ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Can VAR Models Capture Regime Shifts in Asset Returns? A Long-Horizon Strategic Asset Allocation Perspective

Authors	Massimo Guidolin, and Stuart Hyde
Working Paper Number	2010-002B
Revision Date	August 2010
Citable Link	https://doi.org/10.20955/wp.2010.002
Suggested Citation	Guidolin, M., Hyde, S., 2010; Can VAR Models Capture Regime Shifts in Asset Returns? A Long-Horizon Strategic Asset Allocation Perspective, Federal Reserve Bank of St. Louis Working Paper 2010-002. URL https://doi.org/10.20955/wp.2010.002

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

Can VAR Models Capture Regime Shifts in Asset Returns?

A Long-Horizon Strategic Asset Allocation Perspective*

Massimo Guidolin
Federal Reserve Bank of St. Louis and Manchester Business School†

Stuart Hyde
Manchester Business School

August 24, 2010

Abstract

We examine whether simple VARs can produce empirical portfolio rules similar to those obtained under a range of multivariate Markov switching models, by studying the effects of expanding both the order of the VAR and the number/selection of predictor variables included. In a typical stock-bond strategic asset allocation problem on US data, we compute the out-of-sample certainty equivalent returns for a wide range of VARs and compare these measures of performance with those typical of non-linear models that account for bull-bear dynamics and characterize the differences in the implied hedging demands for a long-horizon investor with constant relative risk aversion preferences. In a horse race in which models are not considered in their individuality but instead as an overall class, we find that a power utility investor with a constant coefficient of relative risk aversion of 5 and a 5-year horizon, would be ready to pay as much as 8.1% in real terms to be allowed to select models from the MS class, while analogous calculation for the whole class of expanding window VAR leads to a disappointing 0.3% per annum. We conclude that most (if not all) VARs cannot produce portfolio rules, hedging demands, or out-of-sample performances that approximate those obtained from equally simple non-linear frameworks.

Key words: Predictability, Strategic Asset Allocation, Markov Switching, Vector Autoregressive Models, Out-of-Sample Performance.

JEL codes: G11, C53.

1. Introduction

Since the seminal contributions by Brennan et al. (1997) and Kandel and Stambaugh (1996), the empirical finance literature on normative long-run dynamic asset allocation under predictable returns (i.e. how much

*We would like to thank Jerry Coakley, "Paul" Moon Sub Choi (a discussant), Paolo Colla, Carlo Favero, René Garcia, Abraham Lioui, Patrick Minford, Francesco Saita, Andrea Sironi, Nick Taylor, and participants at the INFINITI Conference on International Finance, Dublin 2010, the 3rd International Conference on Computational and Financial Econometrics, Limmasol 2009 and seminar participants at Bocconi University Milan (Finance dept.), Cardiff Business School, EDHEC Nice, and Essex School of Business.

†Correspondence to: Massimo Guidolin, Federal Reserve Bank of St. Louis, Research Division – St. Louis, MO 63166, USA. Phone: 314-444-8550, Fax: 314-444-8731. E-mail: Massimo.Guidolin@stls.frb.org.

should a risk-averse investor weight each available asset) has exclusively devoted its attention to the portfolio implications of linear predictability models. In a linear predictability model, asset returns are simply forecast by past values of a selected number of predictor variables (such as the dividend yield, the term spread, and the default spread, besides lagged values of the returns themselves) within a vector autoregressive (VAR) framework. The linearity consists of the fact that usually a movement today in one or more of the predictors, commands a proportional response in the expected (predicted) value of future asset returns. However, another strand of the empirical finance literature has in the meantime stressed that returns on most asset classes do contain predictability patterns that are not simply linear, as they involve non-linear patterns (such as regimes, thresholds, self-exciting mean reversions, conditional heteroskedasticity, etc.) that often make not only expected asset returns but also higher-order moments predictable.[1]

Although linear models are key benchmarks in empirical finance and their simplicity makes them obvious choices in many applications, their use in asset allocation applications has often relied on two often-implicit premises. First, that although most normative papers have to be taken as indicative examples of how practical portfolio choice ought to proceed, even when the scope of the investigation is extended beyond the class of small-scale (i.e. with 3-4 predictors at most) VAR(1) models typical in this literature (see e.g. Barberis, 2000, and Lynch, 2001), *some* more complicated VAR must surely exist that is of practical use in terms of consistently improving realized portfolio performances. This means that some VARs can be found that can efficiently summarize the overall balance of predictability in asset returns and that would make the modeling of any residual non-linear effects of second-order importance, at least in terms of impact on portfolio weights and performance. Second, that although more complicated, large-scale VAR(p) models may yield complex portfolio strategies, surely simple, small-scale VAR(1) models must be illustrative already of the first-order effects of linear predictability on dynamic portfolio selection, for instance in terms of implied hedging demands. Our paper tackles both these conjectures at their roots and provides a systematic examination of whether, when, and how small- and medium-scale VAR(p) models typical of the empirical finance literature may deliver dynamic portfolio choices that: (i) are able to approximate the portfolio choices typical of an investor that exploits both linear and non-linear predictability patterns in the data, and (ii) that compete in terms of realized portfolio performance with more complicated models able to capture also any non-linear predictability patterns.

As econometricians would expect on theoretical grounds, our relatively large set of small- and medium-scale (up to 7 predictors are included) VAR(p) models (with $p = 1$, 2, 4, and 12) fails to imply portfolio choices that approximate those from a rather simple (one may say, naive) non-linear benchmark, represented by a plain vanilla 3-state Markov switching (MS) model.[2] This is of course only an ex-ante perspective

[1]The literature on non-linearities in finance is rather voluminous and always growing. A few basic elements are discussed in the books of Campbell et al. (1997) and in Granger and Terasvirta (1993). A much smaller set of papers has also investigated the implications for optimal portfolio choice of non-linear dynamics in asset returns, such as Ang and Bekaert (2002, 2004), Detemple et al. (2003), Guidolin and Timmermann (2008a). Additional references relevant to specific issues of implementation are reported in the main body of the paper.

[2]This alludes to the well-known result (Wold decomposition theorem) that all covariance stationary vector time series may

on the problem: "different" does not imply "worse" in the view of an applied portfolio manager and what could be wrong is not the family of VARs, but the proposed non-linear benchmark. More importantly, VARs systematically fail to perform better than non-linear models in recursive (pseudo) out-of-sample tests, in the sense that VARs generally produce lower realized certainty equivalent returns (i.e. risk-adjusted performances that take into account of the curvature of the utility function under which the portfolio choice program has been solved) than multi-state models. This means that VARs cannot provide approximation results either ex-ante or ex-post.

The easiest way to summarize in quantitative terms the many results in this paper is with reference to the "class-level" horse race we have performed in Section 5.2. Even if we consider an investor that is actually contemplating resorting to a VAR modeling strategy to support her long-horizon SAA decisions, it is very unlikely that this investor will actually decide to specify and estimation one particular VAR model and to stick to it over time. An investor is likely to use statistical criteria to judge the likely performance of competing VAR models at each point in time, with the possibility of occasionally switching among different VARs. We have therefore endowed our VAR investor with the ability to recursively track over time the value of two information criteria, the AIC and the BIC, to decide which VAR model should be used for her asset allocation decisions. To favor comparability, we have applied an identical logic to the Markov switching class. We find that a power utility investor with coefficient of relative risk aversion of 5 and a 5-year horizon, would recursively select among MS models using a BIC minimization, over all possible classes (sets) of VARs. In fact, while this investor would be ready to pay up to 8.1% in real, annualized terms to access portfolio strategies in the MS set, the corresponding real CER is at most 0.3% for VARs.

These results are obtained with reference to a strategic asset allocation (SAA) application that appears to have played a key role in the literature on empirical portfolio choice (see Brennan et al., 1997, Barberis, 2000, Guidolin and Timmermann, 2007): a standard risk-averse (power utility, with constant relative risk aversion) investor wants to allocate at time t her wealth across three macro-asset classes, i.e., stocks (as represented by a standard value-weighted index), long-term, default risk-free government bonds, and 1-month Treasury bills. We use monthly US data for the long period 1953-2008 which also includes the recent financial crisis. We focus on long-horizon portfolio choices (up to 5-year an horizon) of an investor that recursively solve a portfolio problem in which utility derives from real consumption (i.e., cash flows obtained from dividend and coupon payments and from disinvesting securities in the portfolio) and rebalancing is admitted at the same frequency as the data (see Barberis, 2000, and Lynch, 2001). This means that even when the problem solved is characterized by a 60-period ahead horizons (5 years), the investor decides at time t knowing that at times $t+1$, $t+2$, ..., up to $t+59$ she will be allowed to change the structure of her portfolio weights to reflect the fact that at least in principle new information will arrive to her all these

be represented as VARMA processes with appropriate structure. Aside from that the empirical portfolio choice literature seems to only reflect a role for VARs (as opposed to VARMAs), we note that the evidence against the null of covariance stationarity in financial time series is massive and leaves little uncertainty on the usefulness of this result. In fact, no general VAR(MA)-type approximation result is known for strictly (as opposed to covariance-) stationary processes.

future points, possibly requiring a need to re-shuffle portfolio weights. Such a portfolio problem seems to be most appropriate one, not only for its past role in the development of the literature but also for the specific features of our research design. First, a long-horizon is key when discussing the economic value of predictability or – as in our case – of the relative economic value of different types of models in capturing whatever predictability is expressed by the data under investigation. Second, our attention to a problem with continuous/frequent rebalancing of portfolio weights and in which investors care for real consumption streams and real portfolio returns is also consistent with the way predictability is exploited in practice, i.e. with full awareness of the fact that its existence not only affects today's choice but will keep affecting them in all subsequent periods.

Finally, we stress that thrust of our exercise does not consist of investigating the different portfolio implications and out-of-sample performance of linear vs. non-linear models, as this operation has already appeared in the literature for specific linear and non-linear frameworks (see e.g. Detemple et al., 2003, Guidolin and Timmermann, 2007). In essence, these papers try and measure the economic loss from model misspecification in (density) forecasting applications by resorting to portfolio choice metric, as in Bauwens, Omrane, and Rengifo (2010). On the contrary, our point in this paper is to oppose a large set of VAR models potentially spanning a large portion of the models that have appeared in the literature to one single, and also relatively simple, non-linear framework which is selected to be of a Markov switching type as this class model has proven relatively popular and intuitive in the recent finance literature (see e.g. Perez-Quiros and Timmermann, 2000). The large family of VARs is obtained by investigating the forecasting performance, the implied dynamic recursive portfolio choices, and the resulting recursive out-of-sample performance of all VARs one can form using 7 predictors besides lagged values of asset returns themselves (in principle this is a total of 3,628 different VARs, taking into account that all VARs also include lagged values of asset returns and the one candidate VAR is obtained by including only such lagged values), and experimenting with 4 alternative lag orders throughout, $p = 1$, 2, 4, and 12. The seven predictors used are typical in the finance literature and include a few typical macro-finance variables, i.e., the dividend yield, the riskless term spread, the default spread between Baa and Aaa corporate bonds, the CPI inflation rate, the nominal riskless 3-month T-bill rate, the rate of growth of industrial production, and the unemployment rate. Our question is whether it is easy to select a VAR that may approximate portfolio choices and performance that would be given by a slightly more carefully chosen model, in this case with Markov switching features. As we have stated already, it then turns out that under many realistic circumstances it is actually *impossible* (hence it is really not that easy) to achieve this goal, in the sense that VARs do not appear fit to pick-up non-linear predictability patterns. Although this may seem obvious *ex-ante* to some of our Readers, what is not obvious is that in recursive out-of-sample tests such non-linearities seem to be then real and strong enough to condemn most (sometimes all) VARs to disappointing long-run portfolio *ex-post* performances.

The rest of the paper is structured as follows. Section 2 describes the research design of our paper. Although this is generally the case, an empirical exercise such as ours suffers from the fact that all results are the product of the choices we have made in terms of model construction, portfolio choice, and performance measurement. Therefore it is important to try and be as specific as possible on these details, if the goal is to persuade a Reader that our findings are relevant. Section 3 describes the data in our application and devotes some space to both the 3-state Markov switching benchmark employed in this paper to summarize both linear and non-linear predictability patterns and some common features of the adopted VARs. This section also shows that VARs manifest some problems already at the stage of offering sufficiently accurate forecasts of future returns, in particular stock returns. Section 4 computes and presents optimal portfolio weights and hedging demands under the two classes of models entertained in this paper. Section 5 computes realized, recursive out-of-sample portfolio performances. Section 6 performs an important robustness check and asks whether our results may mostly derive by the fact that the non-linear framework specified in Section 3 is fit to capture predictability in second moments, a task obviously impossible to any VAR. Section 7 concludes.

2. Methodology

This Section documents the models and performance indices used in the rest of our analysis, cutting comments and references to the minimum. We also provide details on the portfolio selection problem and required solution methods.[3] Finally but crucially, we describe in detail our recursive (pseudo-) out of sample research design.

2.1. *Econometric Models*

In this paper we perform recursive estimation, assessment of forecasting accuracy, and portfolio weight calculation and assessment for three groups of models. First and foremost, we entertain a large class of VAR(p) models. These VARs consists of a linear relationship linking $\mathbf{r}_{t+1}$, a $N \times 1$ vector of risky real assets at time $t+1$, and $\mathbf{y}_{t+1}$, a $M \times 1$ vector of predictor variables at time $t+1$, to lags of both $\mathbf{r}_{t+1}$ and $\mathbf{y}_{t+1}$. For instance, in the case of a VAR(1), we have

$$\begin{bmatrix} \mathbf{r}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = \boldsymbol{\mu} + \mathbf{A} \begin{bmatrix} \mathbf{r}_{t} \\ \mathbf{y}_{t} \end{bmatrix} + \boldsymbol{\varepsilon}_{t+1} \qquad \boldsymbol{\varepsilon}_{t+1} \sim IID\ N(\mathbf{0}, \boldsymbol{\Omega}), \tag{1}$$

where $\boldsymbol{\mu}$ is a $(N+M) \times 1$ vector of intercepts, $\mathbf{A}$ is a $(N+M) \times (N+M)$ coefficient matrix, and $\boldsymbol{\varepsilon}_{t+1}$ is a $(N+M) \times 1$ vector of IID, Gaussian residuals. The representation of a VAR(1) in equation (1) is without loss of generality as any p order VAR can be re-written as a VAR(1) (see Hamilton, 1994). In this paper we consider multiple values of p, $p = 1, 2, 4$, and 12. Note that – if one accepts to always include the lagged values of real asset returns in (1) – for given value of p there are 2^M different VARs we can obtain according to which of the M predictors are included in $[\mathbf{r}'_{t+1}\ \mathbf{y}'_{t+1}]'$.

[3] References to the econometrics of dynamic portfolio selection can be found in Brandt (2004). The solution of dynamic portfolio choice problems under linear and non-linear predictability is described in Guidolin and Timmermann (2007, 2008b), as well as in Detemple et al. (2003) and Brandt (2004).

The second class of models consists of non-linear models of the k-state Markov switching class with constant transition probabilities (collected in a $k \times k$ matrix $\mathbf{P}$)

$$\mathbf{r}_{t+1} = \boldsymbol{\mu}_{S_{t+1}} + \boldsymbol{\varepsilon}_{t+1} \qquad \boldsymbol{\varepsilon}_{t+1} \sim IID\ N(\mathbf{0}, \boldsymbol{\Omega}_{S_{t+1}}), \tag{2}$$

where the latent Markov state $S_{t+1} = 1, ..., k$ and $\boldsymbol{\mu}$ is a $N \times 1$ vector of state dependent intercepts. One may also allow for the $N \times N$ covariance matrix of residuals $\boldsymbol{\Omega}$ to be state dependent, implying the variance of the asset returns is also state-dependent, i.e., $Var[\mathbf{r}_{t+1}|S_{t+1}] = \boldsymbol{\Omega}_{S_{t+1}}$. Under (2) asset returns are predictable because their density (visibly, the first two moments, although this property extends beyond means, variances, and covariances) are predictable.[4] This obviously derives from the fact that in general—unless particular configurations of the Markov transition matrix apply—Markov chains are predictable processes. Since the state is a complicated non-linear function of all past data before time t, such a predictability pattern is best thought of as a *non-linear* one. There is another sense in which MS implies non-linear predictability: because what is (at most) predictable is when and how the markets will *switch* from one regime to others, these switches may be described as "jumps" in the joint density of the data and as such jumps are best described as non-linear phenomena. In the following, we refer to (2) as MSIH when $\boldsymbol{\Omega}$ is state-dependent, and as MSI when $\boldsymbol{\Omega}$ is constant over time.

The third class of models is obtained at the intersection between the first two classes—these are Markov switching VAR(p) models with structure (e.g., in the simplest case of $p = 1$)

$$\begin{bmatrix} \mathbf{r}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = \boldsymbol{\mu}_{S_{t+1}} + \mathbf{A}_{S_{t+1}} \begin{bmatrix} \mathbf{r}_{t} \\ \mathbf{y}_{t} \end{bmatrix} + \boldsymbol{\varepsilon}_{t+1} \qquad \boldsymbol{\varepsilon}_{t+1} \sim IID\ N(\mathbf{0}, \boldsymbol{\Omega}_{S_{t+1}}), \tag{3}$$

where once more the latent state $S_{t+1} = 1, ..., k$ follows a first-order Markov chain. Clearly, (3) allows the coexistence of both linear and non-linear predictability patterns, as well as of rich interaction effects among the two (see Guidolin and Timmermann, 2007, for further details), the former driven by the classical vector autoregressive structure, the latter by the predictability of the driving Markov state process. However, the fact that the VAR matrices themselves may be a function of the state S_{t+1}, potentially adds to the complexity of the predictability patterns that may be captured. In the following, we refer to (3) as MSVARH(p) when $\boldsymbol{\Omega}$ is state-dependent, and as MSVAR(p) when $\boldsymbol{\Omega}$ is constant over time.

Finally, we also consider a further benchmark class widely adopted in the empirical finance and forecasting literature, a simple Gaussian IID model:

$$\mathbf{r}_{t+1} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_{t+1} \qquad \boldsymbol{\varepsilon}_{t+1} \sim IID\ N(\mathbf{0}, \boldsymbol{\Omega}), \tag{4}$$

which is obviously the single-state restriction of (2). Under (4) asset returns are not predictable. In fact, under appropriate definitions of continuously compounded asset returns (cum dividend), it is easy to show that (4) derives from a simple random walk with drift process for log-asset values.

[4] The predictability of the regimes and hence of the joint $t + H$-ahead density of the data implies that not only moments but more generally densities are predictable under a MS model. See Perez-Quiros and Timmermann (2001) and Guidolin and Ono (2006) for additional details.

2.2. *The Portfolio Choice Problem*

Consider the portfolio and consumption decision of a finite horizon investor with time-separable utility, constant relative risk aversion (CRRA) who maximizes the expected utility of lifetime consumption

$$\max_{\{C_\tau,\omega_{i,\tau}\}_{\tau=1:i=1}^{H-1:N}} \sum_{t=1}^{H} \beta^t E_t \left[\frac{C_t^{1-\gamma}}{1-\gamma} |\mathbf{Z}_t \right] \qquad \beta \in (0,1)\,,\ \gamma > 1, \tag{5}$$

where the discount factor $\beta = 0.9975$ is the subjective rate of time preference (corresponding to an annualized real discount rate of less than 3%), the coefficient γ measures relative risk aversion, C_t is the investor's consumption at time t and $\mathbf{Z}_t$ is the relevant vector of state variables at time t.[5] The investor consumes a proportion of wealth, $\kappa_t \equiv C_t/W_t$, allocating the remainder to an investment portfolio consisting of the N real risky assets. The return on the portfolio, $r_{p,t+1}$ is then given by $\sum_{i=1}^{N} \omega_{i,t} r_{i,t+1}$ where the weights, $\omega_{i,t}$, allocated to each risky asset must sum to unity, i.e. $\sum_{i=1}^{N} \omega_{i,t} = 1$. The intertemporal budget constraint faced by the investor is

$$W_{t+1} = (W_t - C_t)(1 + r_{p,t+1}) = W_t (1-\kappa_t) R_{p,t+1}, \tag{6}$$

where $R_{p,t+1}$ is the gross portfolio return, $R_{p,t+1} \equiv 1 + r_{p,t+1}$. It is easy to show (see Ingersoll, 1987) that the Bellman equation faced by the investor for a CRRA utility function that can be derived from (5) and the budget constraint (6) is

$$\frac{a(\mathbf{Z}_t,t)\, W_t^{1-\gamma}}{1-\gamma} = \max_{\kappa_t,\boldsymbol{\omega}_t} \left\{ \frac{\kappa_t^{1-\gamma} W_t^{1-\gamma}}{1-\gamma} + \frac{\beta (1-\kappa_t)^{1-\gamma} W_t^{1-\gamma}}{1-\gamma} E\left[a(\mathbf{Z}_{t+1}, t+1)\, R_{p,t+1}^{1-\gamma} |\mathbf{Z}_t \right] \right\}, \tag{7}$$

where $a(\mathbf{Z}_t, t)$ is a function that can be computed numerically. Given that this optimization problem is homogeneous of degree $(1-\gamma)$ in wealth, the solution is invariant in wealth. Hence the Bellman equation can be simplified to:

$$\frac{a(\mathbf{Z}_t,t)}{1-\gamma} = \max_{\kappa_t,\boldsymbol{\omega}_t} \left\{ \frac{\kappa_t^{1-\gamma}}{1-\gamma} + \frac{\beta (1-\kappa_t)^{1-\gamma}}{1-\gamma} E\left[a(\mathbf{Z}_{t+1}, t+1)\, R_{p,t+1}^{1-\gamma} |\mathbf{Z}_t \right] \right\}. \tag{8}$$

Equation (8) can then be solved by backward iteration, starting with $t = T-1$ and setting $a(\mathbf{Z}_T, T) = 1$ and then computing $a(\mathbf{Z}_T, t)$ by solving the optimization problem in equation (8) using $a(\mathbf{Z}_{t+1}, t+1)$ from the previous iteration. The backward, recursive structure of the solution reflects the fact that the investor incorporates in the optimal weights computed at time t the fact that such weights will be revised in the future at times $t+1$, $t+2$, ..., $t+H-1$ as new information becomes available through the vector of state variables $\mathbf{Z}_t$. A variety of solution methods are applied in the literature on portfolio allocation under predictable returns. Following Guidolin and Timmermann (2007, 2008b) we employ Monte Carlo methods for integral (expected utility) approximation. Appendix A provides additional details on the numerical methods used in the solution of the portfolio problem.

[5] In the case of a VAR(p), $\mathbf{Z}_t \equiv [\mathbf{r}_t' \ \mathbf{y}_t' \ \mathbf{r}_{t-1}' \ \mathbf{y}_{t-1}' \ ... \ \mathbf{r}_{t-p+1}' \ \mathbf{y}_{t-p+1}']'$ so that the state vectors consists of a combination of lagged value of asset returns and predictor variables. In a MSI/MSIH framework $\mathbf{Z}_t$ consists instead of the vector of state probabilities estimated at time t. Finally, in a MSVAR/MSVARH model, $\mathbf{Z}_t$ consists of both the lagged values of asset returns and predictors, and of the vector of state probabilities.

2.3. *Measuring Forecasting Performance in Recursive Out-of-Sample Experiments*

Our (pseudo) out-of-sample (OOS) experiment has a recursive, expanding structure. This means that at the first iteration we estimate all models (e.g., in the case of VARs these are 512 different linear frameworks) using data for the period 1953:01-1973:01 and then proceed to compute: (i) forecasts at horizons $H = 1$, 12, and 60 months; (ii) portfolio weights at horizons $H = 1$ and 60 months, in the latter case with continuous (i.e., monthly, at the same frequency as the data) rebalancing. The forecasts are produced for both point returns and cumulative returns. For instance, the forecasts will refer to returns predicted for 1978:01, the sum of returns for all months between 1973:02 and 1978:01, and the portfolio weights will be the optimal ones for the period 1973:02-1978:01, when rebalancing can be performed at the end of every month. At this point, the estimation sample is extended by one additional month, to the period 1953:01-1973:02, producing again forecasts at horizons $H = 1$, 12, and 60 months and portfolio weights at horizons of 1 and 60 months. This process of recursive estimation, forecasting, and portfolio solution is repeated until we reach the last possible sample, 1953:01-2008:12 (even though in this case the OOS predictive or portfolio performance cannot be computed as our sample ends in 2008:12).[6]

We also implement a rolling forecasting scheme based on a 10-year window. The 10-year window is selected to allow the estimation of somewhat large models, such as VAR(4) including all predictors (these imply 465 parameters with 1,200 available observations). At the first iteration we estimate all models using data for the period 1963:02-1973:01 and then proceed to compute forecasts and portfolio weights at horizons at horizons $H = 1$, 12, and 60 months. At this point, the estimation sample is updated by adding one additional month at the end of the sample and dropping the first month at the beginning of the sample, so that the resulting period becomes 1963:03-1973:02, producing again forecasts and portfolio weights at the usual horizons. This process of recursive estimation, forecasting, and portfolio solution is repeated until we reach the last possible sample, 1998:01-2008:12.

We define the time t forecast error at horizon H for the real return on asset j (stocks, bonds, 1-month T-bills) as:

$$e^j_{t,t+H} \equiv r^j_{t+H} - \hat{r}^j_{t,t+H}, \tag{9}$$

where $\hat{r}^j_{t,t+H}$ is the generated H-step ahead forecast and r^j_{t+h} is the realized return. In the case of cumulative returns, we have instead that:

$$e^j_{t,H} \equiv \prod_{h=1}^{H}(1 + r^j_{t+h}) - \prod_{h=1}^{H}(1 + \hat{r}^j_{t,t+h}).$$

To evaluate the OOS forecast performance we employ four standard metrics (for simplicity we do not to distinguish between $e^j_{t,t+H}$ and $e^j_{t,H}$):

[6]We need to stress that this OOS experiment does not represent a genuine OOS design since a few (although rather marginal) features of the experiment are designed exploiting end-of-sample hindsight, for instance concerning the most appropriate number of regimes in the specification of the non-linear benchmark (see Section 3.1 for details). Whenever we talk about out-of-sample results we have this important *caveat* in mind.

1. **Root Mean Squared Forecast Error (RMSFE).** The RMSFE is computed as

$$RMSFE_H^j \equiv \sqrt{\frac{1}{T-H}\sum_{t=1}^{T-H}(e_{t,t+H}^j)^2}, \tag{10}$$

where T is the total sample size available for the recursive OOS prediction exercise.

