of asset returns to predict macroeconomic conditions − principally future real growth − seems confined to single-regime models, see e.g. James, Koreisha, and Partch (1985). In this sense, our results are similar to Canova and De Nicolo's (2000). When $k = 4$ most coefficients lose significance. This means that the bulk of predictability for inflation and real growth comes from the regime switching structure (and past macroeconomic conditions), and not from financial markets.[23]

Table 5 also reports the regime-dependent estimated volatilities and pairwise correlations implied by estimated variances and covariances. With limited exceptions, regimes 1 and 2 are characterized by moderate volatilities of the shocks and by correlations which tend to be smaller (in absolute value) than in the single-state VAR(1) model of Panel A. Regimes 3 and 4 imply higher volatilities and (at least for a majority of

[22] For instance, the annualized unconditional mean excess bond return is 1.08% $<$ 1.68% observed in sample; the annualized unconditional inflation rate is 5.88% $>>$ 2.62% observed in sample, more than the double.

[23] Notice that while in a VAR(1), currently high growth forecasts future high inflation, this is not the case under the four-state model. However in both cases currently high inflation forecasts lower future growth, an inverted type of Phillips curve. The effect may have some economic relevance: a one-standard deviation increase in inflation predicts a 0.50% decline in monthly growth in the former case, and 0.49% in the latter.

pairs) larger correlations in absolute value. In fact, Tables 5, 6 and Figure 6 help us giving some economic interpretation to the four regimes. Regime 1 is a bull/rebound state characterized (see unconditional means in Table 6) by high equity risk premia (14.5% on annualized basis), low or negative real short term interest rates, relatively high inflation (4.6% on annual basis), and high dividend yield. In this regime, all variables display moderate volatility, e.g. 13% for excess equity returns, 2.4% for excess bond returns, and 2.5% for inflation. This is a rebound state because its persistence is moderate (approximately 10 months) and it tends to follow bear regimes: the estimated transition matrix in Table 5 shows that starting from a bear/recession regime, 17% of the time the system accesses a rebound (76% it stays in a bear regime). As a result, the mean dividend yield tends to be exceptionally high (5.2% vs. a historical mean of 3.8%), indication of the existence of good bargains in the stock market. The exceptional stock market performance tends to be disjoint from real growth, which has actually an unconditional mean of only 0.84% per annum. Consistently, the yield curve is relative flat (the annualized term premium is 1.9%). Historically this regime coincides with the stock market bubble of 1927-1929, the Great-Depression rebound of 1934-1937, and most of the WWII and immediate post-war years (the 'atomic' age). After one spike in the mid-1950s, the occurrences of this regime have been rather episodic, although some late periods in the tech bubble of 1999-2000 are captured by this state. Interestingly, in this regime the correlation between shocks to inflation and to real-short term rates is not statistically different from -1, i.e. inflation shocks are transmitted one-to-one to real interest rates.

Regime 2 is a stable (low volatility) regime characterized by good real growth (2.9% per year) and moderate inflation (3.2%). This a persistent regime (15 months on average) in which also equity risk premia are fairly high (5.4%), although equity prices correspond to much higher multiples than in regime one (the dividend yield has unconditional mean of 2.8%, below the historical sample mean). As experienced in the 1990s, real short term rates are low and default-free credit cheap, just in excess of 1% per year. In fact, regime 2 captures most of the booming years between the mid 1950s and 1974 (the interruptions simply correspond to officially dated NBER recessions, picked up by state 4). After picking up a portion of the 1980s (but with frequent switches in and out of regime 3), the 1990s are entirely captured by regime 2 as well as − which seems important for current policy perspectives − the more recent, 2002-2004 period. Interestingly − despite the huge debate on the so-called 'New Economy' during the late 1990s, the experience of that decade does not appear completely different from the one from other periods of sustained growth (possibly spurred by substantial productivity gains) and moderate inflation, like the 1960s.[24]

Regime 3 describes periods of intense real growth, essentially the initial stages of the business cycle when the economy emerges from a trough. In fact, regime 3 is scarcely persistent (5 months on average). In this state the picture is dominated by high real growth (almost 15% per year, although this figure must be taken with caution, since the duration is less than 6 months), good equity premia (5.9%), and a clearly upward sloping yield curve. Figure 6 shows that − among other periods − the early 1980s and 1990s are captured by regime 3. Finally, regime 4 represents a classical bear/recession state, in which risk premia are small, dividend yields relatively high (as stock prices decline), and inflation and real growth are both negative. Consistently with this interpretation, the default spread is high in this state (25 basis points vs. a historical mean of 9 points only), while its duration is moderate (4 months), coherently with the fact that recession and bear markets are generally quite short-lived. This state also implies high volatility: e.g. excess stock returns

[24] Figure 6 shows that most of the 1990s (89%) are characterized as the low-volatility regime 2. This is fully consistent with the now widespread notion of an age of "great moderation" started in the early 1990s, see Stock and Watson (2002).

have an annual volatility of 41%, excess bond returns of 7%; even IP real growth is relatively unstable in this regime, 11% per year. Figure 6 shows that regime 4 picks up all major US recessions after WWII, in addition to a long period that matches the so-called Great Depression.[25]

Table 6 also shows that the correspondence between sub-period means $E_{\tau_0 \to \tau_1}[\mathbf{y}_{t+1}]$ under the MSIH(4,0)-VAR(1) model and the data is remarkable and obviously much superior to the simple VAR(1). For instance, the four-state model recognizes that the highest historical excess returns in the US were produced in the 1920s and 30s and then again during the 1990s. Even when the values of $E_{\tau_0 \to \tau_1}[\mathbf{y}_{t+1}]$ depart from sub-sample means, the four state model always gets the ranking of sub-sample right, while the signs are also often correct. This is another indication that − albeit the structure of predictability is constant over time − the presence of time-variation in some of the parameters gives an essential contribution at tracking and predicting the variables at hand. Figure 7 reinforces this conclusion by plotting the in-sample fitted values for the eight series under examination. On the top of each panel, correlations between fitted and observed values are reported. First, we notice that in many instances, the MSIH(4,0)-VAR(1) values are simply much more volatile (hence able to track the underlying series) than in the VAR(1) case, which is to be expected. A careful eye may even detect the existence of periods (e.g. the 1990s) over which the behavior of the fit values across the two models is clearly heterogeneous. Second, the four-state model clearly does a superior job at matching the dynamics of most of the series: for six out of eight, the correlation between actual values and fitted ones is higher under a MSIH(4,0)-VAR(1) than under a VAR(1).[26] In some occasions the distance appears important: e.g. such correlations are $0.61 > 0.50$ for real T-bill yields, $0.65 > 0.53$ for CPI inflation, and $0.30 > 0.22$ for real monetary base growth.

6. **Forecasting Performance**

Ultimately, what matters of a model is not (or not mainly) its ability to produce an accurate in-sample fit, but especially its out-of-sample forecasting performance. In fact, when the models are flexible enough thanks to the presence of a high number of parameters, accuracy of fit is relatively unsurprising. However, rich parameterizations are also well known to introduce large amounts of estimation uncertainty which normally end up deteriorating the out-of-sample performance. In the predictability literature this has been stressed among the others by Chan, Karceski, and Lakonishok (1998) who − studying the out-of-sample predictability of stock returns − found that except for the term and default spreads, macroeconomic variables tend to perform poorly. Goyal and Welch (2003) report that the dividend yield is a good predictor of stock returns only when the forecasting horizon is longer than 5 to 10 years. Otherwise the fit is good only in-sample. Neely and Weller (2000) re-examine the findings of Bekaert and Hodrick (1992) using a predictive metric. They show that VAR models are outperformed by a simple benchmark in which expected excess returns in stock and exchange markets are assumed to be constant and dividend yields and forward premiums are assumed to follow random walks. They suggest that the poor forecasting performance is due to underlying structural changes, which points to some role for regime switching models.

In fact, Bossaerts and Hillion (1999) argue that even the best linear models contain no out-of-sample forecasting power even when the specification of the models is based on statistical criteria that should penalize over-fitting. They speculate that the parameters of the selected models may be changing over time so that

[25]We notice that the last two months of 2004 were classified as a recession by our algorithm.

[26]The exceptions are real IP growth rate and the dividend yield (a tie). However in the latter case the tie is reached at a correlation of 0.98, i.e. both models do an excellent job and no spaces for improvements are left.

the correct model ought to be nonlinear, possibly of a regime switching type. Recent papers (e.g. Guidolin and Timmermann (2005c) for excess stock and bond returns or Guidolin and Timmermann (2005d) for short-term interest rates) have found that regime switching models may prove extremely useful to forecast over intermediate frequencies, such as monthly data. One wonders if a similar result holds for the larger vector under investigation in this paper and when financial and macroeconomic variables are jointly modeled.

To assess whether a four-state model offers any useful prediction performance, we implement the following 'pseudo out-of-sample' recursive strategy. For both the VAR(1) and the MSIH(4,0)-VAR(1) models, we obtain recursive estimates over expanding samples starting with 1926:12 - 1985:01, 1926:12 - 1985:02, etc. up to 1926:12 - 2004:11. This gives a sequence of 239 sets of parameter estimates specific to each of the models. For instance, the regime switching model (3) generates 239 sets of regime-specific intercepts, covariance matrices, and transition matrices, as well as of regime-independent VAR(1) coefficients. At each final date in the expanding sample − i.e. on 1985:01, 1985:02, etc. up to 2004:11 − we calculate 1-month ahead forecasts for each of the 8 variables under study, i.e. including both financial and macroeconomic variables.[27] We call $\hat{y}_t^{(M,h)}$ the forecast generated by model M for variable h. Finally, we evaluate the accuracy of the resulting forecasts, by calculating the resulting forecast errors defined as $e_t^{(M,h)} \equiv y_{t+1}^h - \hat{y}_t^{(M,h)}$.

Figure 8 starts by reporting a few recursive statistics showing that even over the expanding sequence of samples 1926:12 - 1985:01, 1926:12 - 1985:02, etc. up to 1926:12 - 2004:12, there is full justification for using a four-state model instead of a simpler VAR(1). The upper plot in the figure shows the LR tests. Although Section 4.1 has explained the existence of nuisance parameter problems, it seems evident that in the fact of LR tests always exceeding 4,600 significance with p-values below 0.01 can always be established. Equivalently, the bottom panel shows the resulting H-Q information criteria under $k = 1$ and 4. It seems that recursively estimating (3) to generate forecasts does not represent a violation of the properties of the data.

Table 7 reports summary statistics concerning the quality of the relative forecasting performance. In particular, we report three statistics illustrating predictive accuracy: the root mean-squared forecast error (RMSFE),

$$RMSFE^{(M,h)} \equiv \sqrt{\frac{1}{239}\sum_{t=1985:01}^{2003:11}\left(y_{t+1}^h - \hat{y}_{t+1}^{(M,h)}\right)^2},$$

the prediction bias

$$Bias^{(M,h)} \equiv \frac{1}{239}\sum_{t=1985:01}^{2003:11}\left(y_{t+1}^h - \hat{y}_{t+1}^{(M,h)}\right),$$

and the variance (or standard deviation):

$$SD^{(M,h)} \equiv \sqrt{\frac{1}{239}\sum_{t=1985:01}^{2003:11}\left[(y_{t+1}^h - \hat{y}_{t+1}^{(M,h)}) - \frac{1}{239}\sum_{t=1985:01}^{2003:11}\left(y_{t+1}^h - \hat{y}_{t+1}^{(M,h)}\right)\right]^2}.$$

Notice that the three statistics are not independent as it is well known that

$$MSFE^{(M,h)} \equiv \left[Bias^{(M,h)}\right]^2 + \left[SD^{(M,h)}\right]^2,$$

[27] We recognize that the need to accurately predict a few of the variables is probably more pressing (e.g. excess stock returns and inflation rates) than for others (e.g. the dividend yield or the growth rate of the real monetary base). However it would seem overly arbitrary to limit ourselves to a few of the variables only, after having discussed the full-sample evidence with reference to them all.

i.e. the MSFE can be decomposed in the contribution of bias and variance of the forecast errors.

Importantly, we try to frame the exercise in a natural metric in the following sense: while it would be straightforward for us to calculate and evaluate predictive accuracy for the eight variables modeled in this paper, we focus instead on five variables which depart from those in $\mathbf{y}_{t+1}$: value-weighted stock and bond returns (excess stock or bond returns + real T-bills + inflation), T-bill yields (real T-bill yields + inflation), IP nominal growth (real IP growth + inflation), and nominal monetary base growth (real monetary base growth + inflation). It is in fact clear that many decision makers may be interested in nominal asset returns.

The four-state model is generally superior to a simple VAR(1), which is reassuring of the need of both introducing multiple states and − on the other hand − of leaving the VAR coefficients as time homogeneous. $RMSFE^{(M,h)}$ strongly declines when M goes from the VAR(1) to the MSIH(4,0)-VAR(1) for at least 3 variables: excess stock and bond returns, and the dividend yield. Less substantial gains are manifest for real monetary base growth. In only two cases there is a visible loss of performance from adopting $k = 4$, the case of real T-bill rates and of the default spread. For financial variables, the improvement in forecasting accuracy is mostly caused by a reduction in the variance of forecast errors. This is what one expects if the model's regimes accurately identify in real time the economy's turning points. On the contrary, where the performance gets worse, the cause seems to be traceable to an increase in variance. We speculate that a four-state model may then have difficulties at identifying in real time the switches in short-term interest rates and of corporate bonds of different ratings. For CPI inflation and real IP growth, the overall performance is nearly identical, but the composition of the RMSFE is different for the two models, in the sense that a four-state model implies lower variance but higher bias in the forecasts.