2. **Forecast Error Bias.** The bias is simply the signed sample mean of all forecast errors:

$$Bias_H^j \equiv \frac{1}{T-H}\sum_{t=1}^{T} e_{t,t+H}^j. \tag{11}$$

A large, signed value of the bias indicates a systematic tendency of forecasts to either over- or under-predict asset returns.

3. **Forecast Error Variance (FEV).** While the definition is straightforward,

$$FEV_H^j \equiv \frac{1}{T-H}\sum_{t=1}^{T-H}(e_{t,t+H}^j)^2 - \left[\frac{1}{T-H}\sum_{t=1}^{T} e_{t,t+H}^j\right]^2 = \frac{1}{T-H}\sum_{t=1}^{T-h}(e_{t,t+H}^j)^2 - [Bias_H^j]^2, \tag{12}$$

one useful fact is that $FEV_H^j + [Bias_H^j]^2 = MSFE_H^j$, i.e. large MSFEs (poor performance) may derive from either high forecast error variance or from large average bias. We normally report forecast error standard deviation, i.e., the square root of FEV.

4. **Mean Absolute Forecast Error (MAFE).** Similar to the RMSFE, the difference being that signs are neutralized using absolute values rather than by squaring:

$$MAFE_H^j \equiv \frac{1}{T-H}\sum_{t=1}^{T}\left|e_{t,t+H}^j\right|. \tag{13}$$

As it is well known, this statistic is more robust to the presence of outliers than the RMSFE.

2.4. *Performance Measurement*

To evaluate recursive OOS portfolio performance we focus on two key measures. First, we calculate the certainty equivalent return (CER), defined as the sure real rate of return that an investor is willing to accept rather than adopting a particular risky portfolio strategy. We (numerically) compute/solve for CER as:

$$\sum_{t=1}^{H}\beta^t E_t\left[\frac{\hat{C}_t^{1-\gamma}(\hat{\boldsymbol{\omega}}_t)}{1-\gamma}\right] = \sum_{t=1}^{H}\beta^t E_t\left[\frac{\tilde{C}_t^{1-\gamma}}{1-\gamma}\right] \quad with \quad \tilde{C}_t = \frac{1-\beta CER^{1-\gamma}}{1-(\beta CER^{1-\gamma})^{(T-t+1)/\gamma}}, \tag{14}$$

where $\tilde{C}_t$ is the monthly consumption flow an investor receives under a constant investment opportunity set simply composed of a riskless real asset that yields a monthly certainty equivalent of CER. Second, we also compute the out-of-sample Sharpe Ratio for each portfolio strategy, defined as the mean OOS excess portfolio return divided by the standard deviation.

3. Data and Preliminary Evidence

We use monthly data on real asset returns and a standard set of predictive variables sampled over the period 1953:01-2008:12. The data are obtained from CRSP and FRED® at the Federal Reserve Bank of St. Louis. The real asset return data are the CRSP value weighted equity return, the 10-year bond return and the 30-day Treasury bill return, all deflated by the CPI inflation rate. The predictive variables are the dividend yield on equities (computed as a moving average of the past 12-month dividends on the CRSP value-weighted index divided by the lagged index), the short-term interest rate (3 month Treasury bill yield), the CPI inflation rate, the term spread defined as the difference between long- (10 year) and short-term (3 month) government bond yields, the default spread defined as the difference between the yields on Baa and Aaa corporate bonds, the rate of industrial production growth, and the unemployment rate. Our choice of predictor variables is governed by the existing literature on return predictability which provides evidence of the forecasting ability of the dividend yield (e.g., Fama and French, 1988, 1989), short-term interest rates (see Campbell, 1987, Detemple et al., 2003, Ang and Bekaert, 2007), inflation (e.g., Fama and Schwert, 1977, Campbell and Vuolteenaho, 2004), the term and default spreads (Campbell, 1987, Fama and French, 1989), industrial production (e.g., Cutler et al., 1989, Balvers et al., 1990) and the unemployment rate (see Boyd et al., 2005). Notice that 7 predictors and 4 alternative values of p imply that $4 \times 2^7 = 512$ alternative VAR models, as initially stated.

Descriptive statistics for asset returns and predictor variables are reported in Table 1. Mean real stock returns are close to 0.59% per month with mean real long-term bond returns around 0.23% implying annualized returns of 7.1% and 2.8% respectively. Estimates of volatility imply annualized values of around 15% for real stock returns and 7.7% for real bond returns, yielding unconditional Sharpe ratios of 0.11 and 0.06 respectively. In annualized terms (these are useful for comparisons to be performed later), these correspond to Sharpe ratios of 0.39 and 0.21, respectively. Real asset returns are characterized by significant skewness and kurtosis and are clearly non-Gaussian, as signalled by the rejections of the (univariate) null of normality delivered by the Jarque-Bera test.

The rest of this Section is devoted to a number of related sets of estimation results that need brief comment as a way of introducing the main results in Sections 4-6. In Section 3.1 we outline some evidence on the nature and strength of the linear predictability patterns – as picked up by simple VARs typical of the empirical finance literature – that characterize our data on US stock and bond real returns. The objective here is not (and it could not be) to provide an exhaustive quantification of what linear predictability implies, but to at least provide some evidence for how this predictability may appear in a VAR vs. MSI and MSVAR models. This gives us the opportunity to collect signals of misspecifications in linear models and to discuss (at least in an ex-ante perspective) what types of VARs are most likely to succeed in forecasting US real asset returns. In fact, in Section 3.2 we use the estimates from a simple two-state MSVAR model to document the presence of structural instability in VAR models. In Section 3.3 we briefly discuss the properties and implications of our estimates of a simple three-state MSI model. In Section 3.4 we do the same with reference to models in MSVAR class. The number of details and depth of description is kept to a minimum because the goal of our paper is not to analyze the portfolio choice implications of Markov switching models (a task already undertaken by Ang and Bekaert, 2002, and Guidolin and Timmermann, 2007, 2008a) but instead whether standard VAR models can approximate the portfolio implications of MSI and/or MSVAR. Section 3.5 presents a few results on the OOS forecasting performance of Markov switching vs. the VARs models entertained in this paper.

3.1. *Linear Predictability*

Figure 1 plots the own- and cross-correlograms functions for real stock, bond, and T-bill returns (up to lag 24), where the cross-correlograms are computed with reference not only to lagged real asset returns but also to lagged values of the 7 predictors used in this paper.[7] The shaded regions show the interval of values on the vertical axis for which the cross-correlation coefficients fail to be statistically significant (i.e., the null of the coefficient being equal to zero cannot be rejected) at a size of 5% (i.e., absence of predictability). Values of the cross-serial coefficients which are statistically significant are also highlighted by using larger font. Clearly when the plots report values outside the shaded range, we are facing statistically significant (positive or negative) cross-correlation coefficients which may be exploited for prediction purposes and that should be picked up a carefully built VAR.[8] Although each of the panels in Figure 1 contain a large amount of information, some general lessons may be visualized already. First, there is very little predictability in real stock returns. The number of markers that fall outside the (rather large) shaded region is modest, only about a couple dozens out of 250. In particular, there is solid evidence that past values of the dividend yield forecast future real stock returns and that occasionally lagged real bond returns and the term spread may display some forecasting power. While these serial correlations are all positive, there is weak evidence that high inflation in the past forecasts subsequent, lower real stock returns.

There is stronger evidence of linear predictability in real bond returns. Even though the shaded region of no statistical significance is narrower in this case, there are indications that past values of the term spread, the default spread, the short nominal rate, and 1-month real T-bill returns predict higher real returns on long-term government bonds. In many cases, these linear patterns are very persistent over time, i.e., it is long lags of the predictors that forecast real bond returns. Also, the first two lags of inflation forecast lower subsequent real bond returns. Finally, Figure 1 makes it clear that – as one would expect

[7] A cross-correlogram function plots the value of the (sample) cross-correlation coefficient, $\hat{\rho}_q[X,Y] \equiv \widehat{Cov}[X_t, Y_{t-q}]/(\hat{\sigma}_X \hat{\sigma}_Y)$ between variables X and Y, as a function of the lag parameter $q = 0, 1, ..., 24$. When X and Y coincide we have a (own-) serial correlation function; when $q = 0$, we obtain the simultaneous correlation coefficient between X and Y (which is not relevant for prediction purposes); for completeness of information, these coefficients (not reported elsewhere in this paper) are marked on the vertical axis of Figure 1, using a bigger font when the coefficient is statistically significant. Clearly, $\hat{\rho}_0[X,X] = 1$ by construction and such trivial values are not plotted in Figure 1.

[8] However, it is obvious that the estimated (OLS) coefficients of a VAR will not simply correspond (or be proportional) to these cross-serial correlation coefficients. The multivariate nature of a VAR estimation problem breaks down the simple connection between cross-serial correlations and AR coefficients.

in the light of the literature – real 1-month T-bill returns are massively predictable. In this case, almost all predictors as well as lagged values of real bill returns themselves forecast real bill returns. In fact, it is much quicker to comment on which predictors fail to work for real 1-month T-bills: only past values of IP growth, the term spread, real bond and stock returns have weak predictive power. Naturally, a useful VAR ought to be able to pick up these linear predictability patterns and exploit them for SAA purposes.

3.2. *Instability in VARs*

It has been widely documented in the empirical finance literature that most patterns of linear predictability tends to be massively unstable over time: the predictors that forecast asset returns today are hardly the same as those that will forecast the same asset return series at a later point; moreover, even assuming the same predictor maintains some its forecasting power over time, it is common to find that the specific strength and "sign" of this predictability are often subject to sudden reversals (see e.g., Guidolin and Ono, 2006, and Paye and Timmermann, 2008). This pervasive instability also plagues the VAR models examined in this paper. However, dealing with 512 different linear predictability models, it is unusually hard to pin down the patterns and intensity of such instability. At an informal level, we have recursively estimated and examined parameter estimates for a range of VARs that appear to have been commonly employed in the literature, such as parsimonious VAR(1) models including each of the 7 predictors, one at the time, or a VAR(1) model that includes all the predictors proposed in this paper. For instance, Figures 4-6 present recursive OLS coefficient estimates (on an asset-by-asset basis) obtained from a VAR(1) under two alternative assumptions on the predictors: either all our 7 predictors appear or each of the 7 predictors appear one-by-one, in isolation.[9] In practice the plots span 8 different VARs among the 512 we recursively estimate in this paper. Although these are only 8 VARs, they are useful benchmarks to adopt. We have also plotted recursive coefficient estimates for either "intermediate" (i.e., with a number of predictors between 2 and 6, in different combinations) or "larger" VARs (i.e., including most or all predictors and characterized by a higher number of lags) and found qualitatively similar results. In particular, each panel in Figure 4 plots two recursive coefficient series (the solid lines), each with its implied (parametric) 95% confidence bands (the dotted lines): one series is obtained from the full VAR(1) model and the other from the single-predictor VAR(1), when the predicted variable is real stock returns. When *both* sets of 95% confidence intervals fail to include zero (which is an indication of strength of the predictable pattern in a statistical sense), the corresponding period is shaded to stress this is an interval in which linear predictability was present and this finding does not rely on the fine details of the VAR model estimated (hence the choice to require that the intervals do not include zero for both types of VARs plotted). The visual impression offered by Figure 4 on linear predictability of real stock returns is rather stark: there is little predictability in real stock returns and such forecastability essentially ends around 1987 to never re-emerge again. While lagged real bond returns and – to a lesser extent – the lagged term spread had predicted subsequent stock

[9] However, in both cases all lagged real asset returns series have been included as predictors.

market dynamics in the 1970s and early 1980s, such patterns have disappeared during the 1990s and recent years. The predictability from the lagged dividend yield to stock returns much debated in the empirical finance literature has been hardly present for real stock returns, with an isolated episode between 1979 and 1981, even though the p-value of the dividend yield coefficient remains between 0.05 and 0.10 for most of the 1980s and early 1990s (which is consistent with the evidence in papers such as Kandel and Stambaugh, 1996, and Barberis, 2000).

Figure 5 shows instead that—even though it comes with a very small sub-set of predictors—linear forecastability of real bond returns is stronger than in Figure 4 and that it has being increasing over time, appearing to peak after the early 1990s. Clearly, it is lagged real stock returns and the term spread (a variable that is important to understand the dynamics of real bond returns within the expectations hypothesis) that accurately predict subsequent real returns. There is also some weaker, episodic evidence that lagged real T-bill returns (but again, only late in the sample) may forecast long-term real bond returns, which makes sense within frameworks such as the expectations hypothesis. Figure 6 illustrates that, as one would expect, real 1-month bill returns are massively predictable and that this holds throughout our sample period, although the exact identity of the predictors undergoes a few changes. First of all, starting in 1981, there is an increasingly strong autoregressive component in real bill returns, with an AR(1) coefficient that goes from -0.1 in 1973 to 0.35 by the end of 2008. However, also the lagged nominal rate (in the 1970s) and the lagged term spread (after the 1980s) forecast future real bill returns. Although only episodically, also lagged real stock returns and dividend yields have some forecasting power for real bills. Obviously, the evidence in Figures 4-6 is broadly consistent with the patterns already noted in Figure 1 when commenting on cross-serial correlation coefficient patterns. However, it is hard to forget that such a compelling evidence of time-variation in the sign, magnitude, and statistical significance of the estimated coefficients does point towards the existence of pervasive misspecification problems with the family of VAR spanned by the 8 models presented in Figures 4-6. Finally, we notice that with very few (or no) exceptions, macroeconomic predictors such as the default spread, industrial production growth, and unemployment rate are never among the predictors for which the estimated coefficients are statistically significant.

In formal terms, we have exploited the convenience of MSVAR models to use a two-state homoskedastic MSVAR(1) model to try and summarize such instability. We stress that in this Section, the goal is not propose a Markov switching benchmark to be held firm throughout the rest of the paper. Section 3.3 proceeds to a rigorous model specification search to isolate the most sensible Markov switching models. In this section the goal is to simply provide some intuitions for the nature and pervasiveness of the time-variation that affects linear predictability as this may be captured by a simple VAR(1). We have estimated using the EM algorithm the two-state MSVAR(1)

$$\begin{bmatrix} \mathbf{r}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = [S_{t+1}\boldsymbol{\mu}_1 + (1-S_{t+1})\boldsymbol{\mu}_2] + [S_{t+1}\mathbf{A}_1 + (1-S_{t+1})\mathbf{A}_2] \begin{bmatrix} \mathbf{r}_t \\ \mathbf{y}_t \end{bmatrix} + \boldsymbol{\varepsilon}_{t+1} \qquad \boldsymbol{\varepsilon}_{t+1} \sim IID\ N(\mathbf{0}, \boldsymbol{\Omega}), \quad (15)$$

where $S_{t+1} = 1, 2$ and $\mathbf{y}_{t+1}$ includes the 7 predictors. To save space and also to avoid confusing a Reader with reference to the further MSVAR estimates that are reported in Section 3.4, we have not tabulated the estimates of the 277 parameters that such a seemingly innocuous MSVAR(1) implies.[10] We limit ourselves to report the most interesting parameter estimates when they are useful to shed light on instability issues. The first interesting finding is that—even though (using appropriately corrected likelihood ratio-type tests) there is statistical evidence of regimes in the data, for instance the Davies (1977) statistic is 170.2 with a p-value of 0—the underlying two-state Markov chain is hardly persistent, with estimated values $\Pr\{S_{t+1} = 1|S_t = 1\} = 0.46$ and $\Pr\{S_{t+1} = 2|S_t = 2\} = 0.51$ and implied persistence of approximately 2 months for both regimes. This means that (15) does capture instability, but not persistent patterns in the changes of linear, VAR-predictability. As we shall see in Sections 3.3-3.4, this may be a result of a need to specify a higher number of regimes, as it is likely to be required in a multivariate modelling exercises to rather different asset return series spanning both equity and bond markets (as in Guidolin and Timmermann, 2006). Second, the implied instability in linear predictive relationships is massive. For instance, the row of $\hat{\mathbf{A}}_1$ that captures any predictability in real stock returns in regime 1 is (absolute values of robust t-stats are in parenthesis below the estimated coefficients; we have boldfaced coefficients that are significant with a p-value of 0.10 or less)

r_{t-1}^{stock}	r_{t-1}^{bond}	r_{t-1}^{T-bill}	Div. Yield	Short Nom. Rate	Term	Default	CPI Inflation	IP growth	Unempl.
0.252	0.051	2.808	0.008	**-0.008**	-0.004	0.005	1.557	-0.0001	**0.005**
(0.84)	(0.39)	(3.97)	(2.74)	(2.57)	(1.13)	(0.43)	(0.46)	(0.87)	(1.82)

while the row of $\hat{\mathbf{A}}_2$ that captures any predictability in real stock returns in regime 2 is:

r_{t-1}^{stock}	r_{t-1}^{bond}	r_{t-1}^{T-bill}	Div. Yield	Short Nom. Rate	Term	Default	CPI Inflation	IP growth	Unempl.
-0.090	0.180	**9.941**	0.001	-0.005	0.004	-0.007	**9.379**	0.000	-0.0004
(1.61)	(1.36)	(2.67)	(0.40)	(1.58)	(1.30)	(0.82)	(2.56)	(0.19)	(0.15)

Clearly, there is "more" predictability in the first regime than in the second, at least in the sense that 3 predictors (the dividend yield, the short-term nominal rate, and the unemployment rate, plus lagged values of the real short-term rate) forecast one-step ahead real stock returns in the first regime, against one predictor only—and a different one, CPI inflation (besides lagged values of the real short-term rate)—in the second regime. Moreover, a number of coefficients switch signs across different states, although we have no case of switches of sign that preserve statistical significance. For instance, the dividend yield has a famous history as being unreliable and weak among the commonly used predictors of stock returns. The results from (15) stress one possible cause for such a reputation: approximately half of the time, the dividend yields is characterized by an economically small and imprecisely estimated effect on subsequent stock returns. Similarly, past inflation does forecast higher subsequent real stock returns, although this occurs only half of the time, so that the overall, unconditional "loadings" of real stock returns on inflation will be small and imprecisely estimated, as it has been documented by scores of papers. We have also plotted and examined plots over our sample period of predicted, one-step ahead VAR coefficients connecting real asset returns to

[10] Even though estimation proved possible (with 277 parameters and 6,710 observations we have an acceptable saturatio ratio of 24 observations per parameter), it proved very difficult in numerical terms, with considerable evidence of instability due to the presence of local maxima in the log-likelihood function.

predictors (and lagged values of real returns on other asset) and compare them to analogous plots in which one conditions on knowledge of the future regime. The difference is striking: in general, the one-step ahead predicted coefficients tend to be economically small and hardly relevant. However, if one were to condition on perfect-foresight knowledge of the prevailing regime one-month ahead, we have that a few predictors (especially in state 1) would make forecasting possible and somewhat more reliable.[11]

3.3. *Regimes in US Real Asset Returns*

Following common practice in the literature on optimal portfolio choice under Markov switching, as a first step we have estimated and compared a range of homoskedastic Markov models as distinguished by the number of regimes they require, k, and by the number of lags of predictors and real asset returns they employ, p.[12] Of course, when $p \geq 1$, different models will also be determined by which predictors they end up including. Table 2 reports summary statistics for a range of estimated models along the dimensions of k = 1,2, 3, and 4 and p = 0, 1, 2. In the case of $k = 1$—the standard VAR models—we report only a few cases for $p = 1$ and 2 just to provide some ideas on the relative fit provided by single- vs. multi-state models. All the VAR models with $p = 2$, 4, and 12 have information criteria that largely exceeds the tightly parameterized models with $p = 1$ in the Table. The statistics in Table 2 are the maximized log-likelihood function, an approximate nuisance parameter-adjusted likelihood ratio that tests the null of $k = 1$ against $k > 1$, three alternative information criteria (i.e., the Bayes-Schwartz, Akaike, and Hannan-Quinn criteria) that trade off in-sample fit for parsimony, where the latter is considered as an indicator of likely predictive accuracy, and the (saturation) ratio between the total number of observations used in estimation and the total number of parameters estimated. In the case of the information criteria, we have boldfaced the three best (yielding the lowest criteria) models according to each of the three criteria. Homoskedasticity is maintained throughout because we would like at this stage to maximize the degree of comparability between portfolio performance obtained from Markov switching and VAR models, where the latter are models of predictability in the conditional mean only.

[11] Results for real bond returns are

r_{t-1}^{stock}	r_{t-1}^{bond}	r_{t-1}^{T-bill}	Div. Yield	Short Nom. Rate	Term	Default	CPI Inflation	IP growth	Unempl.
-0.055	0.015	-2.649	**-0.003**	0.001	0.002	-0.007	**-3.423**	-0.0001	0.001
(1.87)	(0.21)	(1.60)	(1.88)	(0.90)	(1.30)	(1.38)	(2.06)	(1.08)	(0.68)

in the case of $\hat{\mathbf{A}}_1$ and

r_{t-1}^{stock}	r_{t-1}^{bond}	r_{t-1}^{T-bill}	Div. Yield	Short Nom. Rate	Term	Default	CPI Inflation	IP growth	Unempl.
-0.092	**0.111**	-0.314	**0.003**	**0.004**	**0.004**	**0.011**	-1.821	0.000	**-0.003**
(3.36)	(1.76)	(0.16)	(2.24)	(2.35)	(3.18)	(2.72)	(0.96)	(0.47)	(2.26)

as far as $\hat{\mathbf{A}}_2$ is concerned. Clearly, real bond returns are much more predictable, especially in regime 2, which is consistent with Figure 1. Detailed estimates and results for 1-month real T-bill returns are available from the Authors upon request.

[12] Initially, all models estimated are homoskedastic. Section 6 discusses the estimation and portfolio implications of more complex, heteroskedastic models. In this case, we elect to make second moments depend on the same Markov state as the mean parameters as this seems common in the literature (e.g., Kim et al., 1998, Guidolin and Timmermann, 2006).

A few obvious findings stand out without a need for a careful examination of Table 2. First, independently of the specific Markov switching model considered, it is clear that for all values of k, the null of $k = 1$ is always resoundingly rejected with p-values that are basically nil. The evidence in Section 3.2 makes this finding not surprising: even rather poor Markov switching models—in the sense that their driving Markov state fails to be persistent—do fit the data in-sample better than simple VAR models do. The column devoted to Davies-style LR tests shows that low p-values are systematically achieved when testing the number of regimes. Second, when both $k \geq 2$ and $p \geq 1$ it is very easy to build richly parameterized models with hundreds of parameters. Although our data series are sufficiently long to allow to (try and) estimate some of these large-scale models, it is clear that when the saturation ratios decline below 20, one should not put much faith in the resulting estimates, while it is common to find that a stunning fraction of the conditional mean parameters estimated fails to be statistically significant.

When it comes to model selection, Table 2 shows that—as one would expect—BIC selects very parsimonious models, to the point that only one Markov switching model is among the best three models according to BIC, while the other two models are parsimonious VAR models. The three-state model selected by BIC is also rather parsimonious, a MSI model with 21 parameters only (against the 26 typical of the VARs in the Table). On the other hand, the notoriously lax AIC tends to select heavily parameterized multi-state models, ranging from the intermediate-size three-state MSVAR(1) that use principal components as predictors (see Section 3.4 for additional details) to some larger three-state MSVAR(1) that includes all predictors (in practice, a three-state version of (15) from Section 3.2). The H-Q criterion sits in between, although for our data it tends to yield selections that are similar to AIC. However, H-Q agrees with BIC in returning a simple three-state MSI as a framework that efficiently trades-off fit for parsimony. In fact, under BIC such a model turns out to be the one that yields the lowest BIC, -17.15. As a comparison, keeping fixed the simple MSI structure of the model, the BIC takes values of -16.91 for $k = 1$, of -17.01 for $k = 2$, and -17.06 for $k = 4$. The appropriateness of a three-state MSI model is confirmed by Davies (1977)-corrected likelihood ratio tests that take into account nuisance parameter issues in standard LR tests applied to MSH (see Garcia, 1988). We note that in multivariate applications involving US stock and bond returns more than two regimes may be required for a correct modeling of their joint density appears to be common in the literature, see e.g., Guidolin and Timmermann (2006) or Guidolin and Ono (2006).

Table 3, panel B, shows standard QMLE parameter estimates of the three-state model (see Hamilton, 1994, and Guidolin and Ono, 2006, for additional details on estimation and forecasting in a Markov switching framework). Panel A reports single-state estimates as a benchmark. In this application, the single state model is the Gaussian IID benchmark. Intuition for the properties of the model can be easily gained by commenting the parameter estimates within each regime. The first regime is a bear state in which expected real returns are negative (for 1-month nominal bills and long-term bonds) or zero (stocks, in the sense that the bear state mean parameters fail to be statistically significant). The bear state is moderately persistent with an average duration of approximately 4 months; when the US financial markets leave the bear state, this is usually to switch to the intermediate, normal regime. Notice that differently from other papers in the Markov switching literature, the bear state is in no sense an extreme or "rare events" regime, as it characterizes almost 14% of all long samples one could simulate from the estimated MSI, which in our case is almost 8 years of data. The second regime is a normal state with positive, statistically significant but also moderate mean real returns on all assets. This regime is highly persistent with an average duration in excess of 16 months and characterizes more than 80% of any long sample. The third regime is a bull state, in which all assets yield high real returns, even though the dominant asset class in terms of mean real returns is long-term government bonds. Clearly, the data under investigation lead to the specification of this third regime because they need the flexibility to specify heterogeneous dynamics for bond and T-bill real returns during bull regimes vs. normal states. Further checks confirm that the poor performance of simpler, two-state models fitted to our data largely derives from this need to allow for differential dynamics in stock and bond/bill returns. This third regime is also persistent, with an average duration of almost 5 months. Finally, we notice that the estimated transition matrix in Table 2 has a rather special structure, by which regimes 1 and 2 and to some extent 2 and 3 "communicate" on a frequent basis, while regimes 1 and 3 do not, in the sense that from regime 1 it is difficult to switch to regime 3 and vice versa. The fact that the third state has some persistence but in a sense isolated from regime 1 explains why regime 3 has an ergodic probability of less than 6%.