7. Asset Pricing and Portfolio Implications

Although the general issue of whether the dynamic linkages between financial markets and macroeconomic factors have changed over time has been decided on already, there is a class of economic agents that has a straightforward use for the model in Table 5: portfolio managers. In fact, while financial economists have been worrying about the implications of the predictability findings for issues of market efficiency and the theoretical properties of equilibrium asset prices, money managers have attempted to exploit statistical predictability patterns − including the reaction of asset returns to macroeconomic announcements − to improve the mean-risk properties of their portfolios. In this sense, such decision makers would be mostly interested not in point forecasts of future asset returns (and possibly a few of the macroeconomic aggregates, such as inflation and real growth) or in the ability of (3) to approximate their conditional joint density, but mostly in correctly forecasting the Sharpe ratios,

$$SR_{t,t+1}^{(M,h)} \equiv \frac{E_t^{(M)}[y_{t+1}^h]}{\sqrt{Var_t^{(M)}[y_{t+1}^h]}},$$

where $h = 1$ refers to excess stock returns and $h = 2$ to excess bond returns. Sharpe ratios are the standard measure of the compensation per unit of risk used in the financial sector. Additionally, it is well known (see e.g. Ingersoll (1987)) that in simple mean-variance asset allocation frameworks, the optimal weight to be assigned to some asset h will be:

$$\omega_{t,t+1}^{(M,h)} \equiv \frac{SR_{t,t+1}^{(M,h)}}{\gamma},$$

where γ is an agent-specific risk aversion parameter.

Figure 9 shows 1-month ahead, predicted Sharpe ratios for both stocks and bonds. Such predictions are calculated under both the VAR(1) and the MSIH(4,0)-VAR(1) models and by recursively estimating the two models according to the same expanding-window format employed in Section 7. Once more, the VAR(1) model: (i) generates too flat predictions which are hardly compatible with active portfolio management; (ii) misses the specificity of the 'tech bubble' of the late 1990s. In fact, a simple linear model generates roughly stable $SR_{t,t+1}^{(M,h)}$s between 0.25 and 0.30 for stocks and between 0 and 0.1 for bonds. Such values would have implied roughly constant portfolio weights over time (probably tilted towards stocks), and quite unreasonable investment policies, with large and possibly increasing weight assigned to stocks throughout the 1990s, even at the peak of the bubble.

On the contrary, the four-state model gives reasonable portfolio advice. The equity Sharpe ratio strongly fluctuates over time in a counter-cyclical manner, i.e. $SR_{t,t+1}^{(M,1)}$ is high during recessions and declines during economic booms. Therefore it suggests large commitment to equities in 1985, in the early 1990s, and recently during 2000-2001, when $SR_{t,t+1}^{(M,1)}$ exceeds 0.2 and achieves peaks of 0.3. Importantly, between 1998 and 2000, $SR_{t,t+1}^{(M,1)}$ strongly signals the presence of a bubble, i.e. of a modest compensation for further risks, as the ratio declines below 0.1 and touches 0 in a few months. Under regime switching, $SR_{t,t+1}^{(M,2)}$ is more stable for bonds, but still provides strong and possibly useful signals, as $SR_{t,t+1}^{(M,2)}$ becomes volatile and often exceeds 0.2 at several points between 1998 and 2002, when a rational portfolio might have included much more bonds than stocks. Of course, a much more structured approach to the portfolio effects of time-varying predictability under regime switching would be in order. Ang and Bekaert (2001) and Guidolin and Timmermann (2004, 2005a) are first steps in this direction, although a full integration (possibly in an APT framework inspired by Chen, Roll, and Ross (1986)) of the signals given by macroeconomic factors in portfolio framework awaits further investigation.

8. **Conclusion**

This paper has proposed to use multivariate regime switching models to study the possibility that the predictability patterns involving US asset returns and macroeconomic variables be time-varying. Using a long monthly data set (1926-2004) we find overwhelming evidence of regimes (i.e. of structural breaks in the joint process for returns and macroeconomic factors), although the null of a stable set of dynamic predictability relationships cannot be rejected. In this sense, famous historical experiences concerning the linkages (or absence thereof) between financial markets and the real economy − for instance the Great Depression and the tech bubble of the 1990s − as not as heterogeneous as commonly thought. The good performance of our four-state model at fitting the entire density of the data and its useful forecasting performance stress that payoffs may exist in explicitly modeling the presence of regimes, although it is clear that when switches in intercepts and covariance matrices are accounted for, no need is left for explicitly time-varying predictability patterns.

Several extensions of this paper could be attempted. First of all, although we have tried to focus on a selected number of macroeconomic predictors that had performed well so far at forecasting asset returns, the system might be either expanded to include other variables, to test the robustness of our conclusions. Second, in this paper ours has been a theory-free approach that only focuses on the statistical aspects of predictability and on their possible implications for financial decisions. However, papers like Ang and Piazzesi (2003) and Wickens and Flavin (2001) have recently shown how predictability involving macroeconomic factors, no-arbitrage asset pricing, and optimal asset allocation can be brought together by imposing appropriate restrictions. Finally, there is of course nothing magic about using multivariate regime switching models to study time-varying linkages between financial markets and the macroeconomy. Other modeling approaches could be useful. For instance, Bredin and Hyde (2005) have recently applied smooth transition regression methods to study the nonlinear relationship between eight international stock returns and a few macroeconomic variables.

References

[1] Ang A., and G., Bekaert, 2002, "International Asset Allocation with Regime Shifts", *Review of Financial Studies,* 15, 1137-1187.

[2] Ang, A., and G., Bekaert, 2002, "Regime Switches in Interest Rates", *Journal of Business and Economic Statistics*, 20, 163-182.

[3] Ang, A., and M., Piazzesi, 2003, "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables", *Journal of Monetary Economics*, 50, 745-787.

[4] Banz, R., 1982, "The Relationship Between Return and Market Value of Common Stocks", *Journal of Financial Economics*, 9, 3-18.

[5] Balvers, R., T., Cosimano, and B., Mcdonald, 1990, "Predicting Stock Returns in an Efficient Market", *Journal of Finance*, 45, 1109-1128.

[6] Basu, S., 1983, "The Relationship Between Earnings Yield, Market Value, and Return for NYSE Common Stocks: Further Evidence", *Journal of Financial Economics*, 12, 129-156.

[7] Baum, L., T., Petrie, G., Soules, and N., Weiss, 1970, "A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains", *Annals of Mathematical Statistics*, 41, 164-171.

[8] Bekaert, G., and R., Hodrick, 1992, "Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets", *Journal of Finance*, 47, 467-509.

[9] Berkowitz, J., 2001, "Testing Density Forecasts with Applications to Risk Management", *Journal of Business and Economic Statistics*, 19, 465-474.

[10] Bernanke, B., 1983, "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression", *American Economic Review*, 73, 257-276.

[11] Bernanke, B., and A., Blinder, 1992, "The Federal Funds Rate and the Channels of Monetary Transmission", *American Economic Review*, 82, 901-921.

[12] Bossaerts, P., and P., Hillion, 1999, "Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?", *Review of Financial Studies*, 12, 405-428.

[13] Brandt, M., 1999, "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach", *Journal of Finance*, 54, 1609-1645.

[14] Bredin, D., and S., Hyde, 2005, "Regime Changes in the Relationship Between Stock Returns and the Macroeconomy", mimeo.

[15] Burmeister, E., and M., McElroy, 1988, "Joint Estimation of Factor Sensitivities and Risk Premia for the Arbitrage Pricing Theory", *Journal of Finance*, 43, 721-733.

[16] Campbell, J. 1987, "Stock Returns and the Term Structure", *Journal of Financial Economics*, 18, 373-399.

[17] Campbell, J., A., Lo, and C., Mackinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press.

[18] Campbell, J., and L., Viceira, 1999, "Consumption and Portfolio Decisions when Expected Returns are Time Varying", *Quarterly Journal of Economics*, 114, 433-495.

[19] Canova, F., and G., De Nicolo, 2000, "Stock Returns, Term Structure, Inflation and Real Activity: An International Perspective", *Macroeconomic Dynamics*, 4, 343-372.

[20] Chan, L., J., Karceski, and J., Lakonishok, 1998, "The Risk and Return from Factors", *Journal of Financial and Quantitative Analysis*, 33, 159-188.

[21] Chen, N.-F., R., Roll, and S., Ross, 1986, "Economic Forces and the Stock Market", *Journal of Business*, 59, 383-403.

[22] Cutler, D., J., Poterba, and L., Summers, 1989, "What Moves Stock Prices?", *Journal of Portfolio Management*, 15, 4-12.

[23] Davis, P., and G., Fagan, 1997, "Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in EU Countries?", *Journal of Applied Econometrics*, 12, 701-714.

[24] Davies, R., 1977, "Hypothesis Testing When a Nuisance Parameter Is Present Only Under the Alternative", *Biometrika*, 64, 247-254.

[25] Diebold, F., T., Gunther, and A., Tay, 1998, "Evaluating Density Forecasts", *International Economic Review*, 39, 863-883.

[26] Du, D., 2005, "Monetary Policy, Stock Returns and Inflation", *Journal of Economics and Business*, forthcoming.

[27] Emery, K., 1996, "The Information Content of the Paper-Bill Spread", *Journal of Economics and Business*, 48, 1-10.

[28] Estrella, A., and G., Hardouvelis, 1991, "The Term Structure as a Predictor of Real Economic Activity", *Journal of Finance*, 46, 555-576.

[29] Estrella, A., A., Rodrigues, and S., Schich, 2003, "How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States", *Review of Economic Statistics*, 85, 555-566.

[30] Fama, E., and K., French, 1988, "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics*, 19, 3-29.

[31] Fama, E., and K., French, 1989, "Business Conditions and Expected Returns on Stocks and Bonds", *Journal of Financial Economics*, 25, 23-49.

[32] Fama, E., and K., French, 1992, "The Cross-Section of Expected Stock Returns", *Journal of Finance*, 47, 427-465.

[33] Fama, E., and K., French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33, 3-36.

[34] Fama, E., and W., Schwert, 1977, "Asset Returns and Inflation", *Journal of Financial Economics*, 5, 115-146.

[35] Fenton, V., and R., Gallant, 1996, "Qualitative and Asymptotic Performance of SNP Density Estimation", *Journal of Econometrics*, 74, 77-118.

[36] Ferson, W., and C., Harvey, 1993, "The Risk and Predictability of International Equity Returns", *Review of Financial Studies*, 6, 527-566.

[37] Fischer, S., and R., Merton, 1984, "Macroeconomics and Finance: The Role of the Stock Market", *Carnegie-Rochester Conference Series on Public Policy*, 21, 57-108.

[38] Flannery, M., and A., Protopapadakis, 2002, "Macroeconomic Factors Do Influence Aggregate Stock Returns", *Review of Financial Studies*, 15, 751-782.

[39] Gallant, R., and G., Tauchen, 1997, "SNP: A Program for Nonparametric Time Series Analysis", mimeo, Duke University.

[40] Garcia, R., 1998, "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models", *International Economic Review*, 39, 763-788.

[41] Goyal, A., and I., Welch, 2003, "Predicting the Equity Premium with Dividend Ratios", *Management Science*, 49, 639-654.

[42] Gray, S., 1996, "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process", *Journal of Financial Economics*, 42, 27-62.

[43] Guidolin, M., and A., Timmermann, 2004, "Strategic Asset Allocation under Regime Switching", mimeo, Federal Reserve Bank of St. Louis and UCSD.

[44] Guidolin, M., and A., Timmermann, 2005a, "Economic Implications of Bull and Bear Regimes in UK Stock and Bond Returns", *Economic Journal*, forthcoming.

[45] Guidolin, M., and A., Timmermann, 2005b, "An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns", *Journal of Applied Econometrics*, forthcoming.

[46] Guidolin, M., and A., Timmermann, 2005c, "Term Structure of Risk Under Alternative Econometric Specifications", *Journal of Econometrics*, forthcoming.

[47] Guidolin, M., and A., Timmermann, 2005d, "Optimal Forecast Combination Weights under Regime Shifts: An Application to US Interest Rates", mimeo, Federal Reserve Bank of St. Louis and UCSD.

[48] Hamilton, J., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57, 357-384.

[49] Hamilton, J., and R., Susmel, 1994, "Autoregressive Conditional Heteroskedasticity and Changes in Regime", *Journal of Econometrics*, 64, 307-333.

[50] Hamilton, J., 1993, "Estimation, Inference, and Forecasting of Time Series Subject to Changes in Regime", in Maddala, G., Rao, C., and Vinod, H., *Handbook of Statistics*, vol. 11., Amsterdam: North Holland.

[51] Hansen, B., 1992, "The Likelihood Ratio Test Under Non-Standard Conditions: Testing the Markov Switching Model of GNP", *Journal of Applied Econometrics*, 7, S61-S82.

[52] Hardouvelis, G., 1987, "Macroeconomic Information and Stock Prices", *Journal of Economics and Business*, 39, 131-140.

[53] Homa, K., and D., Jaffee, 1971, "The Supply of Money and Common Stock Prices", *Journal of Finance*, 26, 1045-1066.

[54] Ingersoll, J.E., 1987, *Theory of Financial Decision Making*. Rowland and Littlefield.

[55] Jaditz, T., L., Riddick, and C., Sayers, 1998, "Multivariate Nonlinear Forecasting: Using Financial Information to Forecast the Real Sector", *Macroeconomic Dynamics*, 2, 369-382.

[56] James, C., S., Koreisha, and M., Partch, 1985, "A VARMA Analysis of Casual Relations among Stock Returns, Real Output, and Nominal Interest Rates", *Journal of Finance*, 40, 1375-1384.

[57] Kaul, G., 1987, "Stock Returns and Inflation: the Role of the Monetary Sector', *Journal of Financial Economics*, 18, 253-276.

[58] Kim, C.-J., 1994, "Dynamic Linear Models with Markov-Switching", *Journal of Econometrics*, 60, 1-22.

[59] Krolzig, H.-M., 1997, *Markov-Switching Vector Autoregressions*, Berlin, Springer-Verlag.

[60] Lee, B.-S., 1992, "Causal Relations among Stock Returns, Interest Rates, Real Activity, and Inflation", *Journal of Finance*, 47, 1591-1603.

[61] Leroux, B., 1992, "Maximum Likelihood Estimation for Hidden Markov Models", *Stochastic Processes and their Applications*, 40, 127-143.

[62] Marron, J., and M., Wand, 1992, "Exact Mean Integrated Squared Error", *Annals of Statistics,* 20, 712-736.

[63] Neely, C., and P., Weller, 2000, "Predictability in International Asset Returns: Reexamination", *Journal of Financial and Quantitative Analysis*, 35, 601-620.

[64] Pesaran, H., and A., Timmermann, 1995, "Predictability of Stock Returns: Robustness and Economic Significance", *Journal of Finance*, 50, 1201-1228.

[65] Plosser, C., and G., Rouwenhorst, 1994, "International Term Structures and Real Economic Growth", *Journal of Monetary Economics*, 33, 133-156.

[66] Rapach, D., 2001, "Macro Shocks and Real Stock Prices", *Journal of Economics and Business*, 53, 5-26.

[67] Rapach, D., M., Wohar, and J., Rangvid, 2005, "Macro Variables and International Stock Return Predictability", *International Journal of Forecasting*, 21, 137-166.

[68] Rosenberg, B., K., Reid, and R., Lanstein, 1985, "Persuasive Evidence of Market Inefficiency", *Journal of Portfolio Management*, 11, 9-17.

[69] Sims, C., 1985, "A Comparison of Interwar and Postwar Cycles: Monetarism Reconsidered", *American Economic Review*, 70, 250-257.

[70] Sola, M., and J., Driffill, 1994, "Testing the Term Structure of Interest Rates Using a Stationary Vector Autoregression with Regime Switching", *Journal of Economic Dynamics and Control*, 18, 601-628.

[71] Stock, J., and M., Watson, 1989, "New Indexes of Coincident and Leading Economic Indicators", *NBER Macroeconomics Annual 1989*, MIT Press.

[72] Stock, J., and M., Watson, 2002, "Has the Business Cycle Changed and Why?", *NBER Macroeconomics Annual 2002*, MIT Press.

[73] Stock, J., and M., Watson, 2003, "Forecasting Output and Inflation: the Role of Asset Prices", *Journal of Economic Literature*, 41, 788-829.

[74] Turner, C., R., Startz, and C., Nelson, 1989, "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market", *Journal of Financial Economics*, 25, 3-22.

[75] Wickens, M., and T., Flavin, 2001, "Macroeconomic Influences and Optimal Asset Allocation", mimeo, University of York.

Appendix A - The EM Algorithm

The algorithm is dividend in two logical steps, the Expectation and the Maximization steps. Start from the model written in state-space form (1). For the sake of argument, assume that all the parameters of the model in Ψ and $\mathbf{\Sigma}_M$ are known. We separately describe the expectation and maximization steps and then bring them together.

The Expectation step. It is the product of a few smart applications of Bayes' law that allow to recursively derive a sequence of filtered probability distributions and then (going backwards) a sequence of smoothed probability distributions. Starting from a prior

$$\Pr(\boldsymbol{\xi}_t|\Im_{t-1}) = \sum_{\boldsymbol{\xi}_{t-1}} \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t-1}\right)\Pr\left(\boldsymbol{\xi}_{t-1}|\Im_{t-1}\right)$$

the posterior distribution of $\boldsymbol{\xi}_t$ given $\Im_t = \{\Im_{t-1}, \mathbf{y}_t\}$ $\Pr(\boldsymbol{\xi}_t|\Im_t)$ is given by

$$\Pr(\boldsymbol{\xi}_t|\Im_t) = \frac{\Pr(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1})\Pr(\boldsymbol{\xi}_t|\Im_{t-1})}{\Pr(\mathbf{y}_t|\Im_{t-1})},$$

where $\Pr(\mathbf{y}_t|\Im_{t-1}) = \sum_{\boldsymbol{\xi}_t} \Pr(\mathbf{y}_t, \boldsymbol{\xi}_t|\Im_{t-1}) = \sum_{\boldsymbol{\xi}_t} \Pr(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1})\Pr(\boldsymbol{\xi}_t|\Im_{t-1})$ is the unconditional likelihood of the current observation given its past. For compactness it can also be expressed as

$$\boldsymbol{\eta}_t'\hat{\boldsymbol{\xi}}_{t|t-1} = \boldsymbol{\iota}_k'\left(\boldsymbol{\eta}_t \odot \hat{\boldsymbol{\xi}}_{t|t-1}\right)$$

where $\odot$ denotes the element by element (Hadamard) product and the $k \times 1$ vector $\boldsymbol{\eta}_t$ collects the possible log-likelihood values as a function of the realized state:

$$\boldsymbol{\eta}_t \equiv \begin{bmatrix} p(\mathbf{y}_t|\boldsymbol{\xi}_t = \mathbf{e}_1, \Im_{t-1}) \\ p(\mathbf{y}_t|\boldsymbol{\xi}_t = \mathbf{e}_2, \Im_{t-1}) \\ \vdots \\ p(\mathbf{y}_t|\boldsymbol{\xi}_t = \mathbf{e}_k, \Im_{t-1}) \end{bmatrix} = \begin{bmatrix} (2\pi)^{-1/2}|\mathbf{\Sigma}_1|^{-1/2}\exp\left[(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Psi}\mathbf{e}_1)\mathbf{\Sigma}_1^{-1}(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Psi}\mathbf{e}_1)\right] \\ (2\pi)^{-1/2}|\mathbf{\Sigma}_2|^{-1/2}\exp\left[(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Psi}\mathbf{e}_2)\mathbf{\Sigma}_2^{-1}(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Psi}\mathbf{e}_2)\right] \\ \vdots \\ (2\pi)^{-1/2}|\mathbf{\Sigma}_k|^{-1/2}\exp\left[(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Psi}\mathbf{e}_k)\mathbf{\Sigma}_1^{-1}(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Psi}\mathbf{e}_k)\right] \end{bmatrix}.$$

Since the *filtered* vector $\hat{\boldsymbol{\xi}}_{t|t}$ also corresponds to the discrete probability distribution of the possible states perceived on the basis of the information set $\Im_t$, we can re-write

$$\hat{\boldsymbol{\xi}}_{t|t} = \frac{\boldsymbol{\eta}_t \odot \hat{\boldsymbol{\xi}}_{t|t-1}}{\boldsymbol{\iota}'_k \left(\boldsymbol{\eta}_t \odot \hat{\boldsymbol{\xi}}_{t|t-1}\right)}. \tag{9}$$

The algorithm is completed by the transition equation that implies that

$$E_t[\boldsymbol{\xi}_{t+1}] = \hat{\boldsymbol{\xi}}_{t+1|t} = \mathbf{F}\hat{\boldsymbol{\xi}}_{t|t}. \tag{10}$$

Assuming that the initial state probability vector $\hat{\boldsymbol{\xi}}_{1|0}$ is somehow known, (9)-(10) define an iterative algorithm that allows one to generate a sequence of filtered state probability vectors $\{\hat{\boldsymbol{\xi}}_{t|t}\}_{t=1}^{T}$. [28] Notice that the filtered probabilities are the product of a limited information technique, since despite the availability of a sample of size T, each $\hat{\boldsymbol{\xi}}_{t|t}$ is filtered out of the information set $\Im_t$ only, ignoring $\{\mathbf{y}_\tau\}_{\tau=t+1}^{T}$. However, once $\{\hat{\boldsymbol{\xi}}_{t|t}\}_{t=1}^{T}$ has been calculated, Kim's (1994) smoothing algorithm is then easily implemented to recover the sequence of *smoothed* probability distributions $\{\hat{\boldsymbol{\xi}}_{t|T}\}_{t=1}^{T}$ by iterating the following algorithm backwards, starting from the filtered (and smoothed) probability distribution $\hat{\boldsymbol{\xi}}_{T|T}$ produced by (9)-(10). Observe that

$$\begin{aligned}
\hat{\boldsymbol{\xi}}_{t|T} &= \Pr\left(\boldsymbol{\xi}_t|\Im_T\right) = \sum_{\boldsymbol{\xi}_{t+1}} \Pr\left(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t+1}|\Im_T\right) \\
&= \sum_{\boldsymbol{\xi}_{t+1}} \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t+1}, \Im_T\right) \Pr\left(\boldsymbol{\xi}_{t+1}|\Im_T\right) \\
&= \sum_{\boldsymbol{\xi}_{t+1}} \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t+1}, \Im_t, \{\mathbf{y}_\tau\}_{\tau=t+1}^{T}\right) \Pr\left(\boldsymbol{\xi}_{t+1}|\Im_T\right) \\
&= \sum_{\boldsymbol{\xi}_{t+1}} \frac{\Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t+1}, \Im_t\right) \Pr\left(\{\mathbf{y}_\tau\}_{\tau=t+1}^{T}|\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t+1}, \Im_t\right)}{\Pr\left(\{\mathbf{y}_\tau\}_{\tau=t+1}^{T}|\boldsymbol{\xi}_{t+1}, \Im_t\right)} \Pr\left(\boldsymbol{\xi}_{t+1}|\Im_T\right) \\
&= \sum_{\boldsymbol{\xi}_{t+1}} \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t+1}, \Im_t\right) \Pr\left(\boldsymbol{\xi}_{t+1}|\Im_T\right) \\
&= \sum_{\boldsymbol{\xi}_{t+1}} \frac{\Pr\left(\boldsymbol{\xi}_t|\Im_t\right) \Pr\left(\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t, \Im_t\right)}{\Pr\left(\boldsymbol{\xi}_{t+1}|\Im_t\right)} \Pr\left(\boldsymbol{\xi}_{t+1}|\Im_T\right)
\end{aligned}$$

since the Markovian structure implies that $\Pr\left(\{\mathbf{y}_\tau\}_{\tau=t+1}^{T}|\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t+1}, \Im_t\right) = \Pr\left(\{\mathbf{y}_\tau\}_{\tau=t+1}^{T}|\boldsymbol{\xi}_{t+1}, \Im_t\right)$. Hence $\hat{\boldsymbol{\xi}}_{t|T}$ can be re-written as

$$\hat{\boldsymbol{\xi}}_{t|T} = \left(\mathbf{F}'\left(\hat{\boldsymbol{\xi}}_{t+1|T} \oslash \hat{\boldsymbol{\xi}}_{t+1|t}\right)\right) \odot \hat{\boldsymbol{\xi}}_{t|t}, \tag{11}$$

where $\oslash$ denotes element-by-element division and $\Pr\left(\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t, \Im_t\right)$ equals by construction the transition matrix driving the first order Markov chain and therefore $\mathbf{F}'$ in the transition equation. (11) is initialized by setting $t = T-1$ thus obtaining

$$\hat{\boldsymbol{\xi}}_{T-1|T} = \left(F'\left(\hat{\boldsymbol{\xi}}_{T|T} \oslash \hat{\boldsymbol{\xi}}_{T|T-1}\right)\right) \odot \hat{\boldsymbol{\xi}}_{T-1|T-1}$$

and so forth, proceeding backwards until $t = 1$.[29]

[28] Alternatively, $\hat{\boldsymbol{\xi}}_{1|0}$ might be assumed to correspond to the stationary unconditional probabilities such that $\bar{\boldsymbol{\xi}} = \mathbf{P}\bar{\boldsymbol{\xi}}$. This is the way in which we make our estimation routines operational.

[29] Notice that while $\hat{\boldsymbol{\xi}}_{T|T}$ and $\hat{\boldsymbol{\xi}}_{T-1|T-1}$ will be known from the application of Hamilton's smoothing algorithm, $\hat{\boldsymbol{\xi}}_{T|T-1} = \mathbf{F}\hat{\boldsymbol{\xi}}_{T-1|T-1}$.