Figure 2 completes our description of the MSI model by plotting the smoothed (full-sample, ex-post) probabilities for each of the three regimes. The figure shows plots which are entirely consistent with the interpretation provided above. The first (bear) state characterizes a non-negligible portion of the data and picks up relatively long-lived episodes that consist of either well-known US recessions as dated by the NBER or of periods of crisis in the US financial markets with declining interest rates and negative realized stock and bond returns (e.g., 1974-1975, 1978-1980, 2001-2002, and more recently most of 2008). The second (normal) state is exceptionally persistent and has in fact characterized long chunks of the recent US financial history, such as most of the 1960s and the great moderation period 1990-1999. Finally, the third (bull) state is characterized by three obvious episodes, which are the long period (1981-1986) of declining inflation and short-term rates in the US after the inflationary bouts of the late 1970s, 2005-2006, and (interestingly) the final months of 2008. These are periods of declining short-term rates and of increasing long-term bond prices that lead – consistently with our characterization of the regime – to high and statistically significant real bond returns.

Figure 3 plots the recursive estimates of the mean coefficients under MSI and helps visualize the key result that the nature (e.g. the interpretation of the regimes) of the three-state model tends to be amazingly stable over time, in spite of our recursive implementation. Although one of the dangers is for MSI to make some sense over the full sample but to produce increasingly awkward results when estimated on much shorter samples (e.g. 1953:01-1973:01, our first estimation sample), these dangers seem not materialize in our application: MSI produces stable mean estimates and—with one minor exception concerning real stock returns for a few months in 1985 (when the mean real stock return was identical in regimes 2 and 3)—the interpretation of the regimes has remained the same we have provided in this section. Of course, stability of the coefficient estimates within a multi-state framework is a good indication of absence of misspecification and bodes well for the forecasting properties and performance of the model in OOS tests.

3.4. *Markov Switching VARs*

Table 2 clearly shows that when MSVAR(p) models are specified using the original predictors, their statistical performance is unsatisfactory. Based on the evidence in Section 3.2, we know that two-state models will perform poorly because our data seem to actually need Markov switching models with 3 or more states. However, any three- or four-state MSVAR that employs any significant number of the original predictors normally ends up to be richly parameterized. In fact, we could not even estimate any three- and four-state MSVAR(p) models with $p \geq 2$ when two or more predictors were included because of insurmountable numerical difficulties. However, Table 2 presents also summary statistics for a further, special class of MSVAR models that—differently from other MSVAR models that appear in the Table—use not the predictors of some of sub-sets of them to form predictions of asset returns, but instead first distill our 7 predictors in a relatively small number of principal components and then augments the Markov switching model to include Q principal components

$$\begin{bmatrix} \mathbf{r}_{t+1} \\ \mathbf{pc}_{t+1}^{(Q)} \end{bmatrix} = \boldsymbol{\mu}_{S_{t+1}} + \sum_{j=1}^{p} \mathbf{A}_{j,S_{t+1}} \begin{bmatrix} \mathbf{r}_{t+1-j} \\ \mathbf{pc}_{t+1-j}^{(Q)} \end{bmatrix} + \boldsymbol{\varepsilon}_{t+1} \qquad \boldsymbol{\varepsilon}_{t+1} \sim IID\ N(\mathbf{0}, \boldsymbol{\Omega}), \tag{16}$$

where $\mathbf{pc}_t^{(Q)}$ is a $Q \times 1$ vector that collects Q principal components extracted from the full set of M predictors $\mathbf{y}_t$, with $Q \leq M$. The intuition for why (16) may represent a useful tool to predict real asset returns is that it is possible that the reason for why either the large-scale VARs (characterized by large p) and especially any MSVAR(p) including many predictors fail to deliver appealing information criteria, is that in any (MS)VAR for a vector of $N + M$ variables, any increase in either k or p determines an enormous increase in the number of parameters that need to be estimated, see e.g., Ludvigson and Ng (2007). By resorting to $Q < N$ principal components to replace the 7 predictors we are entertaining, we aim at shrinking the number of parameters while at the same time minimizing the information loss.[13]

We have applied standard (based on correlation matrix decompositions) principal component (PC) methods to $\mathbf{y}_t$, obtaining that the first three components are able to summarize more than 73% of the total variability of $\mathbf{y}_t$. In particular, the first PC accounts for 34%, the second for 25%, and the third for 14%. To save space we do not report in detail the loadings of each of the first 3 PCs on each of 7 original predictors.[14] However, our task is made simple by the fact that PC1-PC3 have a rather straightforward

[13]There is a growing literature that has argued that in the presence of large sets of predictors, a few principal component may deliver substantial OOS forecastability shielding from the perils of over-parameterizations, see, e.g., Heij, Groenen, and van Dijk (2008), Stock and Watson (2002). Our innovation here consists in proposing and estimating a Markov switching mapping between a small number of factors and the variables to be predicted, in the spirit of Bai and Ng (2008).

[14]Detailed results are available upon request from the Authors.

structure. PC1 loads positively with approximately equal weights on four of the seven predictors, the dividend yield, the nominal 1-month T-bill rate, the default spread, and the unemployment rate. PC2 loads positively and with high coefficients on the term spread and (to a lesser extent) the unemployment rate, while it loads negatively on 1-month T-bills and the inflation rate. Finally PC3 can be basically identified with the IP growth rate. Interestingly, all the seven predictors are reflected by at least one PC, and in fact in only one case (the unemployment rate), a predictor is positively correlated with two different PCs. The nominal short-term rate is also the only predictor that would cause a spread between two different PCs (1 and 2), in the sense that a higher short-term rate will increase PC1 while reducing PC2. The ability of $Q = 3 < N = 7$ to summarize more than 73% of the total variability in $\mathbf{y}_t$ suggests constructing three new PC variables to replace $\mathbf{y}_t$ implies that with less than half of the original number of variables it is possible to capture almost three-quarters of the original information.[15]

The fourth panel of Table 2 confirms that our intuition is correct: the MSVAR(p) models that we build using PC1-PC3 perform considerably better than all MSVAR models that include any sub-set of the original predictors. In fact, two information criteria (AIC and H-Q) indicate that MSVAR(1) models using PC1 and PC2 are quite competitive in terms of trade-off between fit and parsimony.[16] In particular a MSVAR(1) that uses PC1 as its only predictor is the model selected by H-Q over any other competing model in spite of its relatively medium-scale size (76 parameters), which is a remarkable finding. We therefore focus our attention on this MSVAR benchmark in the OOS forecasting and portfolio exercises performed in this paper.

Table 4 reports QMLE estimates of (16) when $Q = 1$ and PC1 is the selected summary of the original predictors. Interestingly, the three regimes carry the same interpretations as the regimes in Table 3. However, the regimes are now considerably more persistent. Regime 1 is a bear state in which real T-bill and stock returns are negative (-0.07 and -1.43 percent per month), while real bond returns are essentially zero.[17] In this regime, PC1 predicts all asset returns with coefficients that are statistically significant. Additional, past real stock returns predict their own future and also subsequent real bond and T-bill returns.[18] Linear predictability is rather pervasive and the associated VAR coefficients are estimated with precision. This regime has an average duration of almost 21 months and it characterizes approximately 22% of any long sample. Regime 2 is a normal state that characterizes almost 59% of the sample because of its extreme persistence. In this regime, unconditional mean real returns are positive for all the assets,

[15]In fact, the simplification is even greater: for instance, a VAR(1) matrix for $Q = 3$-dimensional system contains 9 coefficients vs. 49 in the case of a $N = 7$-dimensional system.

[16]In this case we were also able to estimate a few MSVAR(2) models, especially when only PC1 and PC3 were used as predictors. However, all these models are relatively large and severely penalized by the BIC and H-Q information criteria. The three models appearing in Table 2 are the most promising ones in an ex-ante perspective. We have also compared these models with MSVAR(p) models that employ only one predictor at the time, finding that these are dominated by the PC-based MSVARs. This is to be expected because PCs are able to collect much more information that individual predictors.

[17]These estimates of unconditional mean returns are computed as within-regime numbers, $\hat{E}[\mathbf{r}_t|S = s] = (\mathbf{I} - \hat{\mathbf{A}}_s)^{-1}\hat{\boldsymbol{\mu}}_s$.

[18]In Table 4, the regime-dependent VAR(1) matrix have to be read horizontally. For instance, in state 1 the estimate of -0.0009 illustrates the effect of a change of PC1 at time t on the $t + 1$ value of real T-bill returns.

although they are modest in the case of bonds (0.09% per month, against 0.12 and 1.08% for T-bills and stocks, respectively). Because of its high persistence, when markets enter in regime 2, they stay there for almost 41 months on average. In this state, there is less VAR-type predictability, even though PC1 keeps forecasting both real T-bill and stock returns. Finally, regime 3 is a bull state characterized by positive and high unconditional, within-regime mean returns (0.08 and 1.44% per month in the case of T-bills and stocks, respectively), although the bull characterization is particularly strong in the case of bonds (0.89% per month). Also this regime is persistent, with an average duration of 22 months, so to characterize almost 20% of any long sample. Interestingly, in this regime there is hardly any linear predictability left, with the minor exception of real 1-month T-bill returns being forecastable using past real returns on other assets.

Figure 7 shows the smoothed probabilities computed from the estimates in Table 4. Clearly, the considerable regime persistence uncovered from the MSVAR(1)-PC1 model yields a low number of state switches as identified by the smoothed probability series. The bear regime characterizes a number of periods of financial crisis (such as late 1987, or the Summer of 1998) and economic recession (such as 1961, 1973-1974, 2001-2008). The only surprising finding is that most of the recent 2001-2008 period would be characterized as a bear period. However, that was also the case of Figure 2, where the smoothed probabilities for 2002-2008 strikingly resemble those from 1973-1980. US financial markets have historically been most of the time in the normal state, with some long spells that have stretched for almost a decade without interruptions (the last long spell was the 1989-1998 great moderation period). Finally, the Figure shows two bull periods, 1969-1972 and 1979-1986. The final months of 2008 would have been characterized by a strong bull rebound to the long crisis of 2001-2008. Interestingly, Figure 7 appears to be a less jagged, smoother version of Figure 2 that conveys the same basic regime classification. In fact the correlations between smoothed probabilities series of the MSI model in Section 3.2 and the MSVAR(1)-PC model in this section are all positive and statistically significant (ranging from 0.42 to 0.50).

3.5. *Some Evidence on Forecasting Accuracy*

Before proceeding to the recursive computation of optimal SAA weights and of the resulting portfolio performance, it is prudent to examine the forecasting performance of the models using traditional criteria (such as recursive RMSFE). The rationale for this brief diversion is two-fold. First, this is an important preliminary check because it would be wasteful to engage in extensive portfolio calculations opposing a family of VARs to a MSI/MSVAR reference model whenever the latter represents a poor econometric framework unable to produce accurate forecasts. Although, the issue of the performance of Markov switching models in forecasting applications is a much debated one with conclusions that seem to depend on the specific applications (see Guidolin et al., 2009 for a number of examples), one cannot rule out a priori that in spite of its excellent in-sample fit to our SAA data, MSI and MSVAR may fail to be serious competitors in applications that rely on its predictive performance. Second, because we shall adopt a criterion – such as portfolio choice with continuous rebalancing under power utility – that hardly relies only (or even mostly) on point forecasts, a possibility exists that even though MSI and/or MSVAR under-perform the VARs as a forecasting device for the mean, they may represent a useful engine for portfolio choice because it may forecast either higher-order moments (e.g., skewness and kurtosis, besides the mean, as in Guidolin and Timmermann, 2008a) or the entire joint density of real asset returns, which is the object of interest of the portfolio problem introduced in Section 2.3.

Although in practice we have computed recursive forecasts and assessed overall predictive performances for horizons of 1-, 12-, and 60-months and (in the case of 12- and 60-month horizons) we have extended these calculations to the case in which the object of the forecast is not the real return r^j_{t+H} but the cumulative real return, $R^j_{t,H} \equiv \prod_{h=1}^{H}(1 + r^j_{t+h}) - 1$, Tables 5 and 6 only report forecasting performances for the case of $h = 12$ months.[19] Results for the $H = 60$ horizon were qualitatively similar.[20] In the tables, we have listed and reported the forecasting performance for the best 10 forecasting models (among all the VARs we have experimented with, the no-predictability Gaussian IID benchmark that forecasts using a simple recursive sample mean, and of course MSI and MSVAR(1)-PC1) in an overall sense, i.e., scoring all models for their performance in predicting stocks, bond, and T-bill returns.[21] We also report a few additional benchmarks, such as the best performing rolling window VAR, the best performing large-scale VAR (defined as $p \geq 4$, both rolling and expanding), and of course MSI and MSVAR(1)-PC1. Table 5 shows that at least in our SAA application, MSI represents a serious option to any investor interested in 12-month ahead forecasting performance: MSI has the second lowest RMSFE among all models as far as stock and bond returns are concerned, and the best RMSFE in the case of T-bills. The finding for MAFE is similar, apart from the fact that MSI yields now the lowest MAFE for both bonds and T-bills. In the case of stocks and bonds, the lowest RMSFE is instead guaranteed by the MSVAR(1)-PC1 model. In fact, the RMSFE improvement of MSVAR(1)-PC1 over MSI appears massive in the case of stocks, in the order of 40%. However, MSVAR(1)-PC1 performs poorly when it comes to forecast 1-month real T-bill returns, and this responsible for the overall mediocre ranking of the model. MSI provides substantial improvements in RMSFE when compared to linear models, in the sense that its RMSFE is between 14% and 16% lower than the best performing VAR for all asset classes. These improvements come from the uniform ability of MSVAR(1)-PC1 and MSI to reduce the sample standard deviation of forecast errors, while slightly

[19] The choice of a particular forecast horizon is difficult. On the one hand, most of the forecasting literature naturally focusses on the $H = 1$ case, which is however irrelevant for long-horizon portfolio optimizers. On the other hand, even if our goal is to assess portfolio performance at $H = 60$, such a long horizon appears odd in the forecasting literature and implies a severe loss of data. The choice of $H = 12$ in Tables 5-6 is a trade-off between these two considerations.

[20] These are available upon request from the Authors. The results for $H = 1$ month are different, in the sense that MSI fails to be among the best forecasting models. This is interesting because it confirms that when the predictive exercise is performed in ways that differ from Guidolin et al. (2009) (they focus on simple MSI and MSIH predictive univariate regressions), then some results typical of the earlier literature may be still be found. However, for $H = 12$ months, our exercise confirms Guidolin et al.'s findings on US data.

[21] We provide for each model three scores, one per asset, which equals the rankings of the model across all assets (e.g., 1 to the best model, 2 to the second best, etc.). For instance, the best VAR in Table 5 receives scores/ranks of 36, 9, and 39 on forecasting real T-bill, bond, and stock returns, which indicates that it is not particularly accurate for any of the assets, but very robust throughout. The overall rank is based on the sum of these scores, with the best models receiving the lowest total score (3 is the minimum and 990 is the maximum).

better models can usually be found in terms of minimization of the overall (absolute) bias. The last row of Table 5 stresses that a few differences exist between in-sample results on which predictor coefficients are often statistically significant, and what actually pays out in reducing RMSFE in OOS experiments: the term spread and the rate of growth IP and the unemployment rate are the predictors that enter the best performing VARs. Below the tenth position in the ranking, it is clear that rolling window VARs and large-scale VARs all have a hard time providing accurate forecasts.

Results in Table 6 on prediction of cumulative returns are still largely favorable to MSI (which is still ranked as the best predicting model), but are more articulate. As far as cumulative real stock returns are concerned, MSVAR(1)-PC1 remains the best model in terms of both RMSFE and MAFE; the good performance is the result of a low sample standard deviation of forecast errors. MSI is the second best model and it still represents a discrete improvement over the best VAR models. However, the best "cumulative" predictors for real bond and 1-month T-bill returns are VAR models that actually cannot predict real stock returns and that as such are heavily penalized by our overall ranking system, ending up with an overall rank of 59 and 226, respectively. Furthermore, while MSI and MSVAR(1)-PC1 are much worse than the best predicting VAR for real bond returns (their RMSFE are only 4-5% higher than the best VARs), MSI and especially MSVAR(1)-PC1 have big problems at predicting real 1-month T-bill returns (e.g., the RMSFE of MSI is a full 30% higher than the RMSFE of the best performing VAR). The variables that work in making VARs good predictors are the same as in Table 5, although in the case of cumulative returns the unemployment rate seems to be less important and some role is now played by the dividend yield. All in all, the evidence in Table 6 is also indicative that it remains possible for a relatively large set of VAR models to encounter difficulties at producing similar forecasts to (hence, portfolio weights) and better realized SAA performance than Markov switching models, which justifies the rest of our investigation.

4. **Optimal Strategic Asset Allocation and Hedging Demands**

4.1. *Recursive Portfolio Weights*

Figure 8 plots and compares recursive optimal portfolio weights (for $T = 1$ month and 5 years) for two models, MSI and a VAR(1) in which all predictors are included to maximize its overall forecasting power. The left hand plots also report optimal weights under the Gaussian IID (no predictability) benchmark. The recursive exercise is performed on an expanding window over the period 1973:01 - 2008:12, as planned, therefore also including the deep financial crisis of 2008. These weights are computed under the assumption of $\gamma = 5$. Clearly, while VAR(1) implies rich and persistent dynamics in optimal portfolio weights for both short- and long-run horizons, the variability of asset allocations is likewise strong and interesting under MSI, as one would expect given the fact that this models actively draws inference from the nature of the current regime and forecasts H-step ahead market states. In fact, in the case of MSI, asset demands often "jump", reflecting possible switches in the perception of the current regime and—as a result—in the forecasts of future market states. In particular, both linear and non-linear predictability patterns induce strong time variation in optimal weights for a long-horizon (5-year) investor. Here, it is evident that while under MSI the differences between short- and long-run portfolios exist but are generally modest (which means that hedging demands are small, see below), under VAR(1) the opposite occurs: VAR-type linear predictability induces large and persistent differences between optimal decisions by short-horizon investors vs. long-horizon ones.

Even though MSI induces high, regime-linked variations in optimal weights, there are some general trends in portfolio weights that appear both in the left- and right-hand columns of Figure 8. For instance, the optimal demand stocks tends to be non-negative most of the time under both models, with the exception of the period 1977-1981 which—at least in qualitative terms—appears in both models. Similarly, there is a common peak in the demand for 1-month T-bills in correspondence of the same period. In any event, the plots are easier to use to comment on the substantial differences between optimal weights under MSI and VAR(1): as one would expect, the dynamics are rather different in the two cases and it is evident that even a medium-scale VAR(1) model cannot produce the rich, regime-like dynamics in SAA that a MSI model naturally implies. For instance, while MSI implies average weights to stocks that are high by historical norms (around 110%) between 1992 and 1998 this fails to occur under a VAR which for these periods implies instead weights that are either close to unconditional means or actually below such a historical norm. Finally, Figure 8 also offers the first chance to comment two issues briefly touched upon in the Introduction. First, it is clear that while starting in the early 1980s a MSI implies an average demand for stocks that oscillates around a small positive percentage commitment, VAR produces generally high weights that for a long-horizon investor are never below 100% after the late 1980s. Many papers in the empirical SAA literature have complained that this latter implication (for a sensible coefficient of risk aversion such as $\gamma = 5$) seems hardly plausible. Second, the figure shows that while MSI implies a demand for long-term bonds that is generally positive (even though modest and with occasional negative spikes) for both short- and long-run investors, a VAR has odd and counter-factual (i.e. inconsistent with equilibrium) implications by which the demand for bonds ought to be strongly trending but also be characterized by an embarrassingly negative average for short horizons throughout the 1980s and 1990s.

Figure 9 has a structure identical to Figure 8 but compares the weights of MSVAR-PC1 with those characterizing the best performing OOS expanding window VAR (see Section 1), a simple VAR(1) in which there is only predictor, the dividend yield (henceforth called VAR-DY). This creates the impression of higher volatility of asset demands under MSVAR-PC1, which is however not completely correct if we take into account the differences in scales between the plots in Figures 8 and 9 (see also Table 7). Clearly, MSVAR-PC1 implies weights that combine the properties (regime switching-like variation) of MSI with the typical, high frequency persistent dynamics of VAR models. The figure also highlights that a simpler VAR-DY produces uniformly positive (negative) and large hedging demands for stocks (1-month T-bills), while this is not the case for MSVAR-PC1.

Table 7 translates these visual impressions for the case $\gamma = 5$ into summary statistics for our overall sample period. The table reports three types of summary statistics: the mean of recursive portfolio weights, their sample standard deviation, and their 90% empirical range, i.e., the values of the weights that leave 5% of the recursive weights in each of the two tails. The latter measure is offered to avoid undue reliance on sample standard deviations as measures of dispersion when the weights have distributions which are non-normal. These statistics are computed and presented for the MSI and MSVAR-PC1 models, the Gaussian IID benchmark, and a variety of VAR models that are selected in consideration of their pseudo-out-of sample performance at a 60-month horizon in terms of CERs (see Section 6).[22] Table 7 illustrates the existence of major differences across the three types of models (Markov switching, VAR, and no predictability benchmark) according to all types of summary statistics. Interestingly, MSI and MSVAR-PC1 give qualitatively similar outcomes, especially in terms of mean allocations. In the case of recursive mean weights, the differences concern only long-term bonds and 1-month T-bills: while the Gaussian IID benchmark a relatively low demand for long-term bonds (16%) and the MS models an intermediate-level demand (29-40%), the VARs imply rather heterogeneous demands that go from levels of 30% below the typical Markov switching allocations to means in excess of 100% which are typical of rolling window VAR models, where the 10-year scheme occasionally brings to a perception of very high Sharpe ratios, like in the mid-1980s and recently the 2001-2008 period. Similarly, while both the MS models and the IID benchmark deliver on average positive and modest demands for 1-month T-bills (between 9 and 18%), most VAR models make it optimal to actually leverage the portfolios by borrowing at the 1-month real T-bill rate.[23] Finally, although the finding does not concern all the VAR models we have entertained, we notice that a majority of VARs do imply a higher demand for stocks than MS models and the no-predictability benchmark do, say between 80 and 100% on average vs. average allocations between 50 and 70% in the case of MS and IID strategies. This finding echoes the common complaint (see Ang et al., 2005) that asset allocation models calibrated to standard preferences and linear predictability models easily generate "too high" a demand for stocks. Clearly, this is not the case under Markov switching, non-linear predictability.

Table 7 also reports sample measures of dispersion of recursive portfolio weights. Here the finding is clear: given its structure, MSI and MSVAR-PC1 deliver weights which display approximately only half the weight volatility that is typical of VARs. The volatility of portfolio weights of MSVAR-PC1 and MSI are also rather similar, which may be taken as indication that the variability in portfolio decisions will mostly originate from regime switching and not from the linear predictability that is captured by the cross-serial correlations between real asset returns and the first principal component. These findings also apply to the 90% empirical range of optimal weights.[24] These results show that the widespread belief that regime

[22]Table 7 only concerns optimal weights computed for the case of $\gamma = 5$. The results for $\gamma = 2$ and 10 are qualitatively similar. These additional tables are available upon request from the Authors.

[23]The weights mentioned in the main text are the 1-month optimal weights, since this allows a three-way comparison involving the Gaussian IID results. However, most VARs imply a long-run demand for stocks that largely exceeds the 1-month weight and a long-run demands for 1-month T-bills that are negative and large. Hedging demands for long-term bonds tend to be negative but also modest.

[24]As one should expect, the recursive Gaussian IID weights are always the least volatile for all assets and according to all

switching asset allocation frameworks may imply "excessively" volatile portfolio weights may be misleading when applied to long-run SAA under rebalancing.

4.2. *Hedging Demands*

Figure 10 shows the recursive hedging demands for the period 1973:01-2008:12 implied by Figures 8 and 9, for the four competing models covered by these figures. Also in this case, we need to take the results from the "full" VAR(1) (in which all predictors appear) and VAR-DY as representative of the type of hedging demands that may be typically obtained under linear predictability. The VAR hedging demands are not severely affected by the details of the linear framework used: the hedging demand for stocks is large (in excess of 50% over the entire 1973-2008 sample period) and stable, consistent with results reported by Barberis (2000) and Campbell et al. (2003) among the others. On the contrary, hedging demands for 1-month T-bills and long-term bonds contain massive drifts and are considerably volatile in the case of the full VAR(1) model. The negative VAR hedging demand for T-bills under the VARs is at first trending down, for instance falling below -100% in the case of VAR-DY, and then trends up, settling to a negative level between -20 and -60%. The VAR hedging demands for long-term bonds are instead quite different across the "full" and DY models, with a lot of variation in the former case and none in the latter.

Markov switching hedging demands are completely different, in at least two ways. First, they are generally very small when compared to VAR hedging demands. This is consistent with the findings in Ang and Bekaert (2002) and Guidolin and Timmermann (2007) with reference to international portfolio diversification and SAA, respectively. Second, the MS hedging demands are also obviously stationary over time and tend to simply fluctuate around zero. Third, MSI and MSVAR-PC1 hedging demands are qualitatively similar. However, this does not imply that MS hedging demands are zero such that MS non-linear predictability is irrelevant: stocks generally command a positive hedging demand with spikes up to 80%, while long-term bonds usually imply negative but modest hedging demands. These differences between MS and VAR hedging demands are made more explicit in Table 7.[25] For instance, MSI delivers a positive, +3% hedging demand stocks, i.e., the presence of Markov regimes ends up making an investor less cautious in the long-run than in the short-run, which skews her demand towards stocks; the MSI hedging demand for T-bills is also positive on average (+6%) and is negative for long-term bonds (-9%). Interestingly, different VAR models may imply differences in average hedging demands for long-term bonds (although these are generally modest) and especially 1-month T-bills, although it is clear that the average hedging demand for stocks is always positive and often large (also in excess of 100%). Additionally, while the variability of hedging demands is always modest for the Markov switching models, there is much more heterogeneity over time for hedging demands under linear predictability.

The results in Figures 8-10 and Table 7 provide compelling evidence that a model that accounts for

measures.

[25]By construction, the Gaussian IID benchmark implies zero hedging demands in the presence of continuous rebalancing, see Samuelson (1969).

regimes in financial markets delivers recursive, optimal SAA weights that – both in terms of average weights and of their dispersion over time – cannot be approximated by any of the VAR models we have experimented with. The differences are particularly striking for what concerns the level and variability of optimal stock weights and in terms of the implied hedging demands that ought to protect a $\gamma = 5$ investor from stochastic changes in investment opportunities. Since we have experimented with a rather large and encompassing range of VAR models typical of what is commonly found in the empirical literature, this is *prima facie* evidence that simple linear predictability framework may be unable to capture all modes of predictability commonly found in the data, including those summarized by regime switching dynamics.

5. **Realized Recursive Portfolio Performance**

Our finding in Section 4 that VAR models typically produce dynamic (short- and long-run) SAA weights and hedging demands that depart from the implications of a model that accounts for non-linear patterns is suggestive that naive linear frameworks may be too simple to pick up and exploit predictability patterns that are in the data and that may be important in applied portfolio applications. However, these results are suggestive at best: because a model that fits the data better in-sample than another model does not have to out-perform the latter in OOS experiments, a portfolio manager will always want to examine evidence on the recursive, OOS performance of both models before selecting one or – as we aim at – conclude that either of them may be "too simplistic" to be useful. This is exactly what we set out to do in this section: use the recursive experiment outlined in Section 2.3 to assess whether VAR models can yield realized OOS performance that is equivalent (or even superior) to MSH models. In particular, Section 5.1 presents the overall OOS portfolio performance results for the complete set of models examined in Sections 3 and 4. Section 5.2, proceeds to a conceptually tighter and better defined "horse race" between *classes* of models, that allows us to oppose the set of all VARs to the two different Markov switching frameworks—MSI and MSVAR-PC1—that we have developed and estimated.

5.1. *Overall Performance*

Before proceeding further and examine the results of recursive portfolio experiments, it is necessary to briefly discuss two issues with our research design. First, one wonders whether it is sensible to expect that one single (albeit carefully selected, in accordance to the literature) regime switching "champion" may outperform the full set of 896 VARs we have opposed it to. Although there is no unique, compelling answer to this question, two considerations are relevant. In the light of the main research question of this paper, one is tempted to reply that yes: one model out of the set MSI, MSVAR-PC1 ought in principle to be the best performing among all models. The existence of even a few models that might out-perform both the MS frameworks would imply that at least *some* (even if few) VARs could deliver portfolio choices similar or better than MS, which must be a result of the fact that these VARs will be obviously able to capture regime dynamics (or the portion of it that ought to matter for SAA decisions). However, even though both MSI and MSVAR-PC1 were selected as a result of careful model specification search, it cannot be claimed that the rich and ever growing family of Markov switching model for asset returns can be completely represented and summarized by either MSI/MSVAR-PC1 or by the set of models appearing in Table 2. Therefore one may also consider in a light unfavorable to VARs the finding that MSI/MSVAR-PC1 may out-perform a large portion (say 95 or even 99%) of the VARs we have experimented with, according to the idea that if VARs can adequately summarize regime-type dynamics in financial markets, then most of them ought to be able to perform the task, independently of their fine-tuning. In this case – because $(1-0.95)\times 896 \simeq 45$ and $(1-0.99)\times 896 \simeq 9$ – we should find that either MSI or MSVAR-PC1 or both is a "top 50" or even a "top 10" model among all the ones we have tried in our experiments. Second, it must be stressed that even though in what follows we present realized portfolio performances for both 1-month and 5-year horizons, it is sensible to think that the latter sets of results should carry more importance than the former as our stated goal has been to test whether VARs can approximate the performance of models with regimes in the perspective of long-horizon investors. Armed with these considerations, we proceed to present and comment empirical results.

Table 8 reports the key results of the paper.[26] For the case of $\gamma = 5$, to save space we report the best 7 performing models (plus benchmarks, when these are not among the top 10) when all models recursively estimated are ranked according to their real CER. The top panel concerns the 60-month horizon, while the lower panel the 1-month horizon. The models that we report below the 7th position in the CER ranking are selected because they are either benchmarks or representative of wider classes of models, i.e., MSI or MSVAR-PC1 should they fail to be among the top 7 models, the Gaussian IID benchmark, the best performing rolling window VAR (this claim reflects the finding that in general expanding window VARs outperform rolling window VARs), and at least one non-small scale VAR (we class small scale VARs as all those with $p = 1$ or those with $p = 2$, 4 with only one predictor). In the view of a long-horizon ($T = 60$ months) investor, MSI ranks first out of all the models with an annualized CER of 8%; the attached 95% confidence interval is relatively tight, [4.9%, 8.7%], which means that it is likely that a $\gamma = 5$ investor would be ready to pay at least an annualized real, constant return of almost 5% to have access to SAA decisions using the MSI model. This means that our set of VARs fails to include any models that produce CERs which exceed the CER of MSI. In particular, a rather simple VAR that includes only lagged real asset returns and the dividend yield produces a lower CER of 3.7% with a bootstrapped 95% confidence interval of [-3.9%, 11.6%]. However, it is clear that the two confidence intervals for MSI and the best VAR do overlap, which may be taken an indication that there is no strong statistical evidence against the null hypothesis that the two models may give identical CER performance. Interestingly, the richer MSVAR-PC1 model severely underperforms both MSI and the 5 VARs that appear in the top panel of Table 8. Its CER rank is 33, which still places MSVAR-PC1 among the best 5% of all the models we have experimented with

[26] In Tables 8 and 9, the reported 95% confidence bands have been computed by applying a block bootstrap to each of the recursive, realized performance statistics.

in this paper. However, the CER of MSVAR-PC1 is largely disappointing, 2.4% and with a 95% confidence interval ([1.5%, 5.1%] that even its upper bound is at best comparable to lower bounds from the CER of MSI and the best performing VARs.[27] Table 8 also shows the median performance statistics for all 896 VAR models entertained in our paper, distinguishing between expanding and rolling window VARs. The former class performs slightly better than the second, but it is striking to notice that both median real CER measures from VAR models are negative, an indication that a $\gamma = 5$ investor would required *to be paid* in order to accept to perform her SAA using the median VAR model, both in the expanding and in the rolling-window implementations. In other words, blindly exploiting linear predictability (as captured by median performance) incredibly leads to results that are inferior not only to ignoring predictability of all kinds, but also to a passive 100% investment in an asset that gives a constant zero real rate (i.e. which simply protects against inflation dynamics).[28] Clearly, MSI performs considerably better than the median, representative VAR SAA strategy. Interestingly, the no predictability benchmark turns out to be a serious candidate in a long-horizon portfolio perspective, yielding an attractive real CER of 5.5%, which is however inferior to the 8% that can be accessed exploiting a non-linear portfolio strategy.[29]

There is clear structure in the VARs that deliver good portfolio performance: these are very parsimonious models with few lags ($p = 2$ at most, but the majority of the top 20 performers are $p = 1$) and in which only four predictors appear in a variety of combinations: the dividend yield, the default spread, IP growth, and the unemployment rate. Between the possible dimensions of parsimony in our experiment (p vs. choice of M), the latter is more important than the former, in the sense that $p = 2$ sometimes yields interesting performance, but always under the condition that very few predictors are included.

There are also some notable differences in the way in which good realized real CERs are obtained across models, and especially from MS vs. VARs. In particular, MSI gives a lower mean than all other top-performing VARs (e.g., an annualized real 11.1% vs. 22.2% per annum for the best VAR) but also a sensibly lower volatility (e.g. 21.2% per year vs. 53.7% for the best performing VAR). These differences translate in the fact that MSI in fact yields a very appealing Sharpe ratio (0.46 in annualized terms vs. 0.39 for the best VAR), which is second only to the Sharpe ratio for the no predictability benchmark (0.57) but typically much higher than the typical (median) Sharpe ratio among all VAR models (0.10 at best). How is it possible that MSI implies a higher realized OOS CER than the Gaussian IID model does, even though the latter model is characterized by a higher Sharpe ratio? Here we need to notice that especially with a

[27]That MSVAR models may disappoint in recursive OOS portfolio experiments fails to come as a complete surprise. Guidolin and Timmermann (2007) report suggestive evidence that MSVARH models that include the dividend yield do not always outperform simpler MSIH models.

[28]Such inflation-indexed assets exist, at least as a first approximation (e.g. TIPS) and a zero real return can be reasonably taken to be their lower bound for realized real returns. Table 8 reports median performances and not mean performances because of the presence of a few obviously bad models (in general, these are the $p = 12$ models) that produce either negative mean portfolio returns and/or high volatility and therefore largely skew the distribution of portfolio performances.

[29]The Gaussian IID is characterized by a 95% confidence interval of [4.9%, 6.2%] which implies the existence of an overlap with the intervals for the best VARs and MSI. However, the confidence interval for the no predictability benchmark fails to include the CER for MSI. This can be taken as evidence that ignoring predictability would be harmful to long-horizon investors. This is consistent with the bulk of the literature on SAA under predictability, e.g. Barberis (2000) and Lynch (2001).

long-horizon, a power utility investor is different from a mean-variance investor who simply maximizes her portfolio Sharpe ratio. Equivalently, it is well known (see Campbell and Viceira, 2002) that classical mean-variance preferences fail to provide a good approximation to constant relative risk aversion preferences for long-investment horizons, i.e. that isolastic preferences are not locally mean-variance for large T. What can then account for the difference between the Sharpe ratio and the CER-based rankings? The difference must be represented by the role of higher-order moments (skewness, kurtosis, etc. of realized consumption flows financed by the investment strategy), for which a power utility investor cares over and above caring for the mean and the variance. In fact, Table 8 shows that while MSI has positive skewness that is rather close to the asymmetry exhibited by the best VARs, MSI also has the minimal kurtosis among all predictability models investigated. Because excessive kurtosis (i.e., fat tails) in realized portfolio returns hurts a power utility investor, the implication is that MSI is rewarded by a relatively high CER not because of stability per se, but mostly because MSI is a way for a long-run investor to make sure that no excessively poor performances falling in the extreme left tail are obtained.[30] Additionally, the Gaussian IID model displays thin tails which are a positive attribute to a power utility investor, but is also characterized by a rather symmetric distribution of final long-run wealth cumulants, which is inferior to the large and substantial positive skewness coefficients found under MSI.

The lower panel of Table 8 reports on model performance for the best 7 models when the horizon is short. Although this is admittedly less interesting for our paper, here MSI comes in second in the ranking, with a moderate real CER of 5.9% per annum. The most interesting implication of the table is however rather tangential to our main point: in the case of $T = 1$ month, the best realized recursive performance is obtained when all predictability patterns (linear and non-linear) are simply ignored and short-term SAA is implemented using a no predictability benchmark with constant means, variances, and covariances. The Gaussian IID real annualized CER is 6.3% and it is in fact the only CER whose bootstrapped 95% confidence interval ([2.4%, 10.3%]) fails to include zero or values close to zero which an investor can easily purchase in the financial market by simply buying inflation-protected securities. We can summarize this finding as follows: a short-term $\gamma = 5$ should rather ignore predictability than try to use it for portfolio choice; however, conditional on her decision to choose portfolio weights using any predictability patterns, then VARs can neither approximate the portfolio weights computed under MSI nor obtain a comparable recursive OOS performance. It is of some interest to also stress that the acceptable real CER performance of MSI is now generated by properties of portfolio returns which are quite different from those commented for the $T = 60$ months case. Now MSI gives the second best annualized volatility (13.1%), although its realized mean performance remains lower than most VARs (7.6% against median VAR performances of 16-18% per annum). This delivers a MSI Sharpe ratio that is now the highest achievable Sharpe ratio

[30] This claim relies on a difference between variance (the second moment) and kurtosis (the fourth moment scaled by the second), which may be used to illustrate tail thickness above what is allowed under a normal distribution. In general, VAR models tend to produce appreciable Sharpe ratios and positive skewness, but also high excess kurtosis in performance, which means that a VAR model may occasionally "betray" and produce large, negative performance outliers in the left tail which will be wealth-destructive for a long-run investor.

(0.48). Although the results in the lower panel of Table 8 strengthen our earlier conclusion that it is hard for VARs to compete with models that take into account regimes, we leave for future research the task of exploring why ignoring predictability may actually lead to better 1-month recursive performance than in the case predictability is taken into account. Finally, MSVAR-PC1 yields another rather disappointing CER performance of 1.9% with a very wide confidence interval roughly centered around zero.

Table 9 expands the range of portfolio performance results by presenting panels with structure and contents similar to Table 8, but concerning now the cases of $\gamma = 2$ and 10. This addresses the potential concern that our earlier results may be driven by a special (even though, rather typical) assumption on the coefficient of relative risk aversion. In the case of a low risk aversion long-horizon investor, MSI is ranked first on the basis of the annualized real CER (10.2% vs. 4.8% for the best VAR). However, once more the bootstrapped 95% confidence band for MSI ([4.9%, 15.2%]) largely overlaps with the real CER confidence band for the best VARs (e.g., [-2.5%, 12.4%] for the best performing VAR), so that it is hard to actually distinguish MSI from the top 5 models.[31] In this case, MSI has the best Sharpe ratio among all models (0.87), which indicates that for a low risk aversion investor, MSI performs well both in a mean-variance space and in a power utility space in which all moments matter. In the case of low risk aversion, also MSVAR-PC1 becomes a rather competitive model and it comes in third in the CER ranking, however after the Gaussian IID benchmark. However, the CER of MSVAR-PC1 is still only two-third the CER of the more parsimonious MSI, 6.4%. Many other comments expressed with reference to Table 8 apply also in this case. For instance, the best performing VARs are relatively parsimonious models. The lower panel of Table 9 deals instead with the case of a high risk-aversion investor with $\gamma = 10$. In this case, MSI is again the best performing model for long-horizon SAA purposes, with an annualized real CER of 6.2%. Once more, the no predictability benchmark is a serious competitor for a long-horizon investor, with a 4.7% real CER. None of these results obtain for $\gamma = 10$ and a 1-month horizon, where one VAR actually proves useful and better than both MSI and MSVAR-PC1 (even though the bootstrapped confidence interval of the latter remains wide enough to include top CER performances).

5.2. *A Horse Race Between Classes of Models*

While Tables 8-9 highlight the best performing models based on the CER ranking, they suffer from the fact that while the MS performances always appear in the tables, by design, the VAR models covered changes as the parameters of the exercise, specifically γ and H, change across the various panels. Therefore, it would be useful to have a more compact way to summarize and compare the recursive OOS portfolio performances not of each specific econometric model against all other models, but instead in terms of some large macro-classes of models—i.e. all the expanding-window VARs, all the 10-year rolling window VARs, the two Markov switching models, and the Gaussian IID benchmark. Such an experiment actually offers

[31] On the contrary, to tell MSI apart from the median expanding and rolling-windown VARs is easy, as these generates disappointing -0.5% and -0.7% annualized real CERs, respectively.

one additional advantage that does not purely relate to the presentation of the results but that is instead linked to an interesting economic intuition. Consider an investor that is actually contemplating resorting to a VAR modeling strategy to support her long-horizon SAA decisions. It is very unlikely that this investor will actually decide to specify and estimation one particular VAR model and to stick to it over time. Yet this is what our performance assessment in Section 6.1 has assumed. Instead, an investor is likely to use statistical criteria to judge of the likely performance of competing VAR models at each point in time, with the possibility to occasionally switch among different VARs in case this statistical measure of likely OOS performance happens to deteriorate. One may say that such an investor would resort to switching among different VARs instead of building a model of (Markov) switching VAR dynamics as we have done with MSVAR-PC1 in Section 3.4.

As shown by the work by Pesaran and Timmermann (1995, 2000) on switching algorithms to exploit predictability through simple trading strategies, there are a number of ex-ante statistical criteria that an investor may use to determine how and when she would switch from a VAR model to a different one. In this section, we have decided to keep the task simple and endow our VAR investor with the ability to recursively track over time the value of two information criteria already discussed with reference to Table 2, the AIC and the BIC. AIC and BIC are selected over H-Q because the latter is known to generally return indications that are halfway (in terms of parsimony of the selected models) between AIC and BIC. In a sense, we believe that AIC and BIC may span the set of all possible choices. AIC and BIC are selected over in-sample criteria, such as the R-square and the adjusted R-square, because information criteria have been often described as tools to preview the predicting performance of models. Our strategy is as follows. Within the recursive scheme already illustrated in Section 2.3, at each point in time t we model the investor as deciding on which VAR model should be used for her asset allocation decisions between t and $t+H$ based on either AIC or BIC. In fact, to derive distinct evidence on the class-level performance of expanding and rolling-window VARs, we have modeled two different investors, the first selecting among VARs estimated on an expanding window and the second focussing instead on rolling window VARs. Finally, to favor comparability, we have applied an identical logic to the Markov switching class, even though in this case our investor is actually selecting at each point in time between MSI and MSVAR-PC1 only.

Tables 10 and 11 report the recursive OOS portfolio performance following the same structure as Tables 8 and 9. However, by construction, Tables 10 and 11 only feature 7 competing strategies: 1) Switching expanding window VAR set, when the selection criterion is AIC; 2) Switching rolling window VAR set, when the selection criterion is AIC; 3) Switching expanding window VAR set, when the selection criterion is BIC; 4) Switching rolling window VAR set, when the selection criterion is BIC; 5) Markov switching set, when the selection criterion is AIC; 6) Markov switching set, when the selection criterion is BIC; 7) the Gaussian IID benchmark. Before commenting on the OOS performance, let us provide some information on the nature of the switches of the models selected under strategies 1)-6) above. Under 1) the investor would use 64% of the time a model in which real asset returns are predicted using 12 lags of the returns themselves (this strategy produces the lowest AIC for a stunning period of almost 23 consecutive years, between 1985 and 2007), 14% of the time a VAR(12) in which the term spread is the only predictor, and 10% of the time a VAR(12) in which all predictors are included. The remaining 12% of the time is spent using VAR(2) and VAR(12) models with few predictors and with the term spread often entering the mix of predictors. Under 2) the investor would use 49% of the time a model in which real asset returns are predicted using 12 lags of the returns themselves (once more the strategy dominates for long periods, for instance 1978-1995), 21% of the time a VAR(2) in which the term spread is the only predictor, and 18% of the time VAR(2) models in which the term spread is always included (in half of the cases with the short-term nominal bill rate, in the other half with CPI inflation) as predictor. The remaining 12% of the time is spent using VAR(2) and VAR(12) models with few predictors and with the term spread often entering the mix of predictors.[32] The structure of the strategies 5) and 6) is easy to describe. Under a AIC criterion, MSVAR-PC1 is selected 95% of the time. MSI is selected only between 1973 and 1974 and then again in sporadically during the 1980s. Under a BIC criterion MSI is always selected with only 11 exceptions, which occur randomly over our sample (but 3 times during the turbulent 2008).

The results in Tables 10 and 11 completely agree with those already commented in Tables 8 and 9, but are obviously easier to interpret because each row represents now a feasible as well as sensible portfolio strategy based on classes of models, as defined by their econometric structure and whether they are estimated on rolling vs. expanding data sets. For an investor with $\gamma = 5$ and a long, 5-year horizon, the best "class strategy" is based on MS models when these are recursively selected by BIC minimization, which effectively means MSI most of the time. In fact, MS-BIC is ranked first in Table 10 with performance statistics that are very close to MSI in Table 8, for instance the CER is 8.1% vs. 8.0% for MSI. Similarly, the no predictability Gaussian IID benchmark is second in the ranking with a CER of 5.5%, which is by construction identical to the one in Table 8. Interestingly, ignoring predictability implies a CER higher than the CER of MS-AIC (2.6%), which is a mixture of MSI and MSVAR-PC1 tilted in favor of the latter model. However, the key result in Table 10 is the overwhelming evidence that MS models, however they may be recursively selected, outperform the four VAR-class strategies. For instance, the best VAR-class strategy (expanding window, with BIC selection) yields a disappointing real CER of 0.3% with a wide bootstrapped confidence interval that includes negative real CERs. Similarly to Table 8, the strong performance of MS-BIC is not only (or even mostly) the result of a good performance in a simple mean-variance (Sharpe ratio) space, as the Gaussian IID model yields a somewhat higher Sharpe ratio (0.57 vs. 0.45) and yet a lower CER caused by the superior skewness properties of MS-BIC.

[32] Under 3) the investor would use 85% of the time a VAR(4) model that includes all predictors, 6% of the time a VAR(4) in which the predictors are the dividend yield, the short-term nominal rate, the term spread, IP growth, and unemployment, and 5% of the time a similar VAR(4) in which the default spread replaces IP growth. The remaining 4% of the time is spent using similar VAR(4) models that include 3-4 predictors at the time. Under 4) the investor uses the same models, but with slightly different frequencies, for instance 83% of the time a a VAR(4) model that includes all predictors, 8% of the time a VAR(4) in which the predictors are the dividend yield, the short-term nominal rate, the term spread, IP growth, and unemployment. In this case there is a small residual of 3% of the time spent using similar VAR(4) models that include 3-4 predictors at the time.

The bottom panel of Table 10 shows another result that should by now be somewhat familiar: a 1-month horizon investor would derive a higher CER (6.3%) from ignoring predictability altogether—linear and nonlinear—than by either adopting the MS-BIS class strategy (6.0%) or the best among all the VAR-class strategies, which in fact yield zero or negative real CERs. Table 11 repeats the exercise underlying Table 10, but assuming two alternative values for γ, 2 and 10. The implications for the CER rankings across classes of models are identical to Table 10 and consistent with the results in Table 9. For instance, also for low and high risk aversion levels, while a long-horizon investor would prefer MS-BIC over any other class of models—and in particular over strategies that are allowed to switch among different VARs—a 1-month investor would optimally disregard all evidence of predictability and use a simple Gaussian IID model.

6. The Role of Regime Switching Volatilities and Correlations

So far, Sections 3-5 have entertained a systematic comparison of a range of VAR models with two specific, three-state Markov switching model in which second moments are assumed to be constant over time. This appears to be consistent with the fact that by construction, the VAR models that have been featured in the bulk of the literature are themselves homoskedastic. However, it turns out that a model specification search similar to the one performed in Section 3.3 and expanded to include heteroskedastic MS models in which also the covariance matrix is allowed to change as a function of the same Markov states driving conditional mean parameters, often leads to select heteroskedastic MS models. Therefore in this Section we briefly investigate the recursive OOS portfolio performance of heteroskedastic MS models. Moreover, it may be interesting to try and tease out from the data what the economic value of modeling Markov switching in second moments may be when separated from the pure value of switching dynamics in expected real returns. Before proceeding further, let us stress that in an ex-post perspective, it would be incorrect (or at least, naive) to expect that heteroskedastic MS would always perform worse than homoskedastic MS models in recursive OOS experiments. Although in-sample it would be sensible (yet, this is not a necessity in the domain of non-linear models) to expect that homoskedastic MS provide a worse fit than heteroskedastic ones, it is well-known that sometimes simpler and more parsimonious models may perform better than richer models in OOS evaluation. As a result, it is important to stress that we are not performing the exercise in this Section only with the goal of showing that homoskedastic MS does not "fall too far behind" heteroskedastic MS models. To save space, we only report results for our baseline design in which $\gamma = 5$, although findings for $\gamma = 2$ and 10 are qualitatively similar to those reported below.

As a first step, we have expanded Table 2 to also include MSIH and MSVARH models. For simplicity, we have omitted from the resulting Table 12 all the single-state models, for which there is no obvious generalization to heteroskedastic versions, unless one resorts to ARCH-type modeling strategies. For completeness, we have replicated in Table 12 the same statistics for the homoskedastic models from Table 2. As in Table 2, we have boldfaced the best three models selected by each information criterion. One find is striking: whatever the information criterion, the top three models always consist of heteroskedasticity

MS models only. However, once one switches to consider heteroskedastic MS in place of the homoskedastic ones, the set of models that are selected are similar to the ones that have emerged in Table 2. In particular, both BIC and H-Q both highlight the virtues of a three-state MSIH, which is the heteroskedastic analog to the MSI model examined in Section 3.3. Interestingly, the strength of the sample evidence in favor of MSVARH models weakens when compared to what we had found in Table 2. In the light of these results, we next examine the recursive portfolio implications of a MSIH model.

Table 13, panel B, shows QMLE parameter estimates of this three-state model. Panel A reports single-state estimates as a benchmark (these are by definition identical to the single-state estimates in panel A of Table 3. Intuition for the properties of the model can be gained by commenting the parameter estimates within each regime. The first regime is a bear state in which expected real returns are negative (for 1-month nominal bills) or zero (for stocks and bonds, in the sense that the bear state mean parameters fail to be statistically significant). In the bear state, stocks are more volatile than they are unconditionally (in panel A of the table). The bear state is quite persistent with an average duration of almost 19 months. When the US financial markets leave the bear state, this is usually to switch to the intermediate, equity bull regime. Notice that differently from other papers in the Markov switching literature, the bear state is in no sense an extreme or "rare event" regime, as it characterizes more than 37% of all long samples one could simulate from the estimated MSIH. The second regime is a bull state with positive mean real returns on all assets, although the expected real return on stocks is particularly high and statistically significant. In this regime, all assets are less volatile than in the unconditional, single-state case. This regime is highly persistent with an average duration of 34 months and characterizes half of any long sample. This means that in almost half of the time, the US financial markets are characterized by positive real returns on all assets and moderate volatility, which fits historical experience. The third regime is another bull state, but with three interesting features: the dominant asset class in terms of mean real returns is long-term government bonds, while stocks have an estimated mean coefficient which fails to be significant at conventional levels. Bond and stock markets are more volatile in this state than in the single-state, unconditional benchmark; real returns on long-term bonds are highly correlated with both stocks (0.42) and 1-month T-bills (0.40). We have labeled this regime as a "bond bull state" with high volatility. Clearly, the data lead to specifying this third regime because they need the flexibility to specify heterogeneous dynamics for bond and stock returns during bull regimes. This third regime is also highly persistent, with an average duration of 21 months. Finally, the estimated transition matrix in Table 13 has a rather special structure, by which regimes 1 and 2 and 3 and 1 "communicate" on a frequent basis, while regime 2 appears somewhat "isolated". As a result, regime 2 is considerably persistent. The fact that the third state is very persistent but in a sense isolated from regime 2 explains why regime 3 has an ergodic probability of less than 13%.

Figure 11 completes our description of the MSIH model by plotting the smoothed probabilities for each of the three regimes. The figure is consistent with the interpretation provided above. The first (bear) state characterizes a non-negligible portion of the data and picks up relatively long-lived episodes that consist of either well-known US recessions as dated by the NBER (e.g., 1974-1975, 1978-1980, 2001-2004, and more recently 2008) or of periods of crisis in the US financial markets with declining interest rates and negative realized stock and bond returns (such as the early 1970s, 1987-1988, and the international bond market crisis of 1998). The second (bull) state is exceptionally persistent and has in fact characterized long chunks of the recent US financial history, such as most of the 1960s, 1989-1997, and 2000 with some additional spikes during the 1980s. Finally, the third state is characterized by three obvious episodes, which are the long period (1981-1986) of declining inflation and short-term rates in the US after the inflationary bouts of the late 1970s, 2006, and (interestingly) the final months of 2008 and early 2009. These are periods of declining short-term rates and of increasing long-term bond prices that lead – consistently with our characterization of the regime – to high and statistically significant real bond returns.