The Maximization step. Call $\boldsymbol{\theta}$ the vector collecting the parameters appearing in the measurement equation and $\boldsymbol{\rho}$ the vector collecting the transition probabilities in $\mathbf{P}$, i.e. $\boldsymbol{\theta} = [vec(\boldsymbol{\Psi})\,|vec(\boldsymbol{\Sigma}_M)]$ and $\boldsymbol{\rho} = vec(\mathbf{P})$. Write the likelihood function as

$$L\left(\{\mathbf{y}_t\}_{t=1}^T | \{\boldsymbol{\xi}_t\}_{t=1}^T, \boldsymbol{\theta}\right) = \sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta}) \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right) \tag{12}$$

where $\Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right) = \sum_{s_0=1}^M \xi_{s_0} \prod_{t=1}^T p_{s_{t-1},s_t}$ and the first summation spans the space defined by

$$\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2 \otimes ... \otimes \boldsymbol{\xi}_T$$

for a total of k^T possible combinations. Then the parameters $[\boldsymbol{\theta}'\ \boldsymbol{\rho}']'$ can be derived by maximization of (12) subject to the natural constraints:

$$\mathbf{P}\boldsymbol{\iota}_k = \boldsymbol{\iota}_k \quad \boldsymbol{\xi}_0'\boldsymbol{\iota}_k = 1 \tag{13}$$

$$\boldsymbol{\rho} \geq \mathbf{0},\ \boldsymbol{\xi}_0 \geq \mathbf{0}, \text{ and } \boldsymbol{\Sigma}_M\mathbf{e}_j \text{ is positive definite } \forall j = 1, 2, ..., M. \tag{14}$$

At this point it is common place to assume the "nonnegativity" constraints in (14) are satisfied and to take the first-order conditions of a Lagrangian that explicitly enforces the adding-up constraints:

$$L^*\left(\{\mathbf{y}_t\}_{t=1}^T | \{\boldsymbol{\xi}_t\}_{t=1}^T, \boldsymbol{\theta}\right) = \ln\left[\sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta}) \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right)\right] - \boldsymbol{\lambda}_1'\left(\mathbf{P}\boldsymbol{\iota}_M - \boldsymbol{\iota}_k\right) - \lambda_2\left(\boldsymbol{\xi}_0'\boldsymbol{\iota}_k - 1\right). \tag{15}$$

Appendix B - First-order conditions useful in the M-step

Derivation of the logarithm of (15) with respect to $\boldsymbol{\theta}$ gives the score function:

$$\begin{aligned}
\frac{\partial L^*}{\partial \boldsymbol{\theta}'} &= \frac{1}{L} \sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \frac{\partial \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right) \\
&= \frac{1}{L} \sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \frac{\partial \ln\left[\prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta})\right]}{\partial \boldsymbol{\theta}'} \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta}) \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right) \\
&= \sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \sum_{t=1}^T \frac{\partial \ln p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \Pr\left(\boldsymbol{\xi}_t|\Im_T; \boldsymbol{\theta}, \boldsymbol{\rho}\right)
\end{aligned}$$

since from the definition of conditional probability

$$\frac{\prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta}) \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right)}{\sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta}) \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right)} = \frac{\prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1}; \boldsymbol{\theta}) \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0; \boldsymbol{\rho}\right)}{L\left(\{\mathbf{y}_t\}_{t=1}^T | \{\boldsymbol{\xi}_t\}_{t=1}^T, \boldsymbol{\theta}\right)} = \Pr\left(\boldsymbol{\xi}_t|\Im_T; \boldsymbol{\theta}, \boldsymbol{\rho}\right).$$

Therefore

$$\sum_{t=1}^T \hat{\boldsymbol{\xi}}_{t|T}(\boldsymbol{\theta}, \boldsymbol{\rho}) \frac{\partial \ln \boldsymbol{\eta}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \mathbf{0}' \tag{16}$$

provides the first set of FOCs w.r.t. $\boldsymbol{\theta}$. Notice that these conditions involve the smoothed probabilities of the state vector, $\{\hat{\boldsymbol{\xi}}_{t|T}\}_{t=1}^T$. Furthermore, these are simply smoothed probability-weighted standard FOCs (score conditions) in a general MLE problem. For VAR(p) models, GLS-type, closed form solutions are available (see Hamilton (1994)).

The FOCs w.r.t. the transition probabilities are determined as follows. Since

$$\begin{aligned}
\frac{\partial \ln L}{\partial \boldsymbol{\rho}'} &= \frac{1}{L} \sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \frac{\partial \Pr(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}'} \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1};\boldsymbol{\theta}) \\
&= \frac{1}{L} \sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \frac{\partial \ln \Pr(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}'} \prod_{t=1}^T p(\mathbf{y}_t|\boldsymbol{\xi}_t, \Im_{t-1};\boldsymbol{\theta}) \Pr(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0;\boldsymbol{\rho}) \\
&= \sum_{\{\boldsymbol{\xi}_t\}_{t=1}^T} \sum_{t=1}^T \frac{\partial \ln \Pr(\boldsymbol{\xi}_t|\boldsymbol{\xi}_0;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}'} \Pr(\boldsymbol{\xi}_t|\Im_T;\boldsymbol{\theta},\boldsymbol{\rho}),
\end{aligned}$$

for each component p_{ij} of $\boldsymbol{\rho}$ this implies:

$$\begin{aligned}
\frac{\partial \ln L}{\partial p_{ij}} &= \sum_{t=1}^T \sum_{\boldsymbol{\xi}_{t-1}=\mathbf{e}_i} \sum_{\boldsymbol{\xi}_t=\mathbf{e}_j} \frac{\partial \ln \Pr\left(\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t-1};\boldsymbol{\rho}\right)}{\partial p_{ij}} \Pr\left(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}|\Im_T;\boldsymbol{\theta},\boldsymbol{\rho}\right) \\
&= \sum_{t=1}^T \sum_{\boldsymbol{\xi}_{t-1}=\mathbf{e}_i} \sum_{\boldsymbol{\xi}_t=\mathbf{e}_j} \frac{1}{p_{ij}} I(\boldsymbol{\xi}_{t-1}=\mathbf{e}_i, \boldsymbol{\xi}_t=\mathbf{e}_j) \Pr\left(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}|\Im_T;\boldsymbol{\theta},\boldsymbol{\rho}\right) \\
&= \sum_{t=1}^T \sum_{\boldsymbol{\xi}_{t-1}=\mathbf{e}_i} \sum_{\boldsymbol{\xi}_t=\mathbf{e}_j} \frac{\Pr(\boldsymbol{\xi}_{t-1}=\mathbf{e}_i, \boldsymbol{\xi}_t=\mathbf{e}_j|\Im_T;\boldsymbol{\theta},\boldsymbol{\rho})}{p_{ij}},
\end{aligned}$$

which originates the vector expression

$$\frac{\partial \ln L}{\partial \boldsymbol{\rho}'} = \left(\sum_{t=1}^T \left(\hat{\boldsymbol{\xi}}_{t|T}^{(2)}\right)'\right) \oslash \boldsymbol{\rho}'$$

where $\hat{\boldsymbol{\xi}}_{t|T}^{(2)}$ is a k^2 vector of (smoothed) probabilities related concerning the state $\boldsymbol{\xi}_{t-1} \otimes \boldsymbol{\xi}_t$. Since the k adding-up restrictions in $\mathbf{P}\boldsymbol{\iota}_M = \boldsymbol{\iota}_M$ can equivalently be written as $(\boldsymbol{\iota}_k' \otimes \mathbf{I}_k)\boldsymbol{\rho} = \boldsymbol{\iota}_k$, it follows that the FOCs can be written as

$$\frac{\partial L^*}{\partial \boldsymbol{\rho}'} = \left(\sum_{t=1}^T \left(\hat{\boldsymbol{\xi}}_{t|T}^{(2)}\right)'\right) \oslash \boldsymbol{\rho}' - \boldsymbol{\lambda}_1'(\boldsymbol{\iota}_k' \otimes \mathbf{I}_k) = \mathbf{0}'.$$

In other words,

$$\boldsymbol{\rho} = \left(\sum_{t=1}^T \left(\hat{\boldsymbol{\xi}}_{t|T}^{(2)}\right)\right) \oslash (\boldsymbol{\iota}_k \otimes \boldsymbol{\lambda}_1),$$

implying

$$(\boldsymbol{\iota}_k' \otimes \mathbf{I}_k)\left(\sum_{t=1}^T \left(\hat{\boldsymbol{\xi}}_{t|T}^{(2)}\right)\right) \oslash (\boldsymbol{\iota}_k \otimes \boldsymbol{\lambda}_1) = \left(\sum_{t=1}^T \left(\hat{\boldsymbol{\xi}}_{t|T}\right)\right) \oslash \boldsymbol{\lambda}_1 = \boldsymbol{\iota}_k$$

so that $\boldsymbol{\lambda}_1 = \left(\sum_{t=1}^T \hat{\boldsymbol{\xi}}_{t|T}\right)$ results. Finally, we have

$$\boldsymbol{\rho} = \left(\sum_{t=1}^T \left(\hat{\boldsymbol{\xi}}_{t|T}^{(2)}\right)\right) \oslash \left(\boldsymbol{\iota}_k \otimes \left(\sum_{t=1}^T \hat{\boldsymbol{\xi}}_{t|T}\right)\right), \tag{17}$$

which is a highly nonlinear function of smoothed regime probabilities, but that can also be easily evaluated.

Table 1

Summary Statistics for Excess Stock and Bond Returns vs. Prediction Variables

The table reports a few summary statistics for monthly CRSP excess stock and (long-term government bond) return series, and a few macroeconomic variables employed as predictors of excess asset returns. Excess returns are calculated by difference with 30-day T-bill yields. The sample period is 1926:12 – 2004:12. In the case of equities, the CRSP universe spans stocks listed on the NYSE, the NASDAQ, and the AMEX. Data on bond returns refer to the CRSP 10-Year Treasury benchmark. All returns are expressed in monthly percentage terms. LB(j) denotes the j-th order Ljung-Box statistic.

Series	Mean	Median	St. Dev.	Skewness	Kurtosis	Jarque-Bera	LB(4)	LB(4)-squares
	Excess Asset Returns (Risk Premia)							
Value-weighted excess stock returns	0.0065	0.0099	0.0550	0.2133	10.6124	2269**	21.716**	166.87**
Excess bond returns (term premium)	0.0014	0.0014	0.0188	0.2447	5.5932	271.9**	5.1774	176.31**
	Prediction Variables							
12-month cumulated dividend yield	0.0381	0.0363	0.0150	0.9542	5.8183	452.3**	3334**	2829**
Real 1-month T-bill yield	0.0005	0.0007	0.0054	-1.9764	21.0381	13313**	542.13**	79.833**
Default spread	0.0009	0.0007	0.0006	2.4203	11.3805	3657**	3284**	2683**
CPI inflation rate	0.0026	0.0025	0.0055	1.1840	16.7930	7647**	596.9**	82.741**
Real industrial production growth rate	0.0021	0.0023	0.0202	0.7663	13.2813	4219**	268.7**	372.7**
Real adj. monetary base growth rate	0.0004	0.0014	0.0321	1.7034	30.7269	30468**	34.722**	79.514**

* denotes 5% significance, ** significance at 1%.

Table 2

Correlation Matrix of Excess Asset Returns and Predictors

The table reports linear correlation coefficients for monthly excess asset (stocks and bonds) returns and a few common prediction variables. The sample period is 1926:12 – 2004:12. Boldfaced coefficients are significant at 5 percent.

	Value-weighted excess stock returns	Excess bond returns (term premium)	12-month cumulated dividend yield	Real 1-month T-bill yield	Default spread	CPI inflation rate	Real industrial production growth rate	Real adj. monetary base growth rate
Value-weighted excess stock returns	1	**0.1351**	**-0.1072**	-0.0226	0.0028	-0.0138	**0.1457**	0.0008
Excess bond returns (term premium)		1	0.0157	**0.0774**	**0.0990**	**-0.0801**	-0.0143	**0.0750**
12-month cumulated dividend yield			1	0.0040	**0.4926**	**-0.1161**	-0.0538	0.0272
Real 1-month T-bill yield				1	**0.2053**	**-0.8869**	**0.2792**	**0.2506**
Default spread					1	**-0.2345**	0.0341	**0.1650**
CPI inflation rate						1	**-0.3760**	**-0.3006**
Real industrial production growth rate							1	**0.0714**
Real adj. monetary base growth rate								1

Table 3

Model Selection Results

This table reports statistics used to select multivariate regime switching models of the form

$$\mathbf{y}_t = \boldsymbol{\mu}_{s_t} + \sum_{j=1}^{p} A_{js_t}\mathbf{y}_{t-j} + \Sigma_{s_t}\boldsymbol{\varepsilon}_t,$$

where $\mathbf{y}_t$ includes monthly excess stock and bond returns, as well as 6 prediction variables. The switching models are classified as MSIAH(k,p), where *I*, *A* and *H* refer to state dependence in the intercept, autoregressive terms and heteroskedasticity. *k* is the number of states and *p* is the autoregressive order. Models in the class MSI (*k*, 0) – AR(*p*) have regime switching in the intercept but not in autoregressive coefficients. The sample period is 1926:12 – 2004:12.

Model	Number of parameters	Log-likelihood	LR test for linearity	AIC	Hannan-Quinn	BIC	LR- test
Base model: MSIA(1,0)							
MSIA (1,0)	44	24429.52	NA	-52.0502	-51.9635	-51.8228	
MSIA (1,1)	108	29708.90	NA	-63.2498	-63.0368	-62.6911	A: 0.000
MSIA (1,2)	172	29871.33	NA	-63.5280	-63.1885	-62.6375	A: 0.000
MSIA (1,3)	236	29945.19	NA	-63.6171	-63.1508	-62.3943	A: 0.000
Base model: MSIA(2,0)							
MSIA (2,0)	54	24819.30	779.57 (0.000)	-52.8608	-52.7544	-52.5818	
MSI (2,0) – VAR(1)	118	29778.58	139.36 (0.000)	-63.3773	-63.1446	-62.7669	A: 0.000
MSIA (2,1)	182	30213.37	1008.96 (0.000)	-64.1696	-63.8107	-63.2282	A: 0.000
MSIH (2,0)	90	26472.98	4086.94 (0.000)	-56.3137	-56.1364	-55.8486	H: 0.000
MSIAH (2,1)	218	31628.38	3838.98 (0.000)	-67.1162	-66.6863	-65.9886	A: 0.000 H: 0.000
MSIH (2,0) – VAR(1)	154	31542.04	3666.29 (0.000)	-67.0685	-66.7647	-66.2719	A: 0.000 H: 0.000
MSIH (2,0) – VAR(2)	218	31654.08	3565.48 (0.000)	-67.2430	-66.8126	-66.1144	A: 0.000
MSIAH (2,2)	346	31799.46	3856.24 (0.000)	-67.2801	-66.5971	-65.4889	A: 0.000
Base model: MSIA(3,0)							
MSIA (3,0)	66	25263.64	1668.24 (0.000)	-53.7836	-53.6536	-53.4425	
MSI (3,0) – VAR(1)	130	29938.81	459.84 (0.000)	-63.6941	-63.4377	-63.0216	A: 0.000
MSIA (3,1)	258	30734.08	2050.35 (0.000)	-65.1198	-64.6110	-63.7853	A: 0.000
MSIH (3,0)	138	26952.11	5045.18 (0.000)	-57.2340	-56.9620	-56.5207	H: 0.000
MSIAH (3,1)	330	32387.91	5358.03 (0.000)	-68.9998	-68.0490	-66.7928	A: 0.000 H: 0.000
MSIH (3,0) – VAR(1)	202	32235.27	5052.74 (0.000)	-68.4472	-68.0488	-67.4023	A: 0.000 H: 0.000
Base model: MSIA(4,0)							
MSIAH (4,1)	444	32771.52	6125.24 (0.000)	**-69.0159**	-68.2002	-66.6792	
MSIH (4,0) – VAR(1)	252	32547.82	5677.83 (0.000)	-69.0081	**-68.5111**	**-67.7046**	

Table 4 – part a

Density Specification Tests for Regime Switching Models

This table reports tests for the transformed z-scores generated by multivariate regime-switching models

$$\mathbf{y}_t = \boldsymbol{\mu}_{s_t} + \sum_{j=1}^{p} \mathrm{A}_{js_t} \mathbf{y}_{t-j} + \Sigma_{s_t} \boldsymbol{\varepsilon}_t .$$

The tests are based on the principle that under the null of correct specification, the probability integral transform of the one-step-ahead standardized forecast errors should follow an IID uniform distribution over the interval (0,1). A further Gaussian transform described in Berkowitz (2001) is applied to perform Likelihood ratio tests of the null that (under correct specification) the transformed z-scores, z^*_{t+1}, are IIN(0,1) distributed. In particular, given the transformed z-score model

$$z^*_{t+1} = \alpha + \sum_{j=1}^{q} \sum_{i=1}^{l} \beta_{ij} (z^*_{t+1-i})^j + \sigma u_{t+1} ,$$

the Jarque-Bera statistic tests the hypothesis of normality, LR_2 tests the hypothesis of zero mean and unit variance under the restriction $q = l = 0$; LR_3 tests the joint hypothesis of zero mean, unit variance, and $\rho_{11} = 0$ under $q = l = 1$; LR_6 tests the joint null of zero mean, unit variance, and $\beta_{11} = \beta_{12} = \beta_{21} = \beta_{22} = 0$ with $q = l = 2$. Boldfaced statistics indicate that the null of no misspecification should be rejected.

Model	Number of parameters	Jarque-Bera test	LR_2	LR_3	LR_6
Value-weighted excess stock returns					
Linear	108	**3412.4 (0.000)**	**110.22 (0.000)**	**113.66 (0.000)**	**121.64 (0.000)**
MSIAH (3,1)	330	**1412.9 (0.000)**	**47.88 (0.000)**	**49.4 (0.000)**	**76.9 (0.000)**
MSIH(4,1) – VAR(1)	252	1.46 (0.483)	5.06 (0.080)	7.10 (0.069)	**13.52 (0.035)**
Excess 10-Year bond returns (term premium)					
Linear	108	**14.31 (0.001)**	**80.62 (0.000)**	**83.20 (0.000)**	**108.46 (0.000)**
MSIAH (3,1)	330	**90.21 (0.000)**	**61.30 (0.000)**	**77.96 (0.000)**	**93.34 (0.000)**
MSIH(4,1) – VAR(1)	252	**8.96 (0.011)**	5.30 (0.071)	7.74 (0.052)	10.20 (0.116)
12-month cumulated dividend yield					
Linear	108	**13752 (0.000)**	**202.56 (0.000)**	**204.96 (0.000)**	**260.90 (0.000)**
MSIAH (3,1)	330	**611.7 (0.000)**	**130.90 (0.000)**	**135.16 (0.000)**	**168.98 (0.000)**
MSIH(4,1) – VAR(1)	252	**7.63 (0.022)**	1.84 (0.399)	6.18 (0.103)	10.26 (0.114)
Real 1-month T-bill Yield					
Linear	108	**2674.6 (0.000)**	**96.34 (0.000)**	**98.52 (0.000)**	**110.10 (0.000)**
MSIAH (3,1)	330	**30.05 (0.000)**	5.10 (0.078)	**15.62 (0.001)**	**44.14 (0.000)**
MSIH(4,1) – VAR(1)	252	**9.90 (0.007)**	**7.22 (0.027)**	**13.16 (0.004)**	**18.52 (0.005)**

Table 4 – part b

Density Specification Tests for Regime Switching Models

Model	Number of parameters	Jarque-Bera test	LR_2	LR_3	LR_6
Default spread					
Linear	108	**46162 (0.000)**	**304.94 (0.000)**	**308.16 (0.000)**	**313.46 (0.000)**
MSIAH (3,1)	330	**2493.3 (0.000)**	**57.90 (0.000)**	**88.78 (0.000)**	**91.90 (0.000)**
MSIH(4,1) – VAR(1)	252	**19.41 (0.000)**	**9.90 (0.007)**	**30.88 (0.000)**	**56.72 (0.000)**
CPI inflation rate					
Linear	108	**21.66 (0.000)**	**95.34 (0.000)**	**102.80 (0.000)**	**118.38 (0.000)**
MSIAH (3,1)	330	**120.22 (0.000)**	**23.38 (0.000)**	**54.64 (0.000)**	**72.06 (0.000)**
MSIH(4,1) – VAR(1)	252	1.58 (0.454)	5.04 (0.080)	**9.54 (0.023)**	**13.34 (0.038)**
Real industrial production growth rate					
Linear	108	**208.11 (0.000)**	**524.32 (0.000)**	**532.12 (0.000)**	**550.96 (0.000)**
MSIAH (3,1)	330	**842.65 (0.000)**	**494.16 (0.000)**	**550.64 (0.000)**	**562.84 (0.000)**
MSIH(4,1) – VAR(1)	252	5.41 (0.067)	5.56 (0.062)	**10.54 (0.014)**	**18.34 (0.005)**
Real adjusted monetary base growth rate					
Linear	108	**953.52 (0.000)**	**297.92 (0.000)**	**327.22 (0.000)**	**343.62 (0.000)**
MSIAH (3,1)	330	**2382.3 (0.000)**	**149.68 (0.000)**	**310.82 (0.000)**	**482.84 (0.000)**
MSIH(4,1) – VAR(1)	252	3.55 (0.169)	3.68 (0.159)	**10.90 (0.012)**	**25.62 (0.000)**

Table 5 – part a

Estimates of a Four-State Switching Model with Time-Invariant VAR(1) Matrix

This table reports estimates for a single state and a four-state VAR(1) regime switching model (MSIH(4,0)-VAR(1)):

$$\mathbf{y}_t = \boldsymbol{\mu}^*_{s_t} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\Sigma}^*_{s_t}\boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t \sim$ I.I.D. $N(0, \mathbf{I}_8)$ is an unpredictable return innovation. The sample period is 1926:12 – 2004:12.

Panel A – Single State VAR(1) Model

	Stock	Bond	Div. yield	T-bill	Default	Inflation	Growth	Money
1. Intercept	-0.0039	-0.0004	0.0010***	-0.0004	0.0000	0.0006	0.0014	0.0053*
2. VAR(1) Matrix								
Stock excess returns	0.1046***	0.1108	0.2653*	-0.2514	0.5962	-0.3426	0.0653	0.1257**
Bond excess returns	-0.0260**	0.0626*	-0.0020	0.0522	2.2183*	-0.0485	0.0115	-0.0104
Dividend yield	-0.0060***	-0.0011	0.9835***	-0.0403	-0.3175*	-0.0285	-0.0029	-0.0052*
T-bill real yield	-0.0011	0.0013	-0.0249*	0.5768***	1.2793***	0.1068*	-0.0518***	-0.0027
Default spread	-0.0009***	0.0007***	0.0007**	0.0021	0.9628***	0.0002	-0.0005**	0.0000
Inflation	0.0013	-0.0087	0.0232*	0.3867***	-1.2891***	0.8632***	0.0526***	0.0026
IP real growth	0.0760***	-0.0372	-0.0854*	-0.9677***	3.3244***	-0.9506***	0.4008***	0.0396**
Money real growth	-0.0025	0.0219	-0.2341***	-1.3385***	10.338***	-2.3610***	-0.0316	-0.1880***
3. Correlations/Volatilities								
Stock excess returns	0.0542***							
Bond excess returns	0.1377**	0.0187***						
Dividend yield	-0.8830***	-0.1368**	0.0030**					
T-bill real yield	-0.0333	0.0561	0.0404	0.0047***				
Default spread	-0.2596***	0.0712	0.3324***	0.0371	0.0001*			
Inflation	0.0221	-0.0582	-0.0313	-0.9909***	-0.0298	0.0046***		
IP real growth	0.1128**	-0.0104	-0.1012**	0.4401***	-0.1744**	-0.4434***	0.0170***	
Money real growth	-0.0034	0.0594*	-0.0027	0.2329**	0.0677*	-0.2332**	0.0259	0.0031**

Panel B – Four State Model

	Stock	Bond	Div. yield	T-bill	Default	Inflation	Growth	Money
1. Intercept								
Bull-rebound	0.0089**	0.0025***	0.0004**	-0.0023***	0.0000	0.0023**	0.0016*	-0.0099**
Stable-growth	0.0065**	0.0030***	0.0002	-0.0000	0.0000	0.0002	0.0023*	-0.0001
Expansion-peak	0.0058***	0.0101	0.0004**	0.0000	0.0000	0.0004	-0.0002	0.0005*
Bear-recession	-0.0193***	-0.0026***	0.0013***	0.0016***	0.0001*	-0.0016**	0.0048***	0.0044***
2. VAR(1) Matrix								
Stock excess returns	-0.0009	0.1057	0.0870	-1.3822***	4.7186***	-2.1260***	-0.0433**	0.0821*
Bond excess returns	-0.0341*	0.0167	-0.0401*	-0.9605***	2.9146***	-1.0749***	-0.0079	0.0159
Dividend yield	-0.0003	-0.0034	0.9940***	0.0343	-0.3475**	0.0619*	0.0032	-0.0044*
T-bill real yield	0.0004	0.0045*	0.0011	0.4529***	0.3641**	0.0811	-0.0187	-0.0025
Default spread	-0.0003*	0.0003*	0.0001	0.0034*	0.9640***	0.0031**	0.0000	-0.0000
Inflation	-0.0006	-0.0092	-0.0008	0.4872***	-0.3839**	0.8599***	0.0190	0.0028*
IP real growth	0.0235*	-0.0089	-0.0385	-0.8952***	0.9575**	-0.9208***	0.3861***	0.0185
Money real growth	0.0143	0.0233	0.0394	-1.4162***	6.8721***	-2.3985***	-0.0294	-0.1786***
3. Correlations/Volatilities								
Regime 1 (Bull-rebound):								
Stock excess returns	0.0385***							
Bond excess returns	0.1494**	0.0068***						
Dividend yield	-0.8806***	-0.1687***	0.0021***					
T-bill real yield	0.0438	0.0522	-0.0286	0.0071***				
Default spread	-0.1434**	-0.0127	0.1049**	-0.0593	3.9e-05*			
Inflation	-0.0429	-0.0522	0.0270	-0.9990***	0.0589	0.0071***		
IP real growth	-0.0918**	0.0388	0.0678*	0.5054***	-0.1040*	-0.5051***	0.0256***	
Money real growth	0.1742**	0.0220	-0.1864**	0.3201***	-0.0891*	-0.3211***	0.1508**	0.0351***