Figure 12 shows recursive optimal portfolio weights (for $T = 1$ month and 5 years) derived from the MSIH model. These weights are computed under the assumption of $\gamma = 5$. Similarly to Figure 8, one can recognize typical MS-style regime dynamics in implied weights. Also under MSIH, the differences between short- and long-run portfolios exist but are generally modest (which means that hedging demands are small, see below). If these plots are compared to the VAR ones in Figures 8 and 9, one can iterate the comment that even a medium-scale VAR(1) model cannot produce the rich, regime-like dynamics in SAA that a MSIH model generates. For instance, while MSIH implies an average demand for stocks that oscillates around a moderate, positive percentage commitment, VAR produces generally high and wildly oscillating stock weights that for a long-horizon investor easily go from -200 to +400% in a few months only. As in Guidolin and Timmermann (2007) the reason for these more stable, less extreme long-run asset allocations under MSIH comes from the tendency of MSIH to attach considerable importance to the shape of its implied ergodic joint density for real asset returns when the horizon is sufficiently long, which has stabilizing and "moderating" effects on portfolio structure. Figure 12 also shows the recursive hedging demands for the period 1973:01-2009:12. The MSIH hedging demands are generally small when compared to VAR hedging demands and are once more stable over time.

Finally, we have computed and tabulated OOS performance statistics for recursive realized portfolios over the period 1973:01-2008:12 (or the shorter period implied by $H = 60$ months). Focussing on the baseline case of $\gamma = 5$, we have obtained that MSIH leads to a 5-year portfolio strategy that returns a real CER of 4.7% that would place MSIH in third place in the CER ranking of Table 8, after MSI and the Gaussian IID benchmark. The corresponding confidence interval is [3.4%, 6.0%] which tends to overlap to the other confidence intervals we have reported in Table 8. Interestingly, the Sharpe ratio of MSIH is substantially lower than most other models in Table 8 (0.23) so that the positive CER performance of MSIH entirely derives from its ability to inform portfolio strategies of the behavior of asset returns in the tails of their joint conditional density, which leads to a modest, almost nil excess kurtosis of realized performances (0.38 only). This stable performance translates into high and significantly positive real CERs to a $\gamma = 5$ power utility investor. We have also examined the performance of the MSIH strategy at a short-horizon and/or assuming $\gamma = 2$, 10. The general indication we have drawn is that that the results reported in Section 5 in no way depended on the choice of restricting the covariance matrix of real asset returns to be constant over time, in spite the strong indications of heteroskedasticity contained in the data. Additionally, MSIH is clearly superior to all the VAR models entertained in this paper, while Figure 12 has shown that the implied dynamics of portfolio weights bears little or no resemblance to what an investor would have computed in real time using any of the VAR models we have considered. However, it is also interesting that the real CER of MSIH turns out to be inferior to that of clearly misspecified models, such as MSI and the no predictability IID benchmark. This may depend on either the presence of substantial misspecifications in the way time-variation in the covariance matrix of the returns is captured or on the fact that MSIH is a substantially heavier (more richly parameterized) model than MSI is, generating a need to estimate 12 additional parameters. We leave to future research to investigate what the sources of the inferior performance of MSIH may turn out to be.

7. Conclusion

This paper has asked whether it is possible for a large class of VAR models—as defined by the predictors included, their lag structure, and whether they are estimating on a rolling or an expanding window of data—that forecast real asset returns to imply dynamic SAA choices and realized, ex-post performances similar to decisions and performances typical of (slightly) more complicated nonlinear econometric frameworks in which the existence of regimes is accounted for. After identifying the nonlinear framework with a simple three-state MS model of the type recently employed by Ang and Bekaert (2002, 2004) and Guidolin and Timmermann (2007, 2008b), we have obtained a clear negative answer to our main research question: simple VARs are not "sufficient" in either an economic or a statistical sense to summarize the predictability present in U.S. data over the period 1953-2008. Our key result is that in a simple, recursive portfolio experiment no fraction of the VARs estimated can produce SAA choices for long-horizon investors that compete with those obtainable under a three-state MSI model. This result does not depend on the assumed level of relative risk aversion and on the details of the MS models considered, in the sense that also MS models that are richer than a three-state MSI—specifically, a MSVAR model that captures both linear and nonlinear predictability patterns as well as instability in the relationships among real asset returns and predictors—generally outperform simple linear predictability frameworks.

In an attempt to offer a "clean" summary for the differential performance of MS and VAR strategies, we have performed a horse race in which models (both MS and VARs) are not considered in their individuality, but instead as an overall class. In practice, we allow an investor to select over time different models within either the MS or the VAR class on the basis of their recursively computed information criteria (AIC and BIC). We find that a power utility investor with $\gamma = 5$ and a 5-year horizon, would be ready to pay 8.1% in real terms to be allowed to select models from the MS class, while analogous calculation for the class of expanding window VAR yields a disappointing 0.3% per annum. This difference of almost 780 basis points can be taken as a strong indication that taking nonlinearities into accounting in SAA problems may handsomely pay out.

We have disregarded transaction costs: it is possible for a model to imply a superior OOS performance just because the strategy implies frequent and radical portfolio rebalancing, e.g., a more activist stance that aggressively times market regimes. In reality, it may be dubious that such a strategy may actually outperform more passive, and less trading-intense strategies as most investors would have to pay enormous fractions of their wealth in the form of fees, commissions, and bid-ask spreads. Although this an interesting avenue for further research, we can offer two preliminary thoughts. First, a casual investigation of Figures 4 and 5 (taking their left-scales into account), reveals that MS strategies do not imply more aggressive trading than VAR models do. Admittedly, there is a visible trade-off between the infrequent large changes in MS allocations and the continuous variability in the more persistent VAR weights. Second, Tables 6 and 7 have been built with this concern in mind. Although not commented so far, the last column of Tables 6-7 reveals that MS models (especially MSI) imply much less trading (as shown by the average monthly turnover statistics) than VARs do. So the concern above seems to rest on thin grounds.[33]

There are a number of details of our experiment that could have been different. Our investor may have cared for the utility of final wealth only (i.e., the problem may have no interim consumption, as in Avramov, 2002); her preferences could have been different (e.g., Epstein-Zin's preferences as in Campbell et al., 2003, or the wide set of preferences in Ait-Sahalia and Brandt, 2001); many investors would probably impose constraints when solving their portfolio problem, such as short-sale constraints. Of course, it would be sensible to repeat our exercise after either expanding the family of VARs considered (e.g., by adding other predictors, like Ludivgson and Ng's, 2007, *cay*) or adopting alternative nonlinear benchmarks (for instance smooth transition regressions as in Guidolin et al., 2009, or MS models with time-varying transition probabilities as in Ang and Bekaert, 2002). We leave these extensions for future research.

Appendix: Solution of asset allocation problems by Monte Carlo methods

Markov Switching Model

Given the optimization problem is solved backwards at each time t (since the portfolio can be rebalanced every month), such that $a\left(\boldsymbol{\pi}_{t+1}^{i}, t+1\right)$ is known for all values of $i = 1, 2, \ldots, Q$ on a discretization grid. Here $a(.)$ is not a function of the state variables $\mathbf{Z}_{t+1}$ but the regime probabilities $\boldsymbol{\pi}_{t+1}$. Computing a Monte Carlo approximation of the expectation

$$E_t\left[\{\boldsymbol{\omega}_t \mathbf{R}_{t+1,g}\}^{1-\gamma} a\left(\boldsymbol{\pi}_{t+1}^{i}, t+1\right)\right]$$

requires drawing G random samples of asset returns $\{\mathbf{R}_{t+1,g}(\boldsymbol{\pi}_{t+1}^{i})\}_{g=1}^{G}$ from the $t+1$ one-step joint density conditional on the period-t parameter estimates $\hat{\boldsymbol{\theta}}_t = \left(\left\{\hat{\boldsymbol{\mu}}_S, \hat{\boldsymbol{\Omega}}_S\right\}_{S=1}^{k}, \hat{\mathbf{P}}\right)$ assuming that, at each point $\boldsymbol{\pi}_t^i$ is updated to $\boldsymbol{\pi}_{t+1}(\boldsymbol{\pi}_t^i)$. The algorithm consists of the following steps:

[33] Additionally, Table 5 has shown that both the standard deviation and the range of variation of MS weights (especially MSI) are inferior to those of VAR models.

1. For each possible value of the current regime S_t simulate G returns $\{\mathbf{R}_{t+1,g}(S_{t+1})\}_{g=1}^{G}$ in calendar time from the regime switching model:

$$\mathbf{R}_{t+1,g}\left(S_t\right) = \boldsymbol{\mu}_{S_{t+1}} + \boldsymbol{\varepsilon}_{t+1,g} \qquad \boldsymbol{\varepsilon}_{t+1,g} \sim N\left(\mathbf{0}, \boldsymbol{\Omega}_{S_{t+1}}\right).$$

   The simulation enables regime switching as governed by the transition probability matrix $\hat{\mathbf{P}}_t$. For example, starting in state 1, the probability of switching to state 2 between t and $t+1$ is $\hat{p}_{12} \equiv \mathbf{e}_1' \hat{\mathbf{P}}_t \mathbf{e}_2$, while the probability of remaining in state 1 is $\hat{p}_{11} \equiv \mathbf{e}_1' \hat{\mathbf{P}}_t \mathbf{e}_1$. Hence, at each point in time, $\hat{\mathbf{P}}_t$ governs possible state transitions.

2. Combine the simulated returns $\{\mathbf{R}_{t+1,g}\}_{g=1}^{G}$ into a random sample size G, using the probability weights contained in the vector $\boldsymbol{\pi}_t^j$:

$$\mathbf{R}_{t+1,g}(\boldsymbol{\pi}_t^i) = \sum\nolimits_{j=1}^{k} (\boldsymbol{\pi}_t^i \mathbf{e}_j) \mathbf{R}_{t+1,g}\left(S_t = j\right)$$

3. Update the future regime probabilities perceived by the investor using the standard Hamilton-Kim filtering formula

$$\boldsymbol{\pi}_{t+1}\left(\boldsymbol{\pi}_t^i\right) = \frac{\left(\boldsymbol{\pi}_t^i\right)' \hat{\mathbf{P}} \odot \boldsymbol{\eta}\left(\mathbf{R}_{t+1,g}(\boldsymbol{\pi}_t^i); \hat{\boldsymbol{\theta}}_t\right)}{\left(\left(\boldsymbol{\pi}_t^i\right)' \hat{\mathbf{P}} \odot \boldsymbol{\eta}\left(\mathbf{R}_{t+1,g}(\boldsymbol{\pi}_t^i); \hat{\boldsymbol{\theta}}_t\right)\right) \boldsymbol{\iota}_k}.$$

   This gives an $G \times k$ matrix $\left\{\boldsymbol{\pi}_{t+1}\left(\boldsymbol{\pi}_t^i\right)\right\}_{g=1}^{G}$, whose rows correspond to simulated vectors of perceived regime probabilities at time $t+1$.

4. For all $g = 1, 2, \ldots, G$ calculate the value $\tilde{\boldsymbol{\pi}}_{t+1,g}^i$ on the discretization grid $(i = 1, 2, \ldots, Q)$ closest to $\boldsymbol{\pi}_{t+1,g}(\boldsymbol{\pi}_t^i)$ using the distance measure $\sum_{j=1}^{k-1} \left|\boldsymbol{\pi}_{t+1}^i \mathbf{e}_j - \boldsymbol{\pi}_{t+1,g} \mathbf{e}_j\right|$, i.e.

$$\tilde{\boldsymbol{\pi}}_{t+1,g}^i(\boldsymbol{\pi}_t^i) \equiv \arg\min \sum_{j=1}^{k-1} \left|\mathbf{x}\mathbf{e}_j - \boldsymbol{\pi}_{t+1,g}\mathbf{e}_j\right|.$$

   Knowledge of the vector $\left\{\tilde{\boldsymbol{\pi}}_{t+1,g}^i\left(\boldsymbol{\pi}_t^i\right)\right\}_{g=1}^{G}$ allows us to build $\left\{a(\boldsymbol{\pi}_{t+1}^{(i,g)}, t+1)\right\}_{g=1}^{G}$, where $\boldsymbol{\pi}_{t+1}^{(i,g)} \equiv \tilde{\boldsymbol{\pi}}_{t+1,g}^i(\boldsymbol{\pi}_t^i)$ is a function of the assumed, initial vector of regime probabilities $\boldsymbol{\pi}_t^i$.

5. Solve the program

$$\max_{\boldsymbol{\omega}_t(\boldsymbol{\pi}_t^i)} G^{-1} \sum_{g=1}^{G} \left\{ \left[\boldsymbol{\omega}_t \mathbf{R}_{t+1,g}\right]^{1-\gamma} a\left(\boldsymbol{\pi}_{t+1}^{(i,g)}, t+1\right) \right\}$$

   For large values of G this provides an arbitrarily precise Monte Carlo approximation to $E[\{\boldsymbol{\omega}_t \mathbf{R}_{t+1,g}\}^{1-\gamma} a\left(\boldsymbol{\pi}_{t+1}^i, t+1\right)]$. The value function evaluated at the optimal portfolio weights $\hat{\boldsymbol{\omega}}_t(\boldsymbol{\pi}_t^i)$ gives $a(\boldsymbol{\pi}_t^i, t)$ for the ith point on the initial grid. We also check whether $\boldsymbol{\omega}_t \mathbf{R}_{t+1,g}$ is negative and reject all corresponding sample paths.

The algorithm is applied to all possible values $\boldsymbol{\pi}_t^i$ on the discretization grid until all values of $a(\boldsymbol{\pi}_t^i, t)$ are obtained for $i = 1, 2, \ldots, Q$. It is then iterated backwards. We take $a(\boldsymbol{\pi}_{t+1}^i, t+1)$ as given and use the actual vector of smoothed probabilities $\boldsymbol{\pi}_t$. The resultant vector $\hat{\boldsymbol{\omega}}_t$ gives the optimal portfolio allocation at time t, while $a(\boldsymbol{\pi}_t, t)$ is the optimal value function. In our application, Q is selected as $5^2 = 25$ which fits the standard formula 5^{k-1} as in Guidolin and Timmermann (2008b) and the number of Monte Carlo simulations is 30,000.

VAR model

Again the optimization problem is solved by backward iteration for each point t so that $a(\mathbf{Z}_{t+1}, t+1)$. A Monte Carlo approximation of the expectation

$$E_t\left[\{\boldsymbol{\omega}_t \mathbf{R}_{t+1,g}\}^{1-\gamma} a\left(\mathbf{Z}^i_{t+1}, t+1\right)\right]$$

now requires drawing G random samples of the state variables $\{\mathbf{Z}_{t+1,g}\}_{g=1}^G$ from the $t+1$ one-step joint density conditional on the period-t parameter estimates $\hat{\boldsymbol{\theta}}_t = \left(\hat{\boldsymbol{\mu}}, \hat{\mathbf{A}}, \hat{\boldsymbol{\Omega}}\right)$. The algorithm is similar but much simpler than for the Markov Switching model. The G returns $\left\{\mathbf{R}_{t+1}(\mathbf{Z}^i_t)\right\}_{g=1}^G$ need to be simulated from the VAR model. In this case $Q = 20$ delivers quite accurate results (because of the linearity of the prediction framework) and we set again $G = 30,000$.

References

[1] Ait-Sahalia, Y., Brandt, M., 2001. Variable selection for portfolio choice. *Journal of Finance* 56, 1297-1351.

[2] Ang, A., Bekaert, G., 2002. International asset allocation with regime shifts. *Review of Financial Studies* 15, 1137-1187.

[3] Ang, A., Bekaert, G., 2004. How do regimes affect asset allocation? *Financial Analysts Journal* 60, 86-99.

[4] Ang, A., Bekaert, G., 2007. Stock return predictability: is it there? *Review of Financial Studies* 20, 651-707.

[5] Ang, A., Bekaert, G., Liu, J., 2005. Why stocks may disappoint? *Journal of Financial Economics* 76, 471-508.

[6] Ang, A., Chen, J., Xing, Y., 2006. Downside risk. *Review of Financial Studies* 19, 1191-1239.

[7] Avramov, D., 2002. Stock return predictability and model uncertainty. *Journal of Financial Economics* 64, 423-458.

[8] Bai, J., Ng, S., 2008. Forecasting economic time series using targeted predictors. *Journal of Econometrics* 146, 304-317.

[9] Balvers, R., Cosimano, T., McDonald, B., 1990. Predicting stock returns in an efficient market. *Journal of Finance* 45, 1109-1128.

[10] Barberis, N., 2000. Investing for the long run when returns are predictable. *Journal of Finance* 55, 225-264.

[11] Bauwens, L., Omrane, W., Rengifo, E., 2010. Intradaily dynamic portfolio selection. *Computational Statistics and Data Analysis* 54, 2400-2418

[12] Boyd, J., Hu, J., Jagannathan, R., 2005. The stock markets reaction to unemployment news: why bad news is usually good for stocks. *Journal of Finance* 60, 649-672.

[13] Brandt, M., 2004. Portfolio choice problems, in Y. Ait-Sahalia and L.P. Hansen, *Handbook of Financial Econometrics*, North Holland, Amsterdam.

[14] Brennan, M., Schwartz, E., Lagnado, R., 1997. Strategic asset allocation. *Journal of Economic Dynamics and Control* 21, 1377-1403.

[15] Campbell, J. Y., 1987. Stock returns and the term structure. *Journal of Financial Economics* 18, 373-399.

[16] Campbell, J. Y., Chan, Y. L., Viceira, L., 2003. A multivariate model of strategic asset allocation. *Journal of Financial Economics* 67, 41-80.

[17] Campbell, J. Y., Lo, A., MacKinlay, C., 1997. *The econometrics of financial markets.* Princeton University Press, Princeton.

[18] Campbell, J. Y., Viceira, L., 2002. *Strategic asset allocation: portfolio choice for long-term investors.* Oxford University Press, Oxford.

[19] Campbell, J. Y., Vuolteenaho, T., 2004. Bad beta, good beta. *American Economic Review* 94, 1249-1275.

[20] Cutler, D., Poterba, J., Summers, L., 1989. What moves stock prices. *Journal of Portfolio Management* 15, 4-12.

[21] Davies, R. B., 1977. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 64, 247-254.

[22] Detemple, J., Garcia, R., Rindisbacher, M., 2003. A Monte Carlo method for optimal portfolios. *Journal of Finance* 58, 401-446.

[23] Fama, E. F., French, K. R., 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22, 3-25.

[24] Fama, E. F., French, K. R., 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25, 23-49.

[25] Fama, E. F., Schwert, G. W., 1977. Asset returns and inflation. *Journal of Financial Economics* 5, 115-146.

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

[27] Garcia, R., Perron, P., 1996. An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics* 78, 111-125.

[28] Granger, C., Terasvirta, T., 1993. *Modelling nonlinear economic relationships*. Oxford University Press, Oxford.

[29] Guidolin, M., Ono, S., 2006. Are the dynamic linkages between the macroeconomy and asset prices time-varying? *Journal of Economics and Business* 58, 480-518.

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

[31] Guidolin, M., Timmermann, A., 2007. Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control* 31, 3503-44.

[32] Guidolin, M., Timmermann, A., 2008a. International asset allocation under regime switching, skew and kurtosis preferences. *Review of Financial Studies* 21, 889-935.

[33] Guidolin, M., Timmermann, A., 2008b. Size and value anomalies under regime shifts. *Journal of Financial Econometrics* 6, 1-48.

[34] Guidolin, M., Hyde, S., McMillan, D., Ono, S., 2009. Non-linear predictability in stock and bond returns: when and where is it exploitable? *International Journal of Forecasting* 25, 373-399.

[35] Hamilton, J. D., 1994. *Time series analysis*. Princeton University Press, Princeton.

[36] Heij, C., van Dijk, D., Groenen, P., 2008. Macroeconomic forecasting with matched principal components. *International Journal of Forecasting* 24, 87-100.

[37] Ingersoll, J., 1987. *Theory of financial decision making*. Rowman & Littlefield, Totowa.

[38] Kandel, S., Stambaugh, R., 1996. On the predictability of stock returns: an asset allocation perspective. *Journal of Finance* 51, 385-424.

[39] Kim, C. J., Nelson, C. R., Startz, R., 1998. Testing for mean reversion in heteroskedastic data based on Gibbs-sampling-augmented randomization. *Journal of Empirical Finance* 5, 131-154.

[40] Ludvigson, S., Ng, S., 2007. The empirical risk-return relation: a factor analysis approach. *Journal of Financial Economics* 83, 171-222.

[41] Lynch, A., 2001. Portfolio choice and equity characteristics: characterizing the hedging demands induced by return predictability. *Journal of Financial Economics* 62, 67-130.

[42] Paye, B., Timmermann, A., 2006. Instability of return prediction models. *Journal of Empirical Finance* 13, 274-315.

[43] Perez-Quiros, G., Timmermann, A., 2000. Firm size and cyclical variations in stock returns. *Journal of Finance* 55, 1229-1262.

[44] Perez-Quiros, G., Timmermann, A., 2001. Business cycle asymmetries in stock returns: evidence from higher order moments and conditional densities. *Journal of Econometrics* 103, 259-306.

[45] Pesaran, M.H., Timmermann, A., 1995. Predictability of stock returns: robustness and economic significance. *Journal of Finance* 50, 1201-1228.

[46] Pesaran, M.H., Timmermann, A., 2000. A Recursive Modeling Approach to Predicting UK Stock Returns. *Economic Journal* 110, 159–191.

[47] Samuelson, P., 1969. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics* 51, 239-246.

[48] Stock, J., Watson, M., 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147162.

[49] Timmermann, A., 2000. Moments of markov switching models. *Journal of Econometrics* 96, 75-111.

Table 1

Summary Statistics for Portfolio Returns and Predictors

	Mean	Median	Std. Dev.	Uncond. Sharpe Ratio	Minimum	Maximum	Skewness	Kurtosis	J-B test
Real Stock Returns	0.586**	0.967**	4.341	0.112	-22.797	15.570	-0.567*	5.015*	149.7**
Long-term Govt. Bonds Real Returns	0.234**	0.108*	2.226	0.061	-7.478	10.453	0.465*	4.952**	130.9**
1-month T-bill Real Returns	0.099**	0.108**	0.317	__	-1.120	1.938	0.201	5.303*	153.0**
CPI Inflation rate	0.308**	0.297**	0.358	__	-1.915	1.806	0.039	6.129*	274.3**
Dividend Yield (annual MA,)	3.276**	3.185**	1.179	__	1.100	6.260	0.202*	2.556**	10.11**
Short-Term Nominal Rate (annualized)	4.889**	4.618**	2.841	__	0.035	18.190	1.173**	5.165**	285.3**
Riskless Term Spread (annualized)	1.509**	1.470**	1.321	__	-4.300	6.920	-0.114	4.249*	45.11**
Default Spread (Baa-Aaa, annualized)	0.956**	0.830**	0.432	__	0.320	3.380	1.659**	6.780**	708.5**
Industrial production growth (annualized)	2.822*	5.193**	2.393	__	-10.022	9.013	-0.571**	4.084*	69.47**
Unemployment Rate (percentage)	5.714**	5.500**	1.506	__	2.400	11.400	0.652**	3.641	59.20**

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

Table 2

Model Specification Search

Regimes (k)	VAR(p) order	Div. Yield	Short Rate	Term	Default	Inflation	IP Growth	Unempl.	Max Log-Likelihood	Linearity test	AIC	BIC	H-Q	No. Param.	Saturation ratio
		Predictors Included													
(Selected) Single-State Models															
1	0	N	N	N	N	N	N	N	5710.446	–	-16.9686	-16.9082	-16.9452	9	223.67
1	1	Y	N	N	N	N	N	N	5791.261	–	-17.1990	-17.0579	-17.1443	26	103.23
1	1	N	Y	N	N	N	N	N	5799.371	–	-17.2232	**-17.0821**	-17.1685	26	103.23
1	1	N	N	Y	N	N	N	N	5799.366	–	-17.2231	**-17.0820**	-17.1685	26	103.23
1	1	N	N	N	Y	N	N	N	5794.918	–	-17.2099	-17.0688	-17.1552	26	103.23
1	1	N	N	N	N	Y	N	N	5793.887	–	-17.2068	-17.0657	-17.1522	26	103.23
1	1	N	N	N	N	N	Y	N	5792.311	–	-17.2021	-17.0610	-17.1475	26	103.23
1	1	N	N	N	N	N	N	Y	5791.850	–	-17.2007	-17.0596	-17.1461	26	103.23
1	1	Y	Y	Y	Y	Y	Y	Y	5832.249	–	-17.2675	-17.0054	-17.1660	165	40.67
1	2	N	Y	N	N	N	N	N	5799.590	–	-17.1868	-16.9648	-17.1008	42	63.90
Two-State Models															
2	0	N	N	N	N	N	N	N	5760.781	93.588***	-17.1035	-17.0096	-17.0671	14	143.79
2	1	N	N	N	N	N	N	N	5814.817	71.124***	-17.2364	-17.0214	-17.1531	32	62.91
2	Restr.	Y	Y	Y	Y	Y	Y	Y	5857.283	50.067***	-17.3272	-17.0316	-17.2127	177	37.91
2	1	Y	Y	Y	Y	Y	Y	Y	5917.327	170.155***	-17.4168	-16.9195	-17.2242	277	24.22
Three-State Models															
3	0	N	N	N	N	N	N	N	5806.042	184.109***	-17.2174	**-17.1464**	**-17.2528**	21	95.86
3	1	N	N	N	N	N	N	N	5835.473	87.215***	-17.3950	-17.0434	-17.2369	30	67.10
3	Restr.	Y	Y	Y	Y	Y	Y	Y	5874.515	84.531***	-17.3577	-17.0150	-17.2250	191	35.13
3	1	Y	Y	Y	Y	Y	Y	Y	5992.346	320.192***	**-17.5301**	-16.7842	-17.2412	391	17.16
Three-State MSVAR Models with Principal Components of Predictors															
3	1	First principal component							5916.483	246.035***	**-17.4649**	-17.0819	**-17.3166**	76	35.32
3	1	First two principal components							5918.162	234.000***	**-17.4431**	-16.9996	**-17.2713**	111	30.23
3	1	First three principal components							5924.246	240.398***	-17.4344	-16.9305	-17.2392	163	24.70
Four-State Models															
4	0	N	N	N	N	N	N	N	5821.464	232.941***	-17.2622	-17.0606	-17.1841	30	67.10
4	1	N	N	N	N	N	N	N	5899.049	239.133***	-17.4121	-16.9681	-17.2401	66	30.50
4	Restr.	Y	Y	Y	Y	Y	Y	Y	5897.340	130.181***	-17.3989	-16.9958	-17.2428	207	9.72

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

Table 3

Full-Sample Estimates of Three-State Markov Switching Multivariate Model for Real Stock, Bond, and 1-month T-Bill Returns

	Panel A - SINGLE STATE MODEL		
	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns
1. Mean returns	0.0992**	0.2340**	0.5855**
2. Correlations/Volatilities			
Real 1-month T-bill Returns	0.3171**		
Real Long-Term Bond Returns	0.2869*	2.2257**	
Real Stock Returns	0.1228	0.1638*	4.3405**
	Panel B - THREE-STATE MODEL		
	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns
1. Mean returns			
Bear State	-0.3379**	-0.5443*	-0.8662
Normal State	0.1334**	0.1955*	0.7721**
Bull State	0.7225**	2.8695**	1.5373*
2. Correlations/Volatilities			
Real 1-month T-bill Returns	0.2291**		
Real Long-Term Bond Returns	0.1429	2.1201**	
Real Stock Returns	0.0451	0.1424	4.2926**
3. Transition probabilities	Bear State	Normal State	Bull State
Bear State	0.7098**	0.2761**	0.0141
Normal State	0.0481*	0.9392**	0.0127
Bull State	0.0201	0.1964**	0.7834**
	Panel C - MARKOV CHAIN PROPERTIES, THREE-STATE MODEL		
	Bear State	Normal State	Bull State
Ergodic Probabilities	0.1376	0.8061	0.0563
Average Duration (in months)	3.5	16.4	4.6

** = significant at 1% size or lower; * = significant at 5% size.

Table 4

Full-Sample Estimates of Three-State Markov Switching VAR(1) Model for Real Stock, Bond, 1-month T-Bill Returns, and First Principal Component of Predictors

	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns	First PC of Predictors
Panel A - SINGLE STATE MODEL				
1. Intercept	0.0594**	0.1335	0.4209*	-0.0309
3. VAR matrix				
Real 1-month T-bill Returns	0.3984**	1.3302**	0.7649	45.6291**
Real Long-Term Bond Returns	0.0088	0.0833*	0.2138**	-1.2088
Real Stock Returns	-0.0030	-0.0832**	0.0732	-1.0739**
First PC of Predictors	-0.0030	0.0007	0.0010	0.9653**
2. Correlations/Volatilities of shocks				
Real 1-month T-bill Returns	0.2858**			
Real Long-Term Bond Returns	0.2201**	2.1471**		
Real Stock Returns	0.0855*	0.1554**	4.2767**	
First PC of Predictors	-0.4794**	-0.0218	-0.1966**	0.3730**
Panel B - THREE-STATE MODEL				
	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns	First PC of Predictors
1. Intercepts				
Bear State	-0.0959**	-0.0470*	-2.3129**	-0.0151
Normal State	0.0133**	0.0894	0.2158**	0.0174**
Bull State	0.0781	0.3164	-1.2446	0.2160*
2. VAR matrix				
Real 1-month T-bill Returns	0.2067	0.0172	0.0055*	-0.0009**
Real Long-Term Bond Returns	0.7251*	-0.0339	-0.0011**	-0.0020**
Real Stock Returns	-2.6573	0.5819	0.2415**	-0.0174*
First PC of Predictors	57.7428	-3.9443	-1.6758*	0.9491**
Real 1-month T-bill Returns	0.0625	0.0130*	0.0057*	-0.0005**
Real Long-Term Bond Returns	0.7765	0.2006**	-0.0952	0.0005
Real Stock Returns	1.8391	0.1176**	-0.0464	0.0017**
First PC of Predictors	75.1666**	-1.0257	-1.5686	1.0042**
Real 1-month T-bill Returns	0.0817**	0.0092*	0.0061	0.0002
Real Long-Term Bond Returns	0.8836	0.1027	0.0668*	0.0020
Real Stock Returns	1.3457	0.1545*	0.1018	0.0033
First PC of Predictors	14.1380	1.5704	1.0187*	0.0313**
3. Correlations/Volatilities				
Real 1-month T-bill Returns	0.2635**			
Real Long-Term Bond Returns	0.1971*	1.9010**		
Real Stock Returns	0.0581	0.1102	3.4716**	
First PC of Predictors	-0.5046**	-0.0699	-0.1292	0.3213**
3. Transition probabilities	Bear State	Normal State	Bull State	
Bear State	0.9521**	0.0254*	0.0225	
Normal State	0.0176	0.9754**	0.0070	
Bond Bull State	0.0000	0.0454**	0.9546**	
Panel C - MARKOV CHAIN PROPERTIES, THREE-STATE MODEL				
	Bear State	Normal State	Bond Bull State	
Ergodic Probabilities	0.2162	0.5870	0.1968	
Average Duration (in months)	20.9	40.7	22.0	

** = significant at 1% size or lower; * = significant at 5% size.

Table 5

Best Models and Selected Benchmarks Ranked According to Recursive 12-month Ahead Predictive Performance for Real Asset Returns

Panel A – Real Stock Returns

Rank	Window length	Number of lags	Model	Predictors included							Real Stock Returns			
				Div. Yield	Short Rate	Term Spread	Default Spr.	Inflation	IP Growth	Unempl.	RMSFE	Bias	St. Dev.	MAFE
1	Expanding	p=0	MSI	N	N	N	N	N	N	N	0.04110	-0.00069	0.04109	0.03116
2	Expanding	p=2	VAR	N	N	N	N	N	Y	Y	0.04663	-0.00206	0.04659	0.03538
3	Expanding	p=4	VAR	N	N	N	N	N	N	N	0.04663	-0.00095	0.04662	0.03534
4	Expanding	p=2	VAR	N	N	N	N	N	N	Y	0.04665	-0.00155	0.04662	0.03551
5	Expanding	p=2	VAR	N	N	Y	N	N	N	N	0.04661	-0.00107	0.04660	0.03539
6	Expanding	p=2	VAR	N	N	Y	N	N	Y	N	0.04661	-0.00087	0.04660	0.03539
7	Expanding	p=2	VAR	N	N	Y	N	N	N	Y	0.04668	-0.00163	0.04666	0.03564
8	Expanding	p=1	VAR	N	N	Y	N	N	N	N	0.04659	-0.00079	0.04658	0.03537
9	Expanding	p=1	VAR	N	N	Y	N	N	Y	N	0.04659	-0.00064	0.04658	0.03538
10	Expanding	p=2	VAR	N	N	N	N	N	N	N	0.04661	-0.00069	0.04661	0.03534
23	Rolling	p=4	VAR	N	N	N	N	N	N	N	0.04677	-0.00031	0.04677	0.03542
41	Expanding	p=12	VAR	N	N	N	N	N	N	N	0.04719	-0.00111	0.04717	0.03582
125	Expanding	p=1	MSVAR(1) -1PC	First PC of all predictors							**0.02484**	0.00691	**0.02386**	**0.01814**
B	Expanding	p=0	IID	No Predictability							0.04752	0.00062	0.04752	0.03671
			Total	0	0	5	0	0	3	3				

Panel B – Real Bond and 1-month T-Bill Returns

Rank	Window length	Number of lags	Model	Predictors included							Real Bond Returns				Real T bill Returns			
				Div. Yield	Short Rate	Term	Default	Inflation	IP	Unempl.	RMSFE	Bias	St. Dev.	MAFE	RMSFE	Bias	St. Dev.	MAFE
1	Expanding	p=0	MSI	N	N	N	N	N	N	N	0.02163	0.00184	0.02155	**0.01630**	**0.00311**	0.00018	**0.00311**	**0.00236**
2	Expanding	p=2	VAR	N	N	N	N	N	Y	Y	0.02450	0.00196	0.02443	0.01837	0.00354	0.00024	0.00353	0.00267
3	Expanding	p=4	VAR	N	N	N	N	N	N	N	0.02453	0.00229	0.02442	0.01842	0.00354	0.00030	0.00353	0.00268
4	Expanding	p=2	VAR	N	N	N	N	N	N	Y	0.02456	0.00231	0.02445	0.01842	0.00354	0.00033	0.00353	0.00268
5	Expanding	p=2	VAR	N	N	Y	N	N	N	N	0.02449	0.00221	0.02439	0.01839	0.00357	0.00032	0.00356	0.00271
6	Expanding	p=2	VAR	N	N	Y	N	N	Y	N	0.02448	0.00204	0.02440	0.01837	0.00357	0.00033	0.00356	0.00271
7	Expanding	p=2	VAR	N	N	Y	N	N	N	Y	0.02455	0.00220	0.02445	0.01841	0.00355	0.00030	0.00354	0.00269
8	Expanding	p=1	VAR	N	N	Y	N	N	N	N	0.02453	0.00243	0.02441	0.01841	0.00358	0.00036	0.00356	0.00271
9	Expanding	p=1	VAR	N	N	Y	N	N	Y	N	0.02454	0.00249	0.02441	0.01842	0.00358	0.00036	0.00356	0.00271
10	Expanding	p=2	VAR	N	N	N	N	N	N	N	0.02454	0.00247	0.02441	0.01842	0.00358	0.00038	0.00356	0.00271
23	Rolling	p=4	VAR	N	N	N	N	N	N	N	0.02455	0.00106	0.02453	0.01842	0.00357	-0.00006	0.00357	0.00267
41	Expanding	p=12	VAR	N	N	N	N	N	N	N	0.02448	0.00182	0.02441	0.01854	0.00349	0.00012	0.00349	0.00256
125	Expanding	p=1	MSVAR(1) -1PC	First PC of all predictors							**0.01931**	0.001963	**0.01921**	0.01989	0.00422	0.00029	0.00421	0.00316
B	Expanding	p=0	IID	No Predictability							0.02557	0.00237	0.02546	0.01892	0.00381	0.00058	0.00377	0.00284
			Total	0	0	5	0	0	3	3								

Table 6

Best Models and Selected Benchmarks Ranked According to Recursive 12-month Ahead Predictive Performance for Cumulative Real Asset Returns (Between t+1 and t+12)

Panel A – Real Stock Returns

Rank	Window length	Number of lags	Model	Predictors included							Real Stock Returns			
				Div. Yield	Short Rate	Term Spread	Default Spr.	Inflation	IP Growth	Unempl.	RMSFE	Bias	St. Dev.	MAFE
1	Expanding	p=0	MSI	N	N	N	N	N	N	N	0.15991	0.00334	0.15987	0.12393
2	Rolling	p=4	VAR	Y	Y	Y	N	N	N	N	0.17958	0.02535	0.17778	0.14521
3	Rolling	p=4	VAR	Y	Y	Y	N	N	Y	N	0.17916	0.02265	0.17773	0.14564
4	Rolling	p=4	VAR	Y	N	Y	N	Y	N	N	0.17970	0.02573	0.17785	0.14537
5	Rolling	p=4	VAR	Y	N	Y	N	Y	Y	N	0.17926	0.02292	0.17779	0.14576
6	Expanding	p=4	VAR	N	N	Y	N	N	N	N	0.18118	-0.02611	0.17929	0.14514
7	Rolling	p=4	VAR	N	N	Y	N	N	N	N	0.18192	-0.01029	0.18163	0.14365
8	Expanding	p=4	VAR	N	N	Y	N	N	Y	N	0.18033	-0.02454	0.17866	0.14435
9	Rolling	p=4	VAR	N	N	Y	N	N	Y	N	0.18141	-0.00664	0.18129	0.14354
37	Expanding	p=12	VAR	N	N	N	N	N	N	N	0.19437	-0.01701	0.19362	0.15045
70	Expanding	p=12	VAR	N	N	Y	N	N	N	N	0.20224	-0.02780	0.20032	0.16256
106	Expanding	p=1	MSVAR(1) -1PC	N	N	N	N	N	N	N	**0.15198**	0.05183	**0.14287**	**0.11689**
B	Expanding	p=0	IID	No Predictability							0.17971	0.00089	0.17971	0.13774
			Total	4	2	9	0	2	4	0				

Panel B – Real Bond and 1-month T-Bill Returns

Rank	Window length	Number of lags	Model	Predictors included							Real Bond Returns				Real T bill Returns			
				Div. Yield	Short Rate	Term	Default	Inflation	IP	Unempl.	RMSFE	Bias	St. Dev.	MAFE	RMSFE	Bias	St. Dev.	MAFE
1	Expanding	p=0	MSI	N	N	N	N	N	N	N	0.09223	0.02065	0.08989	0.07128	0.02153	0.00221	0.02142	0.01728
2	Rolling	p=4	VAR	Y	Y	Y	N	N	N	N	0.09823	0.01478	0.09711	0.07577	0.02020	0.00325	0.01994	0.01309
3	Rolling	p=4	VAR	Y	Y	Y	N	N	Y	N	0.09859	0.01334	0.09769	0.07644	0.02058	0.00344	0.02029	0.01329
4	Rolling	p=4	VAR	Y	N	Y	N	Y	N	N	0.09824	0.01486	0.09711	0.07578	0.02020	0.00320	0.01995	0.01308
5	Rolling	p=4	VAR	Y	N	Y	N	Y	Y	N	0.09861	0.01339	0.09769	0.07646	0.02058	0.00342	0.02029	0.01329
6	Expanding	p=4	VAR	N	N	Y	N	N	N	N	0.09327	0.01398	0.09221	0.07200	0.02165	**0.00003**	0.02165	0.01750
7	Rolling	p=4	VAR	N	N	Y	N	N	N	N	0.09381	0.00559	0.09364	0.07225	0.02169	-0.00385	0.02135	0.01755
8	Expanding	p=4	VAR	N	N	Y	N	N	Y	N	0.09311	0.01414	0.09203	0.07210	0.02191	0.00062	0.02190	0.01771
9	Rolling	p=4	VAR	N	N	Y	N	N	Y	N	0.09390	0.00632	0.09369	0.07217	0.02171	-0.00332	0.02146	0.01743
37	Expanding	p=12	VAR	N	N	N	N	N	N	N	0.09296	0.01122	0.09228	0.07394	0.01786	0.00037	0.01785	0.01336
70	Expanding	p=12	VAR	N	N	Y	N	N	N	N	0.08382	0.00647	0.08357	0.06679	0.01848	0.00044	0.01848	0.01404
106	Expanding	p=1	SVAR(1) -1I	N	N	N	N	N	N	N	0.09263	0.02380	**0.08952**	0.07132	0.02263	0.00116	0.02260	0.01700
B	Expanding	p=0	IID	No Predictability							0.10457	0.02798	0.10076	0.07945	0.02694	0.00421	0.02661	0.02173
			Total	4	2	9	0	2	4	0								

Table 7

Summary Statistics for Realized, Recursive Optimal Portfolio Weights (over the 1973:01 – 2008:12 Sample) Computed under Power Utility Preferences (γ = 5)

									1-month T-bills			Long-term Bonds			Stocks		
				Predictors Included					T=1	Long	Hedging	T=1	Long	Hedging	T=1	Long	Hedging
Model	Lags	DY	Short	Term	Def.	Infl.	IP grw.	Unempl.	Sample Mean of Portfolio Weights								
MSI	0	N	N	N	N	N	Y	N	0.088	0.147	0.059	0.400	0.314	-0.086	0.512	0.539	0.027
Gaussian IID	0				No Predictability				0.173	—	—	0.156	—	—	0.671	—	—
Exp. VAR	1	Y	N	N	N	N	N	N	-0.143	-1.064	-0.921	0.422	0.290	-0.132	0.721	1.774	1.053
Exp. VAR	1	N	N	N	Y	N	N	Y	-0.179	-1.097	-0.918	0.436	0.302	-0.134	0.743	1.795	1.052
Exp. VAR	1	N	N	N	Y	N	Y	Y	-0.873	-0.254	0.619	0.880	0.206	-0.674	0.993	1.048	0.055
Exp. VAR	1	N	N	N	Y	N	N	N	-0.884	-0.262	0.623	0.884	0.204	-0.681	1.000	1.058	0.058
Exp. VAR	1	N	N	N	Y	N	Y	N	-0.848	-0.278	0.570	0.934	0.298	-0.636	0.915	0.980	0.065
Exp. VAR	2	Y	Y	N	N	N	N	Y	-0.865	-0.294	0.570	0.944	0.306	-0.638	0.921	0.989	0.068
Rolling VAR	1				Lagged real asset returns only				-0.665	-0.468	0.197	1.154	0.855	-0.299	0.511	0.613	0.101
MSVAR(1)-1PC	1				First PC of all predictors				0.177	0.171	-0.006	0.292	0.291	-0.001	0.531	0.538	0.007
		DY	Short	Term	Def.	Infl.	IP grw.	Unempl.	Sample Standard Deviation of Portfolio Weights								
MSI	0	N	N	N	N	N	Y	N	1.375	1.368	0.111	1.053	1.048	0.167	0.653	0.584	0.151
Gaussian IID	0				No Predictability				0.197	—	—	0.201	—	—	0.063	—	—
Exp. VAR	1	Y	N	N	N	N	N	N	2.634	2.671	0.357	2.247	2.169	0.226	1.242	1.520	0.524
Exp. VAR	1	N	N	N	Y	N	N	Y	2.650	2.684	0.358	2.243	2.162	0.226	1.246	1.519	0.524
Exp. VAR	1	N	N	N	Y	N	Y	Y	2.834	2.614	0.387	2.480	2.280	0.422	1.171	1.183	0.107
Exp. VAR	1	N	N	N	Y	N	N	N	2.846	2.628	0.384	2.466	2.263	0.421	1.179	1.193	0.108
Exp. VAR	1	N	N	N	Y	N	Y	N	2.830	2.614	0.333	2.460	2.250	0.402	1.159	1.187	0.099
Exp. VAR	2	Y	Y	N	N	N	N	Y	2.846	2.632	0.330	2.450	2.237	0.400	1.170	1.199	0.101
Rolling VAR	1				Lagged real asset returns only				4.436	4.314	0.203	4.063	3.837	0.327	1.117	1.176	0.168
MSVAR(1)-1PC	1				First PC of all predictors				1.589	1.581	0.103	1.109	1.120	0.149	1.773	1.772	0.125
		DY	Short	Term	Def.	Infl.	IP grw.	Unempl.	Empirical 90% Range								
MSI	0	N	N	N	N	N	Y	N	5.186	5.118	0.175	3.746	3.763	0.378	2.013	1.641	0.286
Gaussian IID	0				No Predictability				0.629	—	—	0.648	—	—	0.200	—	—
Exp. VAR	1	Y	N	N	N	N	N	N	8.294	8.473	1.185	6.639	6.451	0.743	3.915	4.494	1.667
Exp. VAR	1	N	N	N	Y	N	N	Y	8.505	8.488	1.185	6.677	6.375	0.737	3.878	4.560	1.665
Exp. VAR	1	N	N	N	Y	N	Y	Y	8.729	8.164	1.080	7.062	6.417	1.191	3.648	3.777	0.251
Exp. VAR	1	N	N	N	Y	N	N	N	8.799	8.347	1.061	7.024	6.229	1.172	3.687	3.858	0.255
Exp. VAR	1	N	N	N	Y	N	Y	N	8.523	8.026	1.017	6.997	6.227	1.138	3.578	3.689	0.202
Exp. VAR	2	Y	Y	N	N	N	N	Y	8.680	8.201	1.012	6.946	6.248	1.141	3.686	3.750	0.208
Rolling VAR	1				Lagged real asset returns only				13.69	13.54	0.659	11.77	10.65	1.052	3.488	3.863	0.509
MSVAR(1)-1PC	1				First PC of all predictors				4.710	4.705	0.215	3.552	3.537	0.330	5.875	5.900	0.329

Table 8

Best Models and Selected Benchmarks Ranked According to Average Recursive Certainty Equivalent Return Obtained from Optimal Strategic Asset Allocation Choices Under Power Utility Preferences ($\gamma = 5$)

			Predictors Included									Annualized Mean			Annualized Volatility			Sharpe Ratio			CER				
CER Rank	Model	Lags	DY	Short	Term	Default	Infl.	IP grw.	Unempl.	Horizon	Mean returns (yearly)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Volatility (yearly)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Sharpe ratio	95% Conf. Int. -- LB	95% Conf. Int. -- UB	CER (% Ann.)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Skew	Kurtosis	Avg. monthly turnover (adj.)
1	MSI	0	N	N	N	N	N	Y	N	60	11.056	7.704	15.749	21.208	13.387	29.028	0.460	0.336	0.576	**7.955**	4.869	8.730	1.236	6.324	0.510
2	Gaussian IID	0	No Predictability							60	8.523	7.210	11.250	**12.773**	11.016	14.530	**0.565**	0.391	0.745	5.512	4.867	6.219	0.574	**3.171**	**0.054**
3	Exp. VAR	1	Y	N	N	N	N	N	N	60	**22.231**	12.368	33.855	53.683	27.239	80.128	0.390	0.298	0.482	3.713	-3.925	11.571	2.151	12.643	1.581
4	Exp. VAR	1	Y	N	N	N	N	Y	N	60	13.544	7.174	20.852	34.828	13.065	56.591	0.352	0.213	0.496	3.271	-4.065	9.544	2.726	18.289	1.583
5	Exp. VAR	1	N	N	N	Y	N	N	Y	60	14.000	7.406	21.604	35.770	13.759	57.781	0.355	0.214	0.486	3.245	-3.636	10.307	2.326	17.917	1.590
6	Exp. VAR	1	N	N	N	Y	N	Y	Y	60	12.561	6.414	19.467	33.793	12.164	55.421	0.333	0.197	0.457	3.140	-3.014	9.075	2.308	15.180	1.586
7	Exp. VAR	1	N	N	N	Y	N	N	N	60	12.810	6.502	19.901	34.572	12.811	56.333	0.333	0.205	0.455	3.139	-3.212	9.691	3.028	25.382	1.592
33	MSVAR(1)-1PC	1	First PC of all predictors							60	10.836	6.131	16.241	28.846	19.456	38.236	0.331	0.211	0.468	2.379	1.463	5.068	1.133	4.954	1.238
77	Exp. VAR	2	N	N	N	Y	N	Y	N	60	14.771	8.083	19.595	58.898	32.969	84.827	0.134	0.144	0.319	0.846	-4.670	6.055	1.950	6.164	1.652
121	Rolling VAR	1	Lagged real asset returns only							60	5.424	2.316	7.612	27.465	11.763	43.167	0.132	0.020	0.282	-0.189	-0.993	0.553	1.666	7.439	2.364
	Median Expanding VAR performance									60	11.670	5.655	15.863	52.161	28.071	76.251	0.109	0.099	0.310	-0.288	-0.956	0.402	1.737	6.080	2.345
	Median Rolling VAR performance									60	8.220	2.558	11.834	53.950	30.059	77.841	0.100	0.018	0.228	-0.537	-0.949	-0.117	1.610	5.195	3.507
1	Gaussian IID	0	No Predictability							1	6.425	2.775	10.087	**11.282**	10.369	12.195	0.454	0.277	0.623	**6.253**	2.431	10.286	0.172	**3.143**	**0.065**
2	MSI	0	N	N	N	N	N	N	N	1	7.596	0.203	12.098	13.078	7.947	18.209	**0.481**	0.337	0.623	5.935	0.820	11.495	1.170	4.797	0.376
3	Exp. VAR	1	Lagged real asset returns only							1	13.769	0.521	20.315	36.924	26.275	47.718	0.338	0.208	0.475	5.240	-5.430	15.904	2.313	14.483	1.611
4	Exp. VAR	1	N	N	N	N	N	Y	N	1	13.744	0.206	20.451	37.454	26.770	48.356	0.332	0.200	0.475	3.602	-7.388	15.954	0.484	3.230	1.627
5	Exp. VAR	1	Y	N	N	N	N	N	N	1	13.715	0.094	20.477	38.479	27.414	49.286	0.323	0.186	0.445	3.287	-9.046	15.783	1.520	8.569	1.628
6	Exp. VAR	1	N	N	N	N	N	N	Y	1	14.334	0.194	21.358	39.507	28.987	50.292	0.330	0.205	0.448	2.791	-8.477	14.876	2.168	11.441	1.598
7	Exp. VAR	1	Y	N	N	N	N	Y	N	1	13.634	-0.326	20.559	38.817	27.787	49.679	0.318	0.197	0.454	2.105	-9.721	13.730	1.860	7.921	1.633
51	MSVAR(1)-1PC	1	First PC of all predictors							1	5.601	-5.333	11.565	13.289	9.199	17.380	0.324	0.137	0.506	1.934	-29.831	30.440	0.033	6.779	1.197
57	Rolling VAR	1	Lagged real asset returns only							1	13.459	-3.586	21.944	40.608	17.966	63.250	0.299	0.170	0.439	-0.981	-18.782	18.374	2.628	20.567	2.410
106	Exp. VAR	2	Y	N	Y	N	N	Y	N	1	22.670	-0.039	33.985	53.214	27.401	79.028	0.402	0.267	0.533	-7.446	-39.132	27.068	1.947	11.349	2.525
	Median Expanding VAR performance									1	18.087	-4.173	28.986	48.669	26.668	70.671	0.106	0.206	0.470	0.214	-6.656	7.716	1.423	7.976	2.562
	Median Rolling VAR performance									1	16.208	-10.378	29.064	65.151	26.766	103.536	0.109	0.088	0.386	-2.619	-10.687	4.887	1.448	7.662	4.143
	Predictors	Total	5	0	1	4	0	7	3																

Note: in the table performance statistics are boldfaced when these are the best (maximum for mean, volatility, Sharpe ratio, CER, and skewness; minimum for kurtosis) among all the econometric models considered (including those not covered by the table). Because models are ranked in the table on the basis of their CERs, it is possible that the best model under other metrics may fail to appear in the table.