Table 5 – part a

Estimates of a Four-State Switching Model with Time-Invariant VAR(1) Matrix

	Panel B – Four State Model							
	Stock	Bond	Div. yield	T-bill	Default	Inflation	Growth	Money
3. Correlations/Volatilities								
Regime 2 (Stable-growth):								
Stock excess returns	0.0365***							
Bond excess returns	0.1433**	0.0180***						
Dividend yield	-0.8999***	-0.1807***	0.0011***					
T-bill real yield	0.0784	0.0593	-0.0721	0.0023***				
Default spread	-0.0271	-0.0144	0.0204	0.0989*	3.98e-05*			
Inflation	-0.0917*	-0.0554	0.0884*	-0.9857***	-0.1034**	0.0023***		
IP real growth	0.0953*	-0.0492	-0.1158**	0.5290***	-0.0144	-0.5406***	0.0083***	
Money real growth	0.0539	0.0009	-0.0449	0.3928***	0.0472	-0.3953***	0.2029**	0.0108***
Regime 3 (Expansion-peak):								
Stock excess returns	0.0606***							
Bond excess returns	0.1965***	0.0277***						
Dividend yield	-0.9691***	-0.2080***	0.0026***					
T-bill real yield	0.0063	0.0896	-0.0427	0.0043***				
Default spread	0.0773*	0.3746***	-0.0809**	0.0494	0.0001**			
Inflation	-0.0665	-0.1141*	0.1161**	-0.9535***	-0.0045	0.0041***		
IP real growth	0.1126**	-0.0253	-0.1372**	0.5943***	-0.0540	-0.5789***	0.0125***	
Money real growth	-0.0977**	0.0234	0.0702*	0.3791***	-0.0690	-0.4093***	0.2576**	0.0198***
Regime 4 (Bear-recession):								
Stock excess returns	0.1196***							
Bond excess returns	0.1040*	0.0197***						
Dividend yield	-0.9225***	-0.1873***	0.0082***					
T-bill real yield	-0.1698**	0.1577**	0.1597**	0.0067***				
Default spread	-0.4836***	-0.1435**	0.4864***	0.0744	0.0003***			
Inflation	0.1727**	-0.1605**	-0.1631**	-0.9981***	-0.0797*	0.0067***		
IP real growth	0.3401***	0.0683	-0.2432***	0.1544**	-0.4347***	-0.1601**	0.0317***	
Money real growth	-0.0456	0.2287***	0.0456	0.0262	0.1042*	-0.0180	-0.1768**	0.0769***

4. Transition probabilities	Bull-rebound	Stable-peak	Expansion	Bear-recession
Bull-rebound	0.8939***	0.0237*	0.0207	0.0616
Stable-growth	0.0100	0.9311***	0.0523**	0.0066
Expansion-peak	0.0230	0.1517***	0.8030***	0.0381
Bear-recession	0.1683***	0.0340*	0.0381**	0.7596

* denotes 10% significance, ** significance at 5%, *** significance at 1%.

Table 6

Implied Monthly Means from a Four-State Switching Model

This table reports estimates for a single state and a four-state VAR(1) regime switching model (MSIH(4,0)-VAR(1)):

$$\mathbf{y}_t = \boldsymbol{\mu}^*_{s_t} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\Sigma}^*_{s_t}\,\boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t \sim$ I.I.D. $N(0, \mathbf{I}_8)$ is an unpredictable return innovation. The sample period is 1926:12 – 2004:12.

	Stock	Bond	Div. yield	T-bill	Default	Inflation	Growth	Money
	Panel A – Data							
Overall mean	0.0065	0.0014	0.0381	0.0005	0.0009	0.0025	0.0021	0.0004
Mean 1926-1946	0.0072	0.0027	0.0493	0.0001	0.0015	0.0006	0.0018	0.0064
Mean 1947-1966	0.0090	-0.0001	0.0421	0.0000	0.0005	0.0018	0.0005	0.0012
Mean 1967-1986	0.0032	0.0008	0.0372	0.0009	0.0010	0.0051	-0.0081	-0.0040
Mean 1987-2004	0.0065	0.0026	0.0223	0.0012	0.0007	0.0025	0.0026	0.0025
	Panel B – Single state VAR(1) Model							
Overall mean	0.0065	0.0014	0.0381	0.0005	0.0009	0.0025	0.0021	0.0004
Unconditional mean	0.0057	0.0009	0.0397	0.0004	0.0007	0.0034	0.0049	0.0040
Mean 1926-1946	0.0118	0.0027	0.0489	0.0003	0.0015	0.0005	0.0028	0.0062
Mean 1947-1966	0.0078	0.0003	0.0422	-0.0006	0.0005	0.0025	0.0014	-0.0034
Mean 1967-1986	0.0040	0.0015	0.0371	0.0015	0.0010	0.0044	-0.0073	-0.0049
Mean 1987-2004	0.0018	0.0011	0.0226	0.0011	0.0007	0.0026	-0.0014	0.0064
	Panel C – Four State Model							
Overall mean*	0.0063	0.0014	0.0381	0.0005	0.0009	0.0025	0.0021	0.0004
Mean 1926-1946*	0.0102	0.0029	0.0492	-0.0002	0.0015	0.0009	0.0021	0.0051
Mean 1947-1966*	0.0090	0.0006	0.0421	-0.0003	0.0005	0.0022	0.0004	0.0018
Mean 1967-1986*	0.0018	0.0015	0.0373	0.0013	0.0010	0.0045	-0.0079	-0.0038
Mean 1987-2004*	0.0059	0.0007	0.0224	0.0011	0.0007	0.0026	0.0023	0.0013
Unconditional mean	0.0061	0.0016	0.0335	0.0005	0.0009	0.0025	0.0018	0.0006
Regime 1 – unc. mean	0.0121	0.0006	0.0523	-0.0030	0.0006	0.0038	0.0007	-0.0068
Regime 2 – unc. mean	0.0045	-0.0005	0.0282	0.0009	0.0006	0.0027	0.0024	0.0023
Regime 3 – unc. mean	0.0049	0.0073	0.0251	0.0020	0.0015	0.0040	0.0122	0.0001
Regime 4 – unc. mean	0.0021	0.0052	0.0318	0.0034	0.0025	-0.0045	-0.0084	0.0242

* Based on smoothed probabilities.

Table 7

Out-of-Sample, Recursive Predictive Performance

The table reports the root-mean-square forecast error, the predictive bias, and the forecast error variance for two models, a single state and a four-state VAR(1) regime switching model (MSIH(4,0)-VAR(1)):

$$\mathbf{y}_t = \boldsymbol{\mu}^*_{s_t} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\Sigma}^*_{s_t} \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t \sim \text{I.I.D. } N(0, \mathbf{I}_8)$ is an unpredictable return innovation. The (pseudo) out-of sample period is 1985:01 – 2004:11. The models are recursively estimated on expanding windows 1926:12 – 1985:01, 1926:12 – 1985:02, up to 1926:12 – 2004:11. One-month ahead forecast errors are calculated.

	VAR(1)			**Four-state MSIH(4,0)-VAR(1)**		
	Root-MSFE	**Bias**	**St. dev.**	**Root-MSFE**	**Bias**	**St. dev.**
Value-weighted stock returns	0.0458	-0.0104	0.0446	0.0376	-0.0022	0.0376
Long-term bond returns	0.0216	-0.0001	0.0216	0.0183	-0.0010	0.0183
Dividend yield	0.0158	-0.0133	0.0086	0.0144	-0.0143	0.0018
T-bill yields (annualized)	0.0341	-0.0276	0.0201	0.0585	-0.0272	0.0518
Default spread	0.0003	-0.0002	0.0002	0.0015	0.0011	0.0011
CPI inflation	0.0025	-0.0008	0.0023	0.0028	-0.0022	0.0017
Industrial production growth	0.0058	-0.0013	0.0057	0.0069	0.0057	0.0038
Monetary base growth	0.0296	0.0028	0.0294	0.0260	0.0043	0.0257

Figure 1

Series Plots

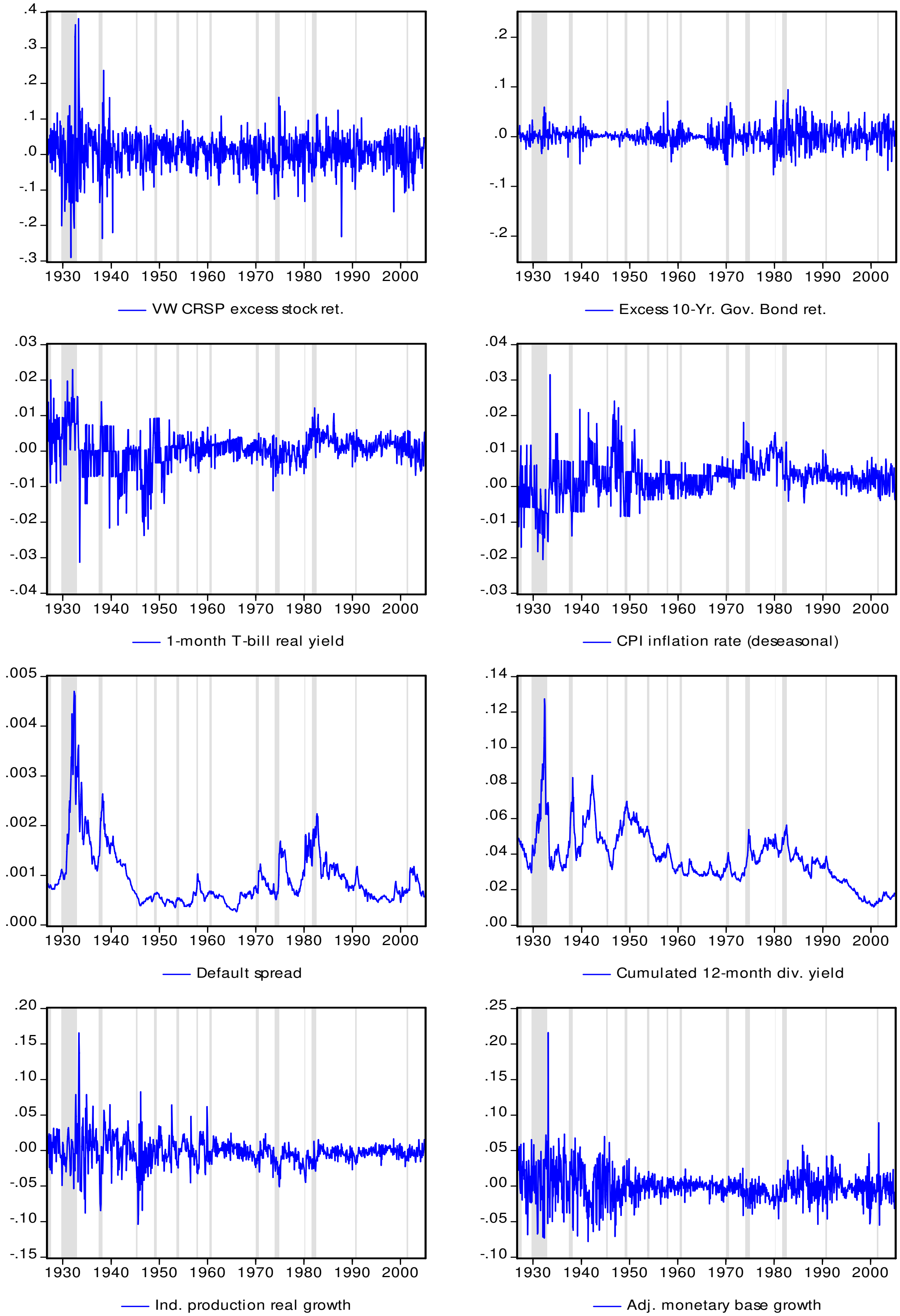

Figure 2

Distribution of Transformed (Generalized) z-Scores from Three-State VAR(1) Switching Model