Table 9

Best Models and Selected Benchmarks Ranked According to Average Recursive Certainty Equivalent Return Obtained from Optimal Strategic Asset Allocation Choices Under Power Utility Preferences (γ = 2 and 10)

γ = 2

CER Rank	Model	Lags	Predictors Included: DY	Short	Term	Default	Infl.	IP grw.	Unempl.	Horizon	Annualized Mean: Mean returns (yearly)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Annualized Volatility: Volatility (yearly)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Sharpe Ratio: Sharpe ratio	95% Conf. Int. -- LB	95% Conf. Int. -- UB	CER: CER (% Ann.)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Skew	Kurtosis	Avg. monthly turnover (adj.)
1	MSI	0	N	N	N	N	N	N	N	60	27.226	-51.895	75.274	29.962	-6.409	66.333	**0.865**	0.678	1.053	**10.234**	4.852	15.161	0.602	**3.203**	0.294
2	Gaussian IID	0				No Predictability				60	22.101	5.293	32.912	**25.634**	9.155	42.113	0.811	0.548	1.062	9.512	6.248	13.069	1.010	3.810	**0.279**
3	MSVAR(1)-1PC	1				First PC of all predictors				60	25.386	14.223	38.382	33.954	23.219	44.689	0.709	0.592	0.807	6.420	3.055	9.999	1.252	4.826	1.898
4	Exp. VAR	4	Y	N	Y	N	N	Y	Y	60	29.613	-5.830	58.645	101.024	49.546	152.503	0.280	0.076	0.502	4.784	-2.509	12.435	0.562	3.015	6.834
5	Exp. VAR	4	N	N	Y	N	Y	Y	Y	60	36.174	-2.796	69.525	108.011	57.720	158.303	0.323	0.090	0.543	3.954	-4.741	13.572	0.497	3.495	3.503
6	Exp. VAR	1	N	N	N	N	Y	Y	N	60	29.639	-0.388	56.090	80.687	31.248	130.127	0.351	0.174	0.535	3.754	3.245	4.227	0.667	4.562	1.454
7	Exp. VAR	1				Lagged real asset returns only				60	29.232	0.486	54.943	78.380	30.688	126.072	0.356	0.191	0.537	3.674	3.153	4.191	0.719	4.899	1.435
						Median Expanding VAR performance				60	27.344	-4.457	53.918	89.468	39.123	139.814	0.291	0.084	0.496	-0.487	-0.590	-0.391	0.617	3.804	2.856
						Median Rolling VAR performance				60	23.778	-7.079	48.083	91.513	35.608	147.418	0.246	0.051	0.454	-0.680	-0.741	-0.613	0.671	4.162	6.046
1	Gaussian IID	0				No Predictability				1	10.217	-12.869	24.163	**11.209**	7.421	14.998	**0.796**	0.488	1.130	**9.945**	8.105	11.695	-0.047	4.351	**0.262**
2	MSI	0	N	N	N	N	N	N	N	1	14.183	-18.787	35.407	26.232	-16.758	69.223	0.123	-0.039	0.302	7.427	4.502	10.586	-0.116	5.866	1.302
3	Exp. VAR	1				Lagged real asset returns only				1	13.467	-19.712	30.135	45.053	-6.725	96.832	0.270	0.143	0.405	6.355	-0.713	14.033	0.927	5.549	1.498
4	MSVAR(1)-1PC	1				First PC of all predictors				1	11.494	-6.702	21.482	16.202	12.240	20.163	0.629	0.462	0.789	4.343	-0.156	9.104	-0.696	6.045	1.789
5	Exp. VAR	1	N	N	N	N	N	Y	N	1	13.441	-21.295	30.701	45.709	-7.042	98.461	0.266	0.139	0.395	4.268	-3.455	12.385	0.729	3.553	1.491
6	Exp. VAR	1	N	N	N	N	N	N	Y	1	14.034	-22.298	32.055	48.243	-3.027	99.514	0.264	0.139	0.399	3.946	-4.810	11.922	1.087	3.334	1.806
7	Exp. VAR	1	N	N	N	N	N	Y	Y	1	14.025	-22.167	31.905	48.975	-2.220	100.169	0.260	0.129	0.391	0.545	-8.333	9.600	0.850	4.179	1.736
						Median Expanding VAR performance				1	18.567	-23.173	58.833	78.469	4.408	152.531	0.220	0.078	0.352	-1.393	-4.936	2.300	1.724	8.455	8.230
						Median Rolling VAR performance				1	17.607	-29.458	60.324	128.295	48.761	207.828	0.127	-0.002	0.270	-10.433	-14.635	-6.189	2.004	9.903	19.610
	Predictors	Total	1	0	2	0	1	5	4																

γ = 10

CER Rank	Model	Lags	Predictors Included: DY	Short	Term	Default	Infl.	IP grw.	Unempl.	Horizon	Annualized Mean: Mean returns (yearly)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Annualized Volatility: Volatility (yearly)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Sharpe Ratio: Sharpe ratio	95% Conf. Int. -- LB	95% Conf. Int. -- UB	CER: CER (% Ann.)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Skew	Kurtosis	Avg. monthly turnover (adj.)
1	MSI	0	N	N	N	N	N	N	N	60	5.377	4.973	5.784	**9.007**	8.406	9.593	**0.453**	0.239	0.653	**6.191**	5.480	6.919	0.419	3.120	**0.025**
2	Gaussian IID	0				No Predictability				60	6.011	5.361	6.670	14.291	13.149	15.400	0.330	0.189	0.467	4.734	4.132	5.375	0.881	3.101	0.192
3	Exp. VAR	1	Y	N	N	Y	N	Y	Y	60	8.072	2.457	11.591	34.190	-11.890	80.270	0.198	0.017	0.387	2.233	1.601	2.893	2.451	15.131	0.810
4	Exp. VAR	1	Y	N	N	Y	N	Y	N	60	7.123	2.387	10.181	28.314	-6.471	63.100	0.206	0.033	0.398	2.215	1.648	2.849	2.229	13.883	0.805
5	Exp. VAR	1	Y	N	N	Y	N	N	Y	60	8.011	2.330	11.596	34.299	-12.973	81.572	0.196	-0.009	0.404	2.207	1.609	2.883	**2.828**	20.507	0.810
6	Exp. VAR	1	Y	N	N	Y	N	N	N	60	7.103	2.366	10.170	28.515	-8.260	65.291	0.204	-0.017	0.421	2.191	1.543	2.738	1.940	9.564	0.808
55	MSVAR(1)-1PC	1				First PC of all predictors				60	4.566	3.570	5.576	22.048	20.301	23.698	0.148	-0.021	0.322	1.239	-1.444	4.301	0.453	**3.077**	0.634
						Median Expanding VAR performance				60	7.938	1.626	11.667	38.149	-10.753	87.051	0.174	0.041	0.298	-4.590	-4.739	-4.397	2.355	15.841	1.144
						Median Rolling VAR performance				60	3.771	-0.275	5.976	25.999	0.874	51.124	0.095	-0.026	0.206	-4.709	-4.822	-4.606	1.664	9.149	1.675
1	Gaussian IID	0				No Predictability				1	3.694	1.860	5.558	**5.571**	5.137	6.011	0.430	0.270	0.607	**5.981**	2.655	9.209	0.299	4.356	**0.025**
2	Exp. VAR	1				Lagged real asset returns only				1	9.578	1.161	15.300	18.599	13.102	24.130	0.445	0.330	0.579	4.873	-3.935	14.444	**2.229**	16.704	0.800
3	MSI	0	N	N	N	N	N	N	N	1	4.552	2.086	7.006	7.413	6.669	8.195	0.439	0.267	0.603	4.784	3.803	5.828	-0.253	5.931	0.239
4	Exp. VAR	1	N	N	N	N	N	Y	N	1	9.562	0.859	15.485	18.859	13.468	24.422	0.438	0.310	0.569	2.619	0.326	4.829	0.541	4.726	0.809
5	Exp. VAR	1	Y	N	N	N	N	N	N	1	9.542	0.835	15.372	19.353	13.753	24.946	0.426	0.290	0.545	2.601	0.276	4.789	2.534	19.959	0.809
6	Exp. VAR	1	N	N	N	N	N	N	Y	1	9.955	0.831	16.158	19.871	14.554	25.420	0.436	0.298	0.566	2.459	0.417	4.868	0.509	3.790	0.794
27	MSVAR(1)-1PC	1				First PC of all predictors				1	5.891	-0.783	10.059	14.749	12.789	16.663	0.311	0.135	0.462	0.713	-1.937	3.224	0.354	**4.038**	0.504
						Median Expanding VAR performance				1	12.399	-1.729	21.572	28.729	22.439	35.456	0.386	0.240	0.532	-0.175	-11.059	10.536	1.400	12.197	1.162
						Median Rolling VAR performance				1	10.663	-6.213	21.374	38.776	27.842	51.003	0.241	0.101	0.389	-4.967	-51.081	41.182	1.382	12.674	1.730
	Predictors	Total	5	0	0	4	0	3	3																

Note: in the table performance statistics are boldfaced when these are the best (maximum for mean, volatility, Sharpe ratio, CER, and skewness; minimum for kurtosis) among all the econometric models considered (including those not covered by the table).

Table 10

Recursive Out-of-Sample Portfolio Performance from Alternative Classes of Models Ranked According to Average Certainty Equivalent Return (γ = 5)

				Annualized Mean			Annualized Volatility			Sharpe Ratio		CER				
CER Rank	Model	Horizon	Annualized mean returns	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Annualized volatility	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Sharpe ratio	95% Conf. Int. -- LB	95% Conf. Int. -- UB	CER (% Annualized)	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Skew	Kurtosis	Average monthly turnover (adj.)
1	MS-BIC	60	10.647	6.151	14.372	20.613	13.837	28.944	0.453	0.310	0.584	**8.132**	4.745	9.641	**1.201**	5.719	0.597
2	Gaussian IID	60	8.523	7.210	11.250	**12.773**	11.016	14.530	**0.565**	0.391	0.745	5.512	4.867	6.219	0.574	**3.171**	**0.054**
3	MS-AIC	60	9.736	5.607	16.008	28.086	17.803	43.354	0.300	0.253	0.543	2.572	1.883	5.528	1.062	4.644	1.100
4	Expanding VAR-BIC	60	11.448	0.334	16.238	50.570	7.794	93.347	0.201	-0.071	0.469	0.280	-6.213	7.513	0.726	6.147	1.572
5	Expanding VAR-AIC	60	10.549	-0.249	15.114	49.258	7.659	90.856	0.188	-0.129	0.464	-0.820	-3.739	2.463	0.927	9.150	2.058
6	Rolling VAR-BIC	60	**20.607**	0.067	29.398	44.956	20.596	69.315	0.429	0.159	0.692	-2.326	-3.763	-1.017	-0.662	4.051	3.428
7	Rolling VAR-AIC	60	20.404	-0.139	29.181	44.968	20.841	69.095	0.425	0.135	0.690	-2.326	-3.670	-0.867	-0.608	3.375	3.477
1	Gaussian IID	1	6.425	2.775	10.087	**11.282**	10.369	12.195	0.454	0.277	0.623	**6.253**	2.431	10.286	0.172	**3.143**	**0.065**
2	MS-BIC	1	7.567	0.916	12.023	14.778	6.539	18.312	0.424	0.351	0.636	5.956	1.579	12.237	**1.344**	4.004	0.405
3	MS-AIC	1	5.709	-7.772	12.012	15.238	9.344	17.894	0.289	0.130	0.555	2.161	-27.966	32.862	0.070	4.641	0.928
4	Expanding VAR-BIC	1	12.710	-13.031	25.751	15.677	5.132	26.221	**0.728**	0.813	1.378	0.057	-6.498	7.034	0.402	6.093	2.005
5	Expanding VAR-AIC	1	9.651	-21.260	26.109	14.832	4.551	25.112	0.563	0.421	1.007	-0.729	-3.690	2.544	0.336	6.316	1.884
6	Rolling VAR-BIC	1	15.271	-31.411	40.679	22.057	-0.923	45.037	0.633	0.554	1.075	-1.804	-3.151	-0.340	0.668	14.910	1.189
7	Rolling VAR-AIC	1	**15.459**	-31.567	40.927	22.225	-0.616	45.067	0.637	0.548	1.097	-1.804	-3.186	-0.556	1.018	14.513	1.213

Note: in the table performance statistics are boldfaced when these are the best (maximum for mean, volatility, Sharpe ratio, CER, and skewness; minimum for kurtosis) among all the econometric models considered (including those not covered by the table). Because models are ranked in the table on the basis of their CERs, it is possible that the best model under other metrics may fail to appear in the table.

Table 11

Recursive Out-of-Sample Portfolio Performance from Alternative Classes of Models Ranked According to Average Certainty Equivalent Return (γ = 2 and 10)

γ = 2

CER Rank	Model	Horizon	Annualized Mean: Annualized mean returns	Annualized Mean: 95% Conf. Int. -- LB	Annualized Mean: 95% Conf. Int. -- UB	Annualized Volatility: Annualized volatility	Annualized Volatility: 95% Conf. Int. -- LB	Annualized Volatility: 95% Conf. Int. -- UB	Sharpe Ratio: Sharpe ratio	Sharpe Ratio: 95% Conf. Int. -- LB	Sharpe Ratio: 95% Conf. Int. -- UB	CER: CER (% Annualized)	CER: 95% Conf. Int. -- LB	CER: 95% Conf. Int. -- UB	Skew	Kurtosis	Average monthly turnover (adj.)
1	MS-BIC	60	**26.445**	-7.161	43.056	29.941	9.618	60.205	0.462	0.378	0.547	**10.844**	7.920	13.967	0.730	**2.904**	0.883
2	Gaussian IID	60	22.101	5.293	32.912	**25.634**	9.155	42.113	**0.811**	0.548	1.062	9.512	5.248	14.069	1.010	3.810	**0.279**
3	MS-AIC	60	25.241	14.229	38.148	33.888	20.196	47.580	0.706	0.516	0.903	9.108	8.717	9.506	1.252	4.826	1.582
4	Expanding VAR-BIC	60	15.710	7.103	20.962	50.518	6.159	94.878	0.285	0.097	0.459	4.968	4.251	5.684	1.807	11.690	2.698
5	Expanding VAR-AIC	60	10.977	1.814	16.506	53.133	3.505	115.973	0.182	0.020	0.337	4.911	4.265	5.574	**2.468**	18.837	2.037
6	Rolling VAR-BIC	60	17.691	0.703	27.555	47.479	18.696	76.262	0.345	0.094	0.570	4.156	4.050	4.264	1.701	9.054	3.548
7	Rolling VAR-AIC	60	17.321	0.250	27.183	47.387	18.253	76.521	0.338	0.094	0.592	4.154	4.044	4.267	1.694	8.813	3.947
1	Gaussian IID	1	10.217	-12.869	24.163	**11.209**	7.421	14.998	0.796	0.488	1.130	**9.945**	8.105	11.695	-0.047	**4.351**	**0.262**
2	MS-BIC	1	13.232	-18.501	35.275	25.068	4.465	48.416	0.476	0.168	0.781	7.945	4.111	11.823	-0.075	5.670	1.234
3	MS-AIC	1	11.824	-7.223	20.484	16.596	12.196	19.916	0.634	0.292	0.968	4.954	-0.075	10.724	-0.206	6.951	1.439
4	Expanding VAR-BIC	1	35.867	-8.982	62.119	46.319	16.191	76.447	0.746	0.624	0.867	3.316	1.349	5.302	2.651	16.638	2.512
5	Expanding VAR-AIC	1	47.641	-16.465	95.395	62.537	30.043	95.031	0.741	0.621	0.872	-5.093	-7.318	-3.011	2.034	9.040	2.510
6	Rolling VAR-BIC	1	83.918	-11.056	112.716	80.972	33.552	128.393	1.020	0.898	1.150	-25.359	-27.105	-23.434	**2.755**	17.632	4.653
7	Rolling VAR-AIC	1	**84.931**	-10.644	113.299	81.649	34.790	128.509	**1.024**	0.913	1.149	-25.453	-27.234	-23.648	1.824	7.981	4.712

γ = 10

CER Rank	Model	Horizon	Annualized Mean: Annualized mean returns	Annualized Mean: 95% Conf. Int. -- LB	Annualized Mean: 95% Conf. Int. -- UB	Annualized Volatility: Annualized volatility	Annualized Volatility: 95% Conf. Int. -- LB	Annualized Volatility: 95% Conf. Int. -- UB	Sharpe Ratio: Sharpe ratio	Sharpe Ratio: 95% Conf. Int. -- LB	Sharpe Ratio: 95% Conf. Int. -- UB	CER: CER (% Annualized)	CER: 95% Conf. Int. -- LB	CER: 95% Conf. Int. -- UB	Skew	Kurtosis	Average monthly turnover (adj.)
1	MS-BIC	60	4.961	3.749	6.086	**9.748**	9.221	11.123	**0.376**	0.215	0.602	**6.452**	5.518	7.262	0.563	3.310	0.309
2	Gaussian IID	60	6.011	5.361	6.670	14.291	13.149	15.400	0.330	0.189	0.467	4.734	4.132	5.375	0.881	**3.101**	**0.192**
3	MS-AIC	60	4.246	2.532	6.237	22.615	19.757	24.935	0.130	-0.001	0.334	2.515	1.960	3.151	0.583	3.666	0.486
4	Expanding VAR-BIC	60	9.125	2.990	12.892	38.479	9.724	67.234	0.203	0.101	0.316	0.645	0.459	0.825	**2.606**	11.911	0.921
5	Expanding VAR-AIC	60	12.774	4.028	18.030	53.824	24.436	83.212	0.213	0.124	0.299	0.625	0.490	0.749	2.493	11.013	1.034
6	Rolling VAR-BIC	60	23.939	10.911	31.844	67.879	45.116	90.641	0.334	0.233	0.441	0.582	0.517	0.652	1.551	6.598	1.521
7	Rolling VAR-AIC	60	**24.519**	11.213	32.741	69.663	46.728	92.599	0.333	0.232	0.424	0.581	0.518	0.653	1.523	6.514	1.564
1	Gaussian IID	1	3.694	1.860	5.558	**5.571**	5.137	6.011	0.430	0.270	0.607	**5.981**	2.655	9.209	0.299	4.356	**0.025**
2	MS-BIC	1	4.081	1.685	6.926	6.392	4.732	8.053	**0.435**	0.248	0.610	5.050	3.925	6.007	0.113	**3.936**	0.310
3	MS-AIC	1	5.946	-0.248	11.087	13.828	11.268	17.388	0.336	0.112	0.458	1.038	0.017	2.097	0.219	4.703	0.404
4	Expanding VAR-BIC	1	9.173	-9.925	19.576	19.457	12.311	26.603	0.405	0.238	0.561	0.063	-12.494	12.616	0.017	7.737	1.636
5	Expanding VAR-AIC	1	**9.465**	-5.022	19.619	14.150	8.381	19.919	0.577	0.416	0.749	-1.843	-16.323	11.420	-0.156	5.040	1.246
6	Rolling VAR-BIC	1	8.758	-12.898	16.638	18.400	9.292	27.508	0.405	0.232	0.557	-4.301	-9.222	0.541	0.773	8.994	2.119
7	Rolling VAR-AIC	1	8.992	-12.686	16.836	18.521	9.428	27.614	0.415	0.252	0.571	-6.616	-11.392	-1.962	**0.786**	7.510	2.179

Note: in the table performance statistics are boldfaced when these are the best (maximum for mean, volatility, Sharpe ratio, CER, and skewness; minimum for kurtosis) among all the econometric models considered (including those not covered by the table).

Table 12

Model Specification Search Extended to Heteroskedastic Markov Switching Models

Regimes (k)	VAR(p) order	Predictors Included							Hetero-skedastic	Negative Log-Likelihood	Linearity test	AIC	BIC	H-Q	No. Param.	Saturation ratio
		Div. Yield	Short Rate	Term	Default	Inflation	IP Growth	Unempl.								
Two-State Models																
2	0	N	N	N	N	N	N	N	No	5760.781	93.588***	-17.1035	-17.0096	-17.0671	14	143.79
2	1	N	N	N	N	N	N	N	No	5814.817	71.124***	-17.2364	-17.0214	-17.1531	32	62.91
2	Restr.	Y	Y	Y	Y	Y	Y	Y	No	5857.283	50.067***	-17.3272	-17.0316	-17.2127	177	37.91
2	1	Y	Y	Y	Y	Y	Y	Y	No	5917.327	170.155***	-17.4168	-16.9195	-17.2242	277	24.22
2	0	N	N	N	N	N	N	N	Yes	5831.461	237.145***	-17.2960	-17.1618	-17.2440	20	100.65
2	1	N	N	N	N	N	N	N	Yes	5900.596	223.462***	-17.4742	**-17.2189**	-17.3753	38	52.97
2	Restr.	Y	Y	Y	Y	Y	Y	Y	Yes	5929.339	194.180***	-17.5241	**-17.1881**	**-17.3940**	232	28.92
2	1	Y	Y	Y	Y	Y	Y	Y	Yes	5926.268	188.038***	-17.4255	-16.8880	-17.2173	332	20.21
Three-State Models																
3	0	N	N	N	N	N	N	N	No	5806.042	184.109***	-17.2174	-17.0764	-17.1628	21	95.86
3	1	N	N	N	N	N	N	N	No	5832.473	87.215***	-17.2950	-17.0934	-17.2169	30	67.10
3	Restr.	Y	Y	Y	Y	Y	Y	Y	No	5874.515	84.531***	-17.3577	-17.0150	-17.2250	191	35.13
3	1	Y	Y	Y	Y	Y	Y	Y	No	5992.346	320.192***	-17.5301	-16.7842	-17.2412	391	17.16
3	0	N	N	N	N	N	N	N	Yes	5886.030	356.074***	-17.5368	**-17.2650**	**-17.4209**	33	61.00
3	1	N	N	N	N	N	N	N	Yes	5939.016	319.066***	-17.5493	-17.1457	-17.3930	60	33.55
3	Restr.	Y	Y	Y	Y	Y	Y	Y	Yes	5964.330	264.162***	**-17.5897**	-17.1663	**-17.4257**	301	22.29
3	1	Y	Y	Y	Y	Y	Y	Y	Yes	6043.570	422.640***	**-17.6470**	-16.8205	-17.3269	501	13.39
Three-State MSVAR Models with Principal Components of Predictors																
3	1	First principal component							No	5916.483	246.035***	-17.4649	-17.0819	-17.3166	76	35.32
3	1	First two principal components							No	5918.162	234.000***	-17.4431	-16.9996	-17.2713	111	30.23
3	1	First three principal components							No	5924.246	240.398***	-17.4344	-16.9305	-17.2392	163	24.70
3	1	First principal component							Yes	5954.353	321.775***	-17.5420	-17.0784	-17.3625	96	27.96
3	1	First two principal components							Yes	5974.655	346.986***	**-17.5757**	-17.0516	-17.3727	141	23.79
3	1	First three principal components							Yes	5981.964	355.834***	-17.5707	-16.9861	-17.3443	163	24.70
Four-State Models																
4	0	N	N	N	N	N	N	N	No	5821.464	232.941***	-17.2622	-17.0606	-17.1841	30	67.10
4	1	N	N	N	N	N	N	N	No	5899.049	239.133***	-17.4121	-16.9681	-17.2401	66	30.50
4	Restr.	Y	Y	Y	Y	Y	Y	Y	No	5897.340	130.181***	-17.3989	-16.9958	-17.2428	207	9.72
4	0	N	N	N	N	N	N	N	Yes	5907.185	404.384***	-17.4640	-17.1415	-17.3391	48	41.94

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

Table 13

Full-Sample Estimates of Three-State Heteroskedastic Markov Switching Multivariate Model for Real Stock, Bond, and 1-month T-Bill Returns

	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns
Panel A - SINGLE STATE MODEL			
1. Mean returns	0.0992**	0.2340**	0.5855**
2. Correlations/Volatilities			
Real 1-month T-bill Returns	0.3171**		
Real Long-Term Bond Returns	0.2869*	2.2257**	
Real Stock Returns	0.1228	0.1638*	4.3405**
Panel B - THREE-STATE MODEL			
	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns
1. Mean returns			
Bear State	-0.0560*	-0.0390	-0.1506
Equity Bull/Low Volatility State	0.1249**	0.2066*	1.1074**
Bond Bull State	0.4702**	1.2043**	0.6018
2. Correlations/Volatilities			
Bear State			
Real 1-month T-bill Returns	0.3291**		
Real Long-Term Bond Returns	0.2042*	2.4192**	
Real Stock Returns	0.0734	-0.0179	5.1063**
Equity Bull/Low Volatility State			
Real 1-month T-bill Returns	0.2085*		
Real Long-Term Bond Returns	0.1873	1.5593**	
Real Stock Returns	0.2153*	0.2683*	3.4589**
Bond Bull State			
Real 1-month T-bill Returns	0.3297**		
Real Long-Term Bond Returns	0.3979**	3.4523**	
Real Stock Returns	-0.0321	0.4327**	4.8150**
3. Transition probabilities	Bear State	Equity Bull/Low Volatility State	Bond Bull State
Bear State	0.9514**	0.0359*	0.0127
Equity Bull/Low Volatility State	0.0284*	0.9714**	0.0002
Bond Bull State	0.0265	0.0018	0.9717**
Panel C - MARKOV CHAIN PROPERTIES, THREE-STATE MODEL			
	Bear State	Equity Bull/Low Volatility State	Bond Bull State
Ergodic Probabilities	0.3648	0.4689	0.1662
Average Duration (in months)	20.6	35.0	35.3

** = significant at 1% size or lower; * = significant at 5% size.

Figure 1

Preliminary Evidence of Serial and Cross-Serial Correlation for Real Asset Returns