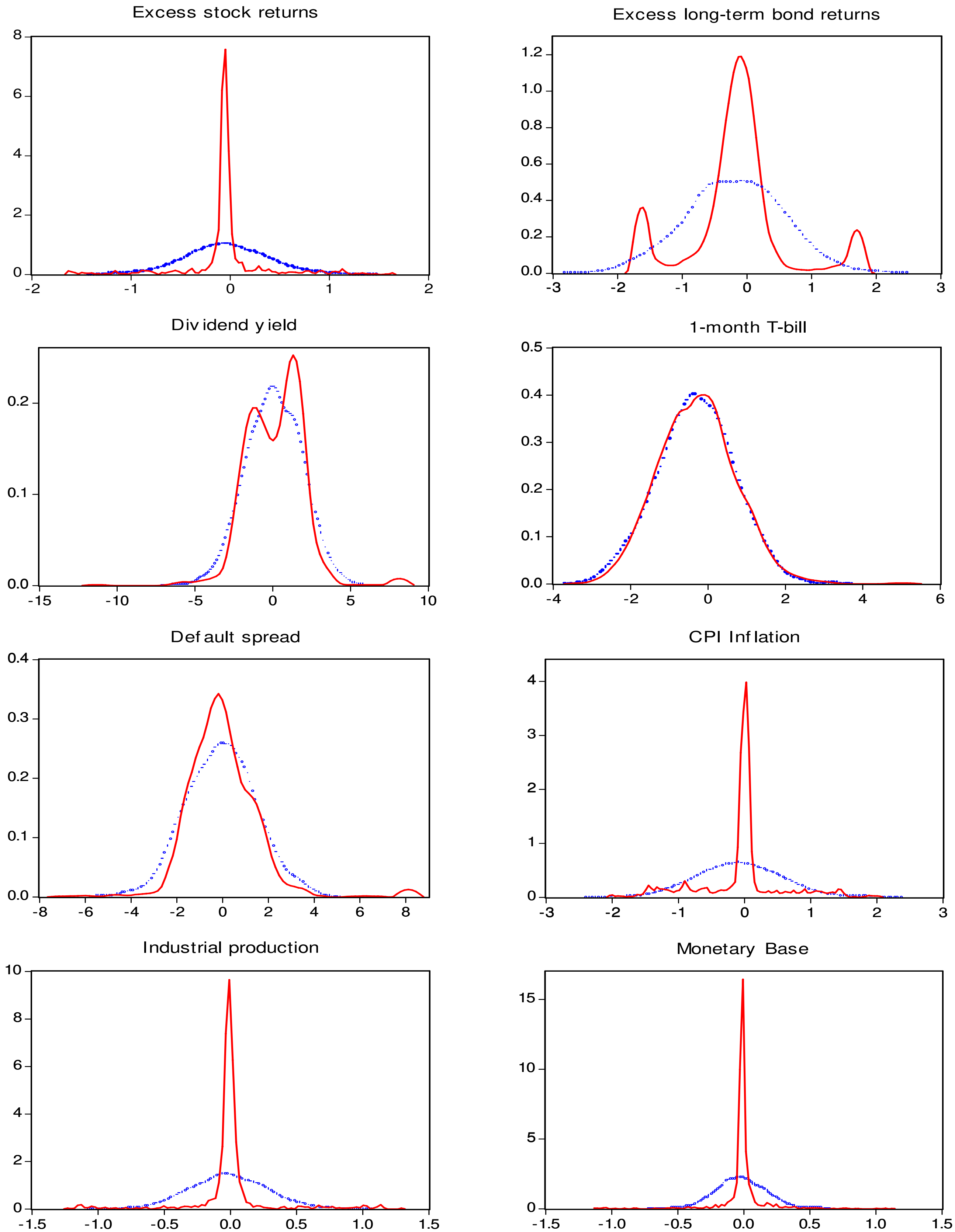

Figure 3

Quantile-Quantile Plots (vs. a Gaussian with identical mean and variance) for Transformed z-Scores from Three-State VAR(1) Switching Model

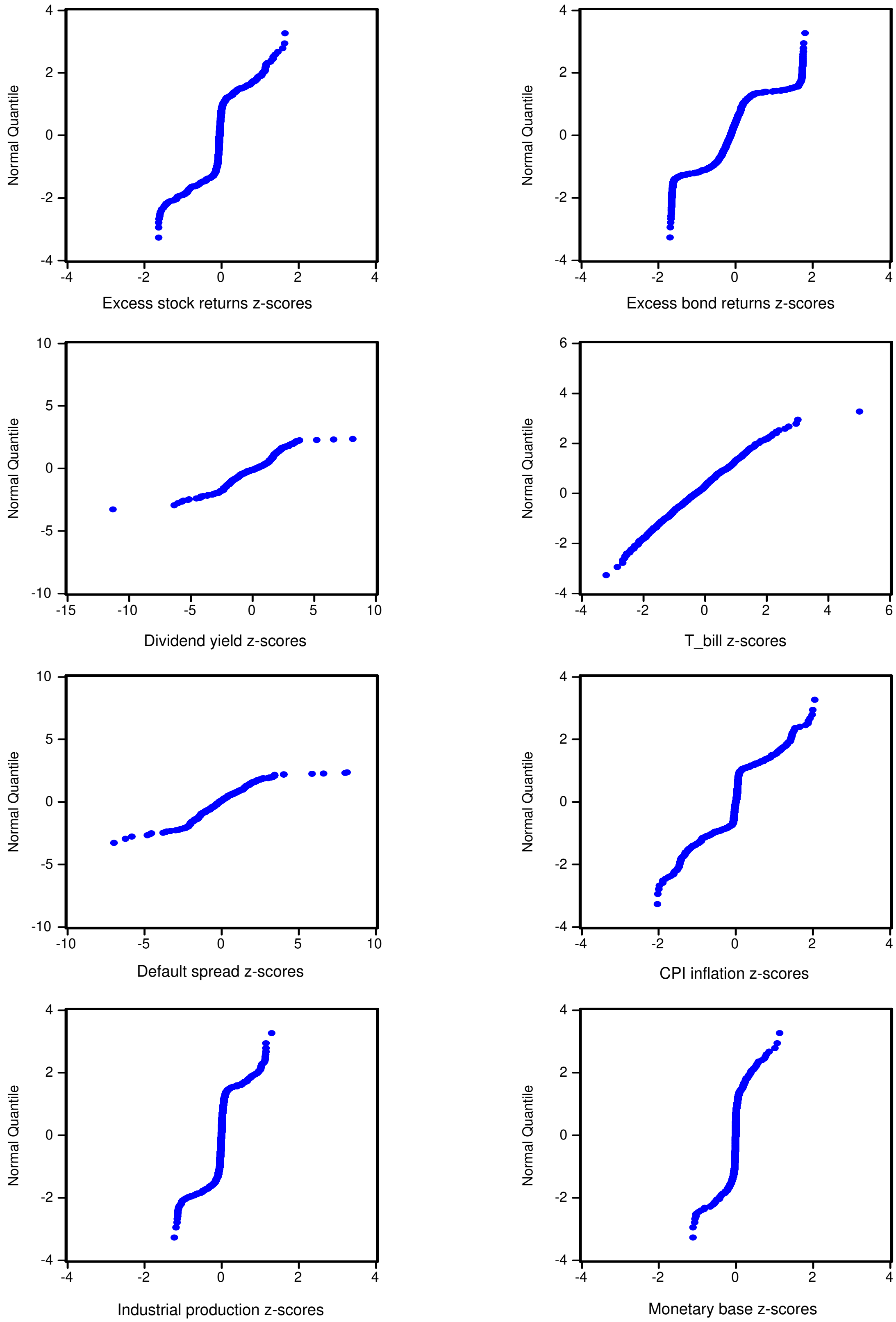

Figure 4

Distribution of Transformed (Generalized) z-Scores from Four-State Model with Time-Invariant VAR(1) Matrix

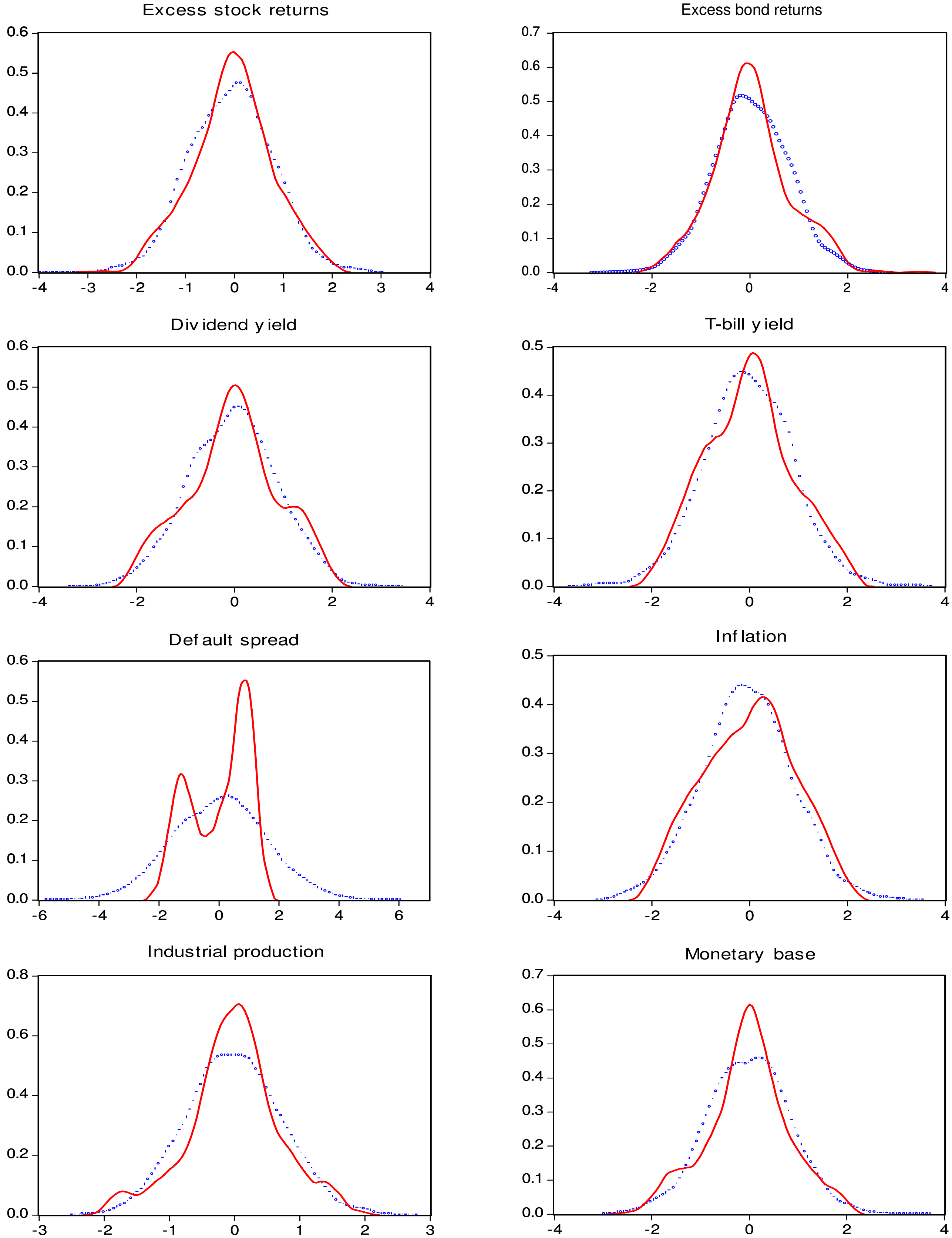

Figure 5

Quantile-Quantile Plots (vs. a Gaussian with identical mean and variance) for Transformed z-Scores from Four-State Model with Time-Invariant VAR(1) Matrix

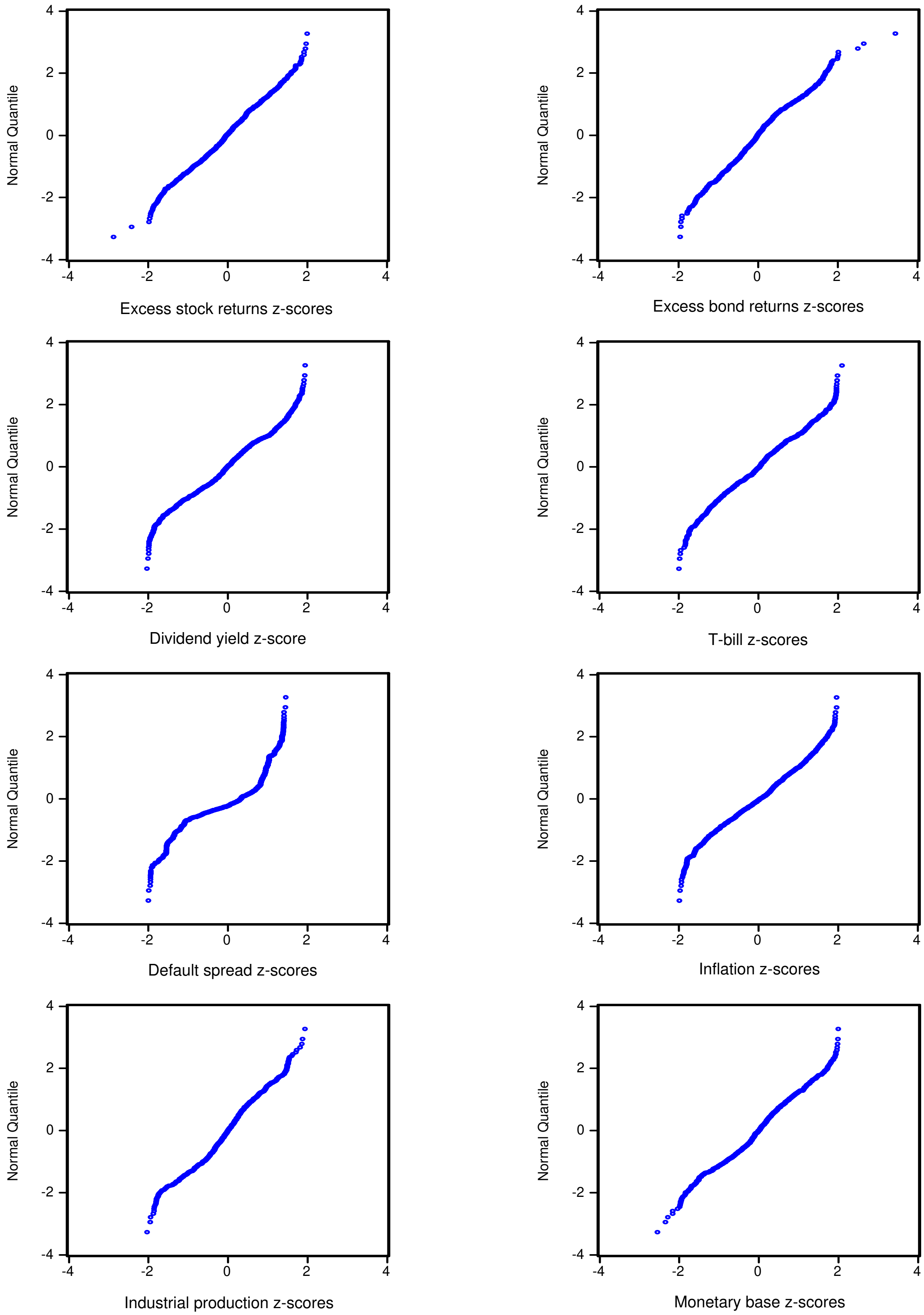

Figure 6

Smoothed Probabilities from Four-State Model

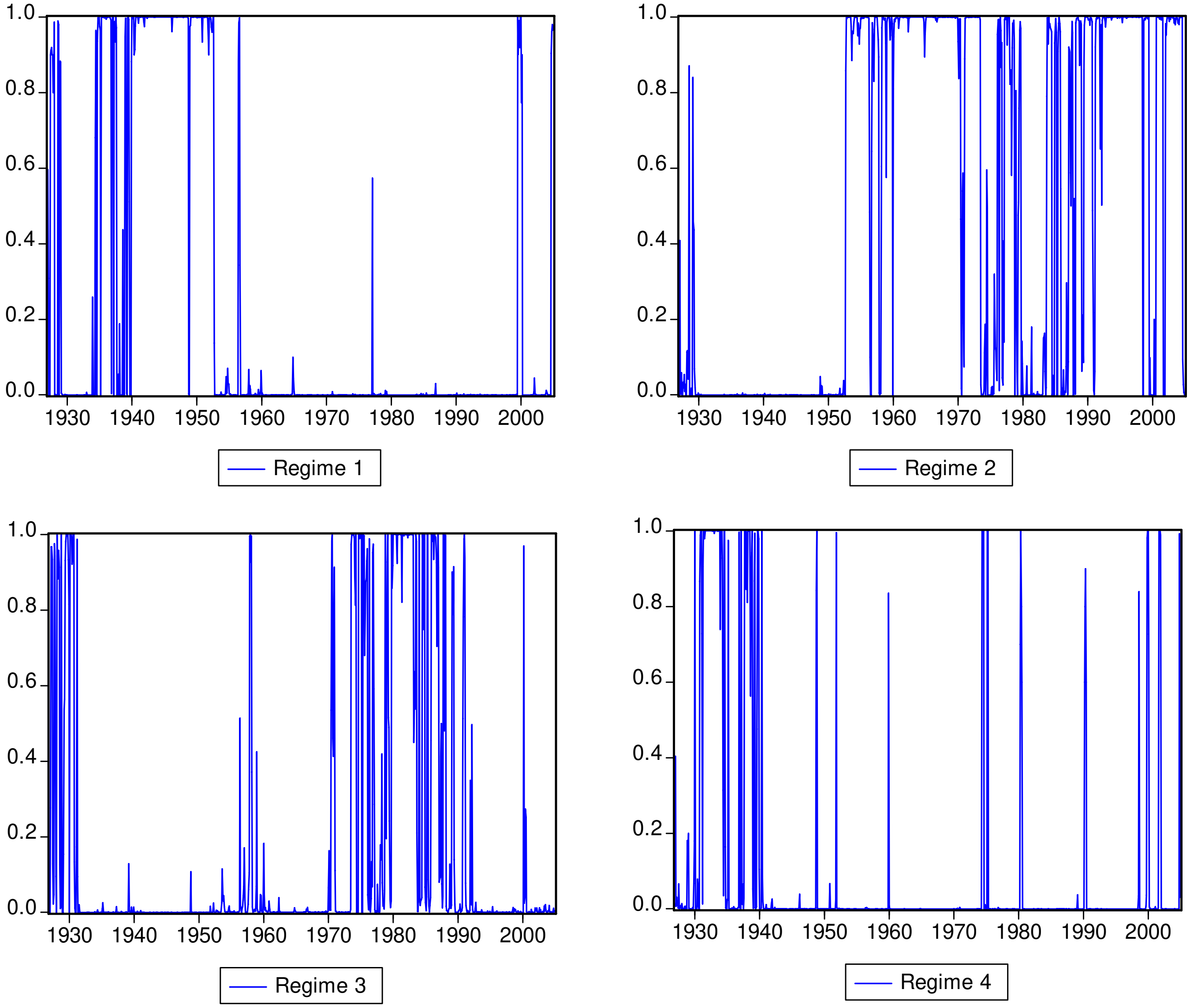

Figure 7

Comparing Fitted and Realized Values: Four-State Model with Time-Invariant VAR(1) Matrix

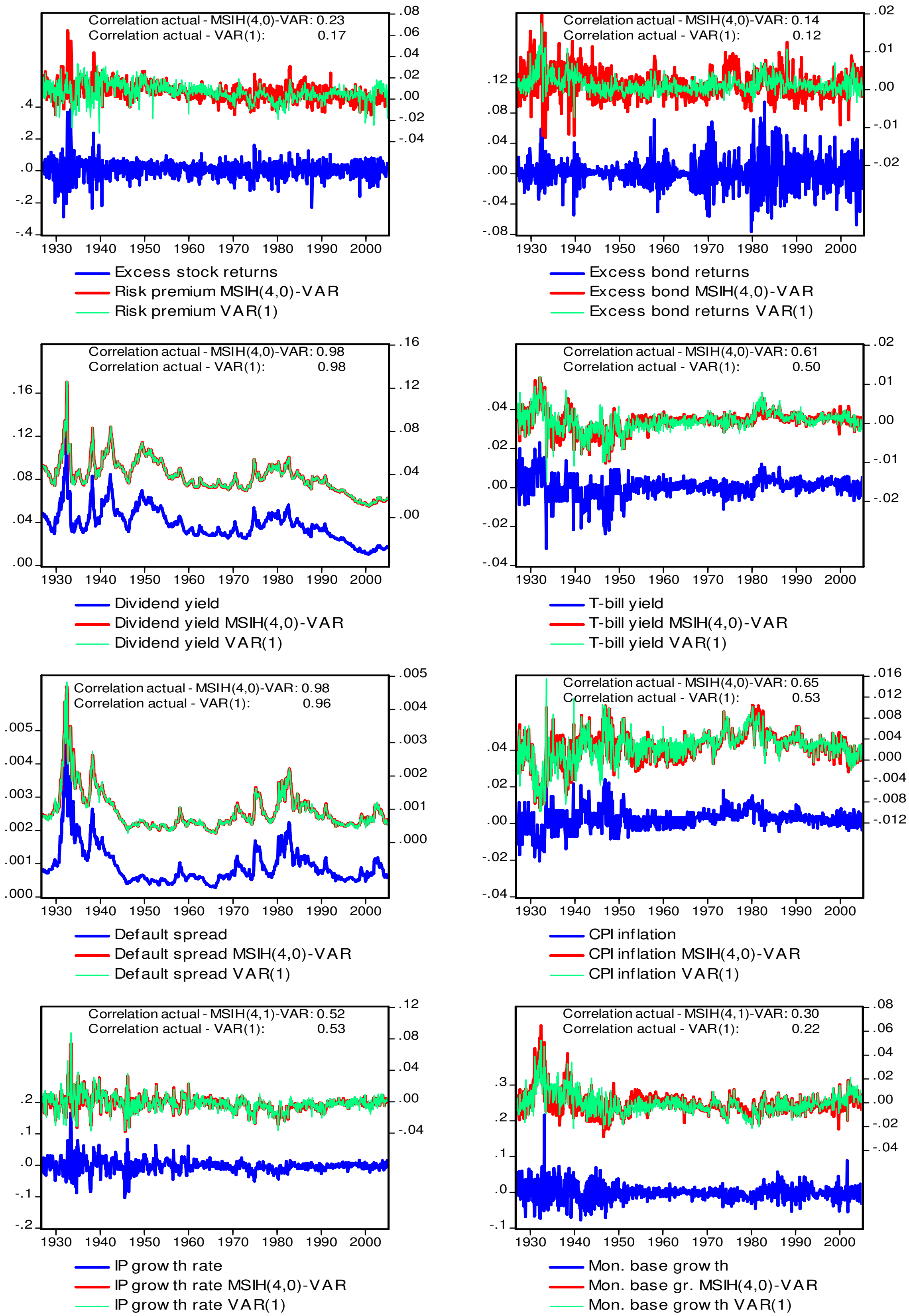

Figure 8

Recursive Likelihood Ratio Statistics and Hannan-Quinn ICs: VAR(1) vs. Four-State Model

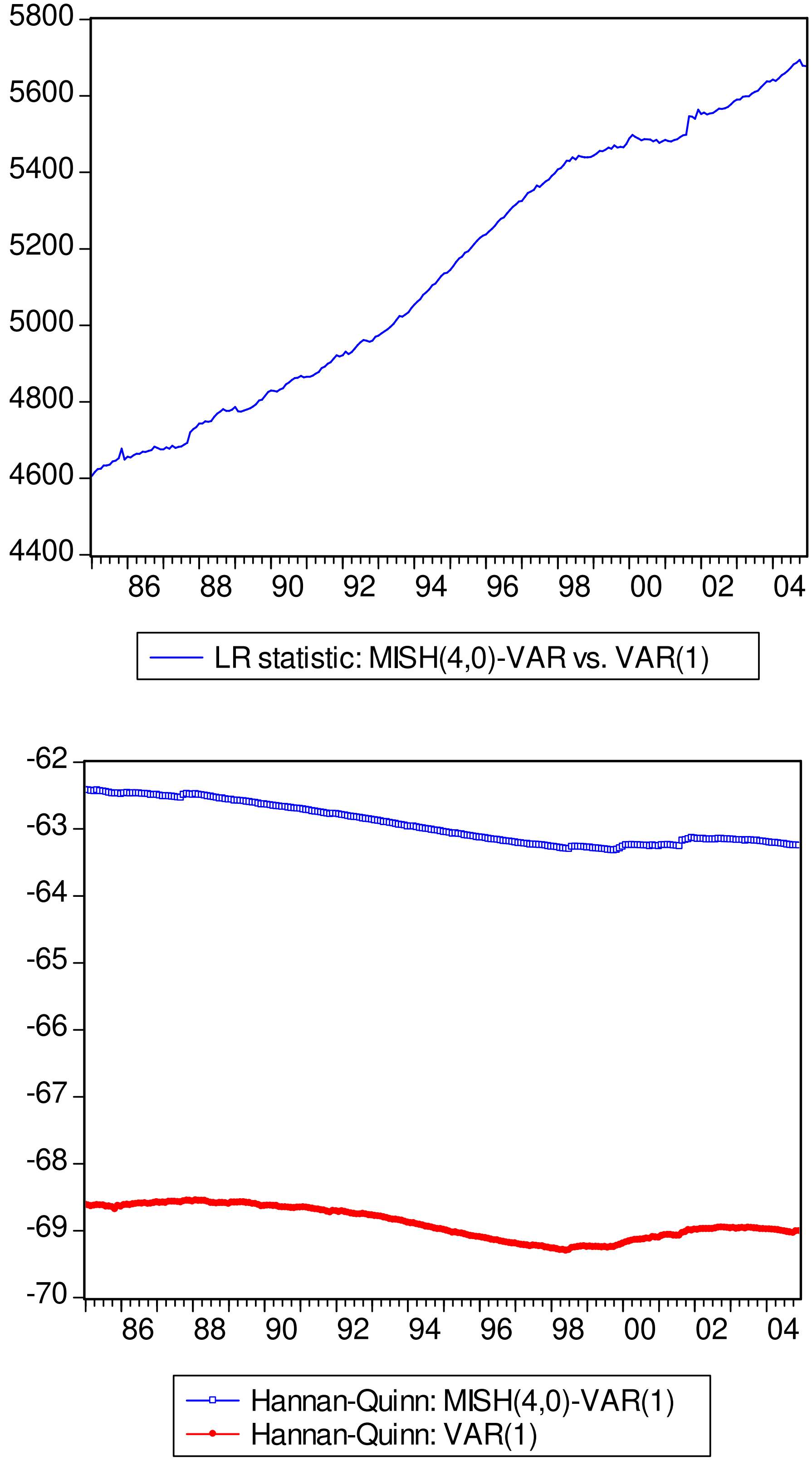

Figure 9
Implied Monthly Predicted Sharpe Ratios from a Four-State Model