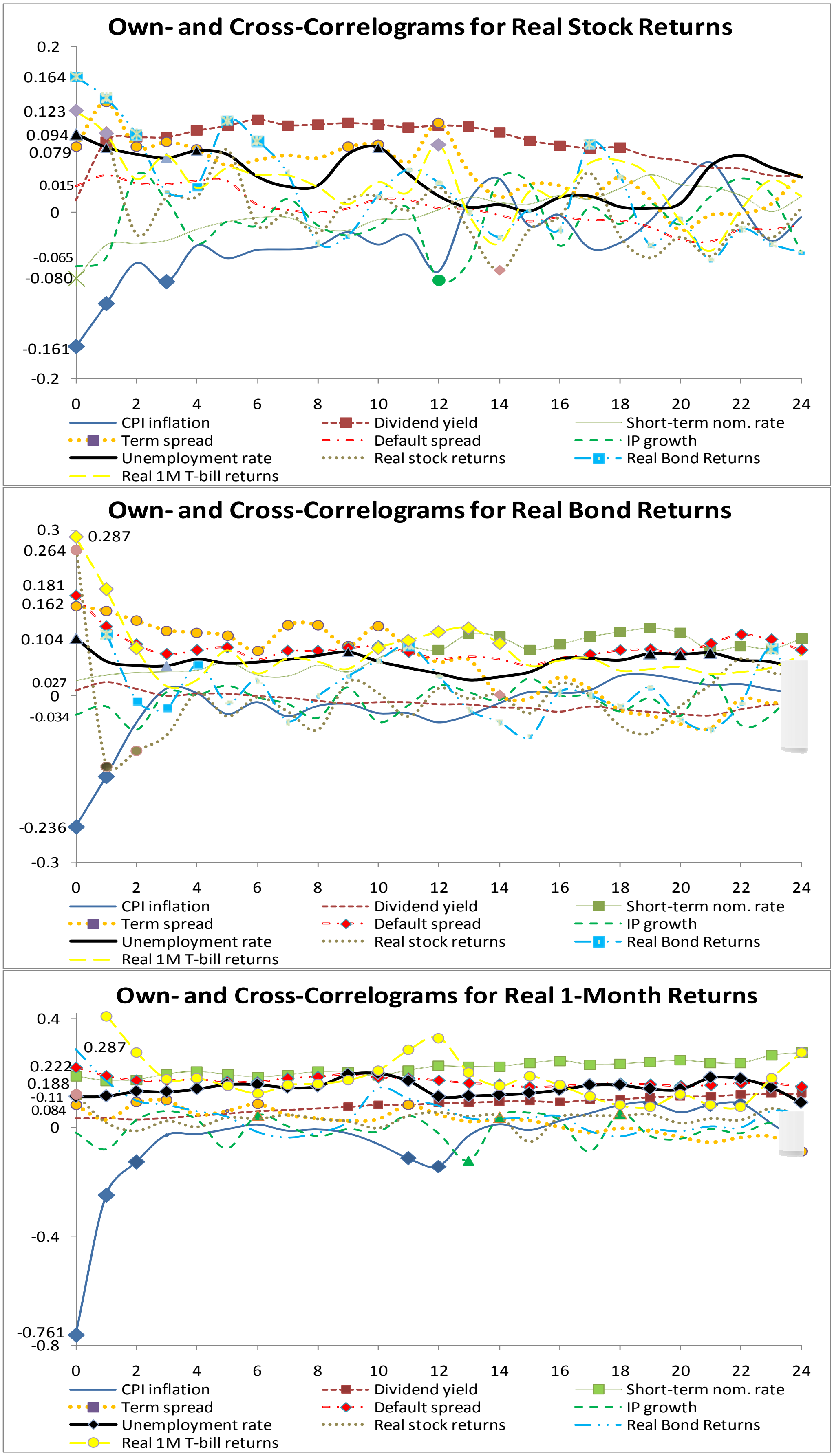

Figure 2

Smoothed Regime Probabilities from Three-State Homoskedastic Markov Switching Model

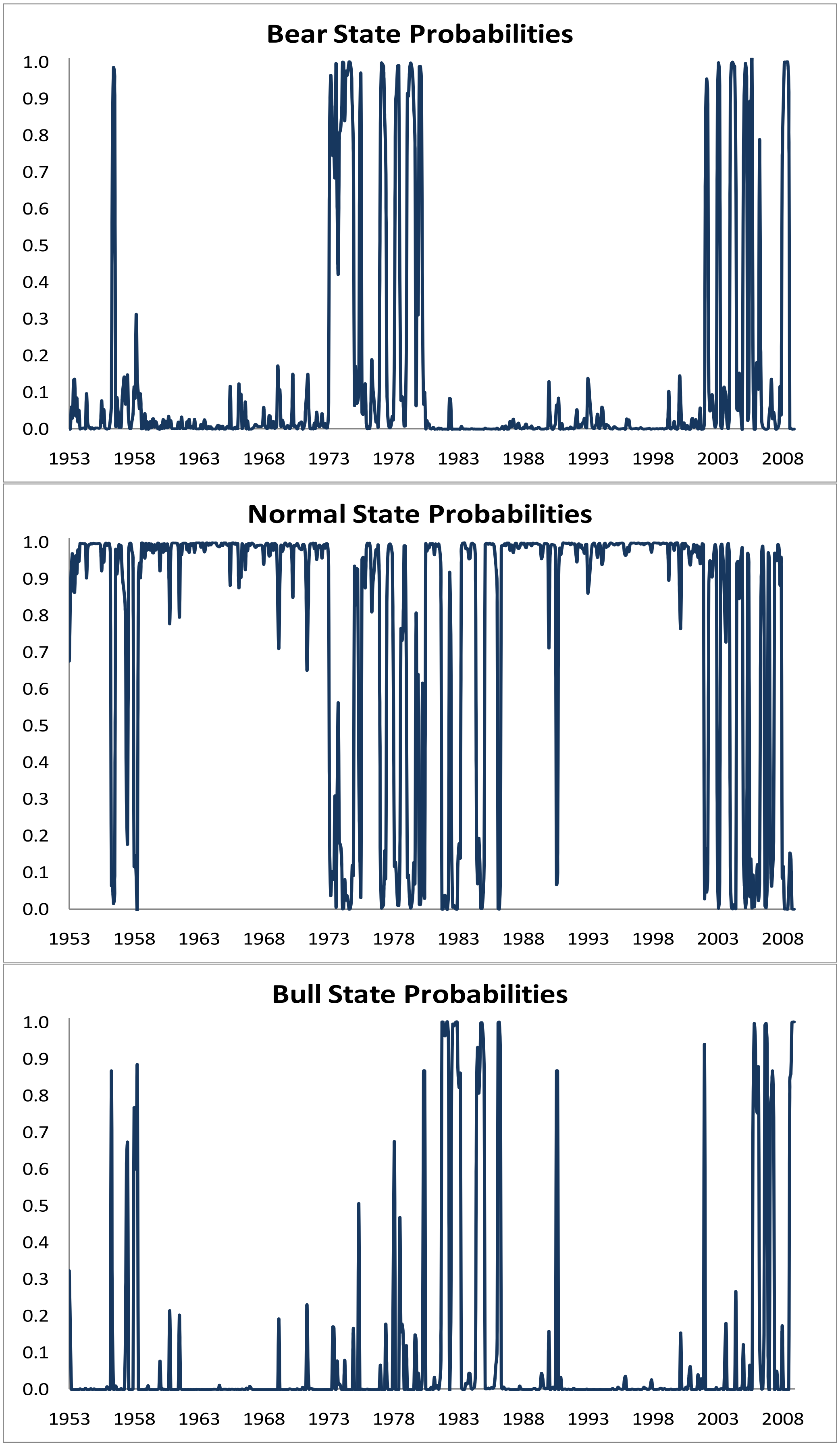

Figure 3

Recursive Mean Estimates from Three-State Homoskedastic Markov Switching Model

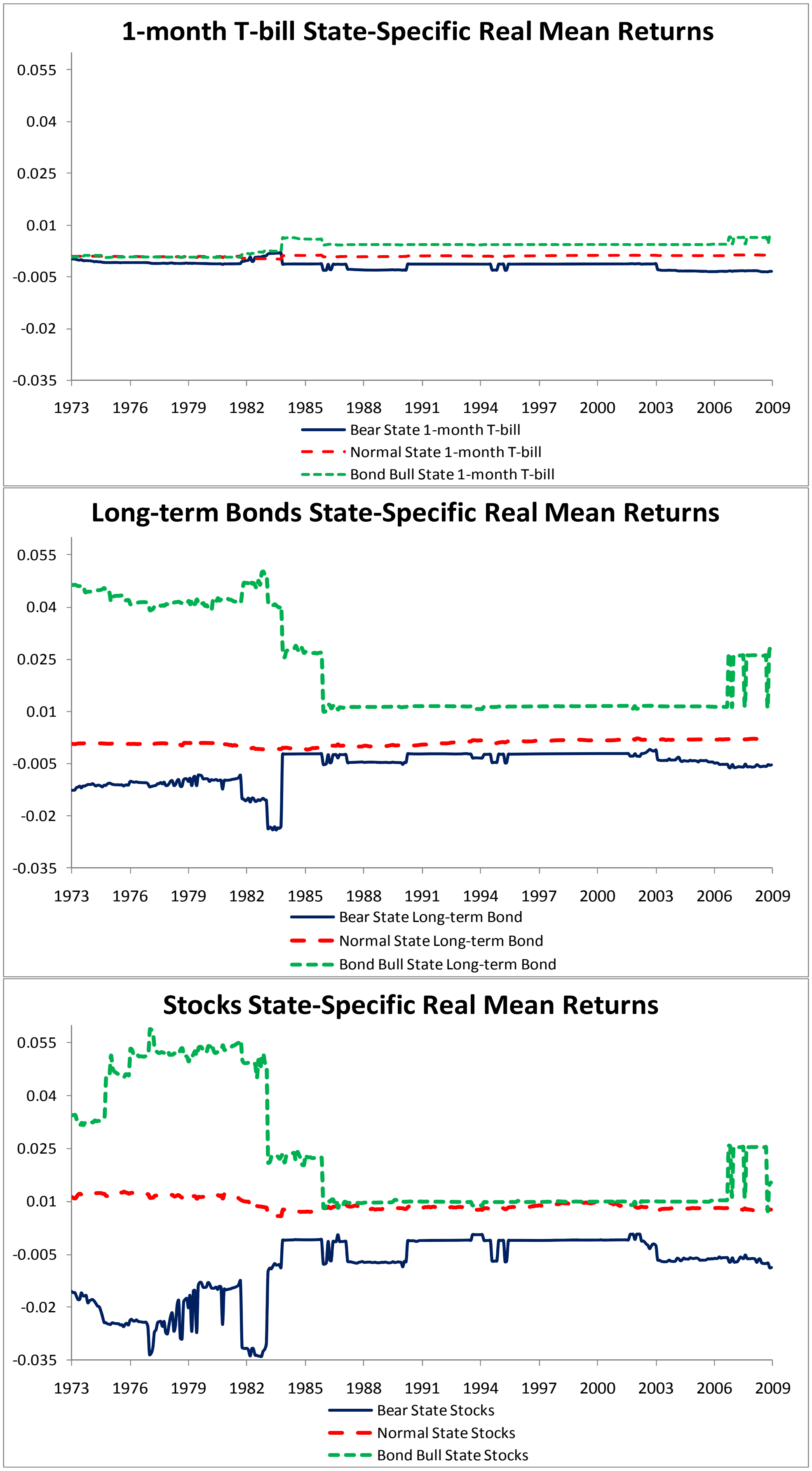

Figure 4

Recursive Coefficient Estimates from VAR(1) Linear Predictability Model: Real Stock Returns

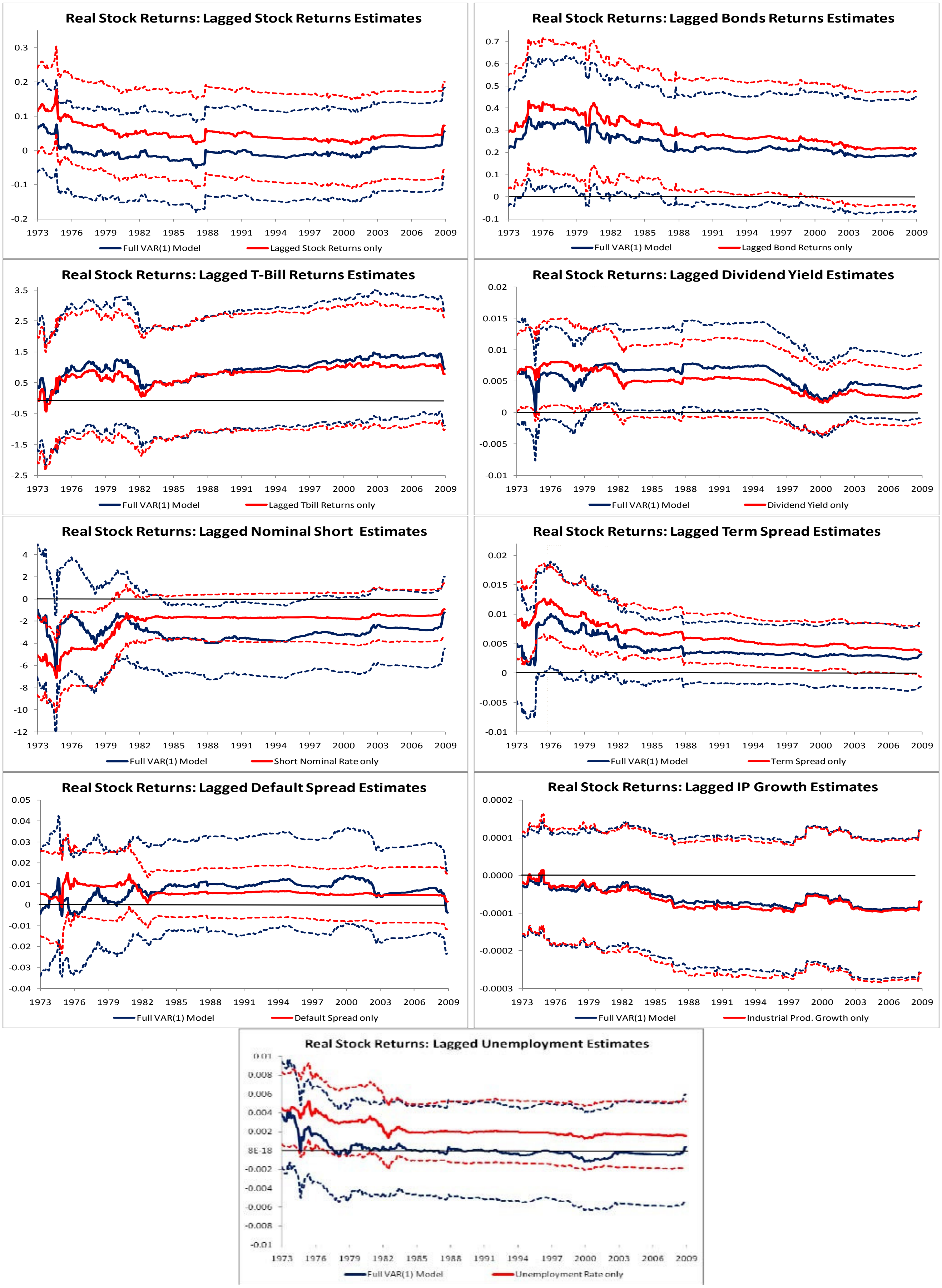

Figure 5

Recursive Coefficient Estimates from VAR(1) Linear Predictability Model: Real Bond Returns

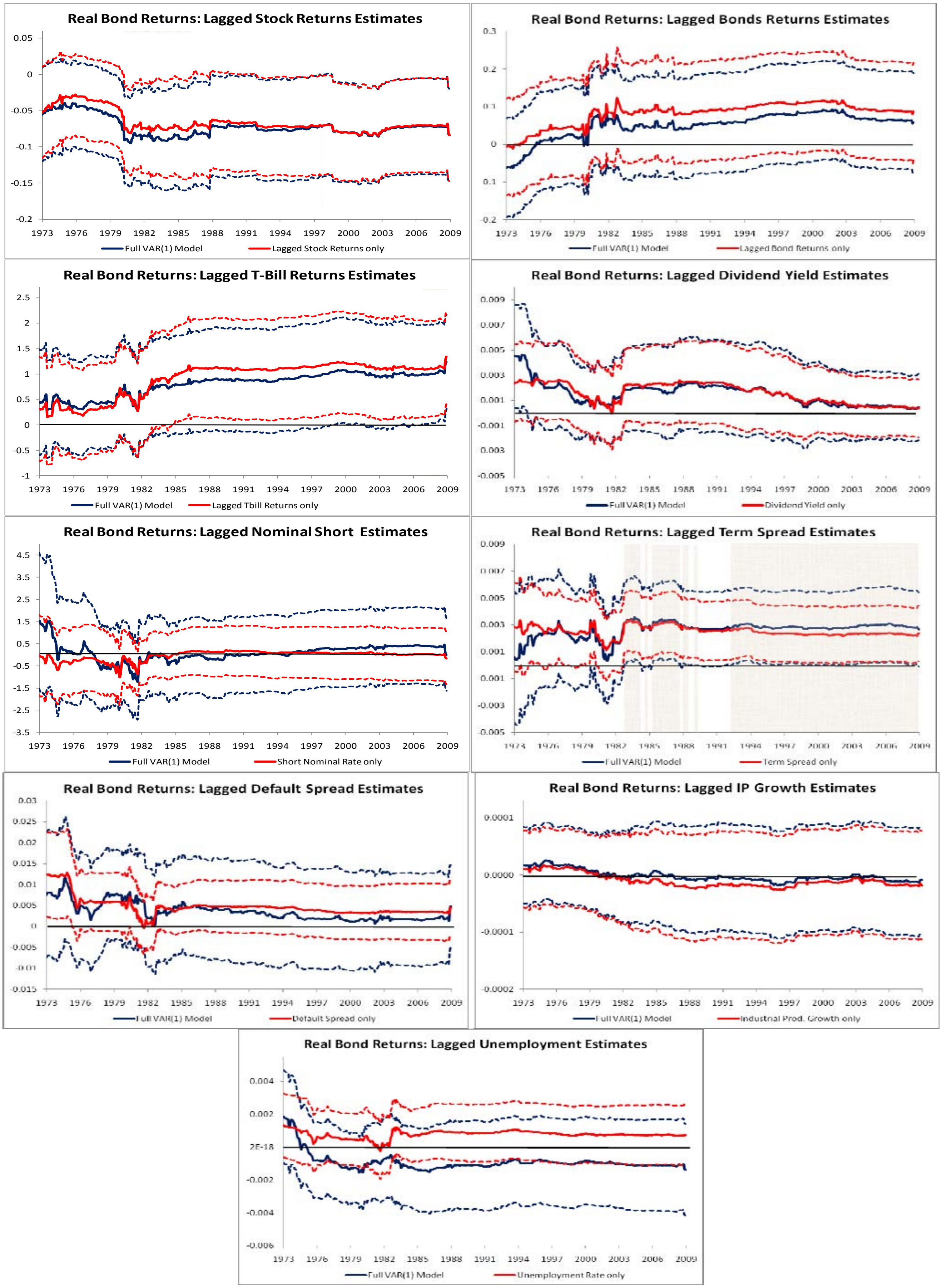

Figure 6

Recursive Coefficient Estimates from VAR(1) Linear Predictability Model: Real Bond Returns

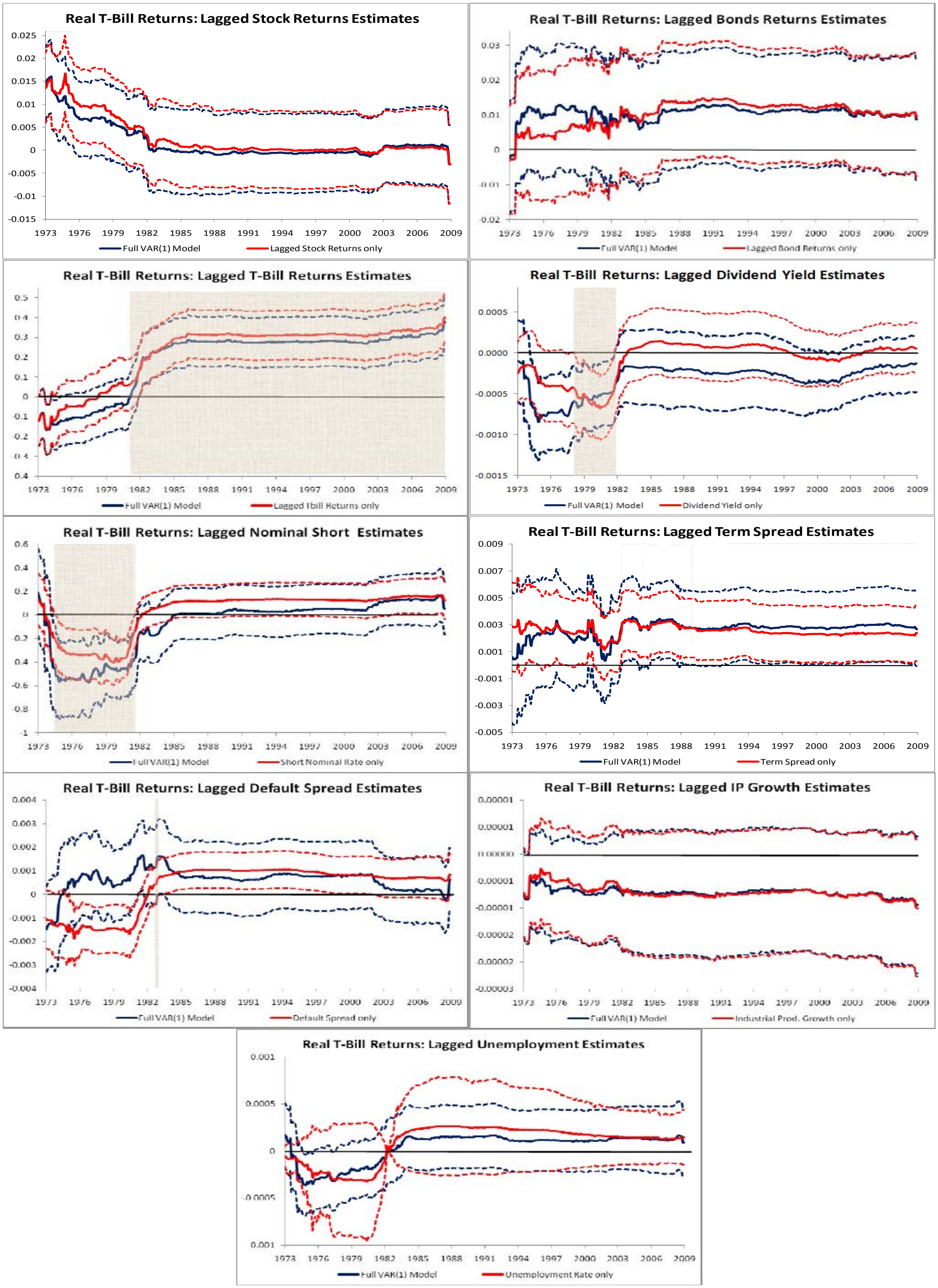

Figure 7

Smoothed Regime Probabilities from Three-State Homoskedastic MSVAR(1) Model

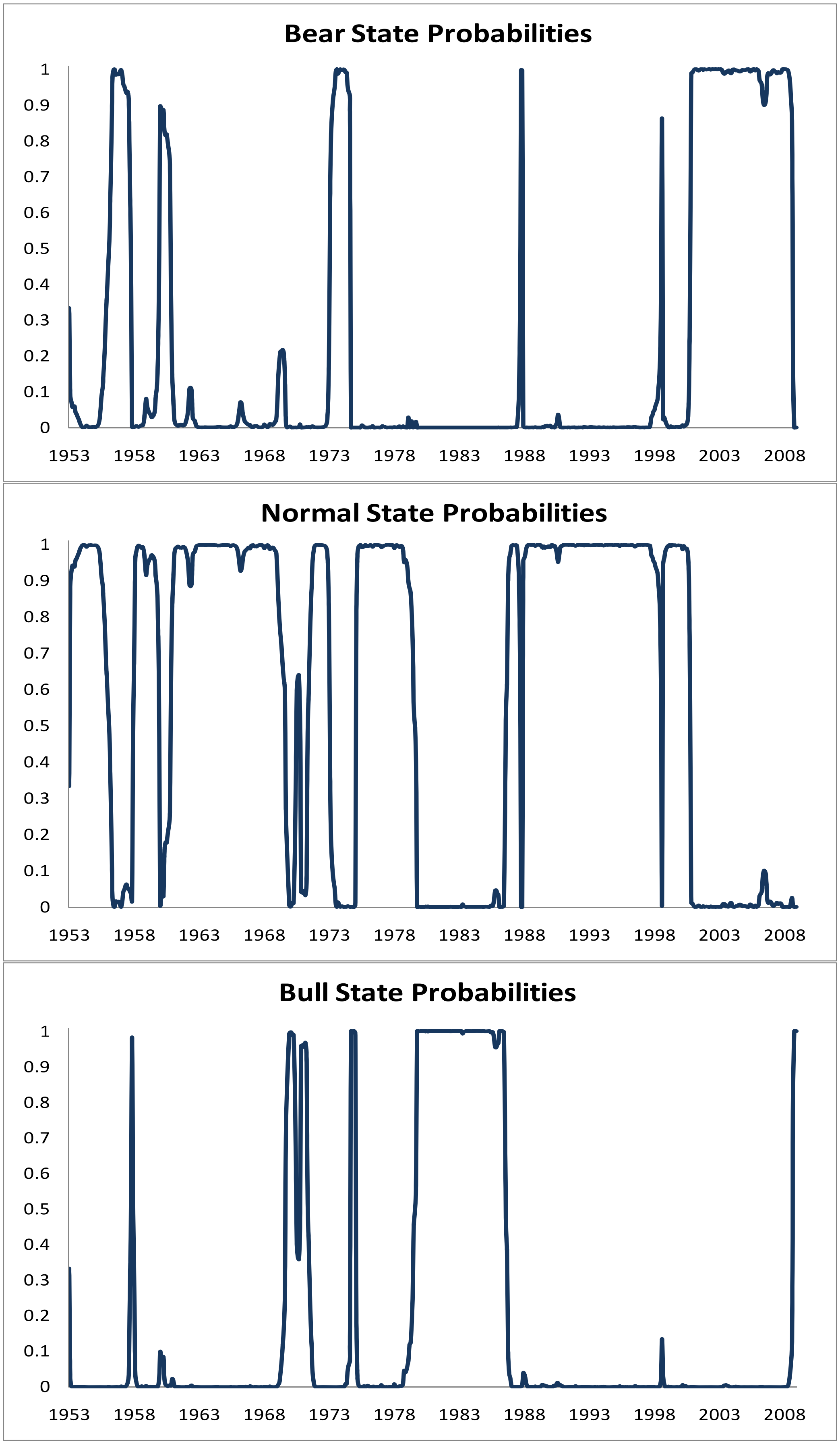

Figure 8

Dynamics of Portfolio Weights under Markov Switching vs. Full VAR(1), $\gamma = 5$

Markov Switching Homoskedastic | **Full VAR(1)**

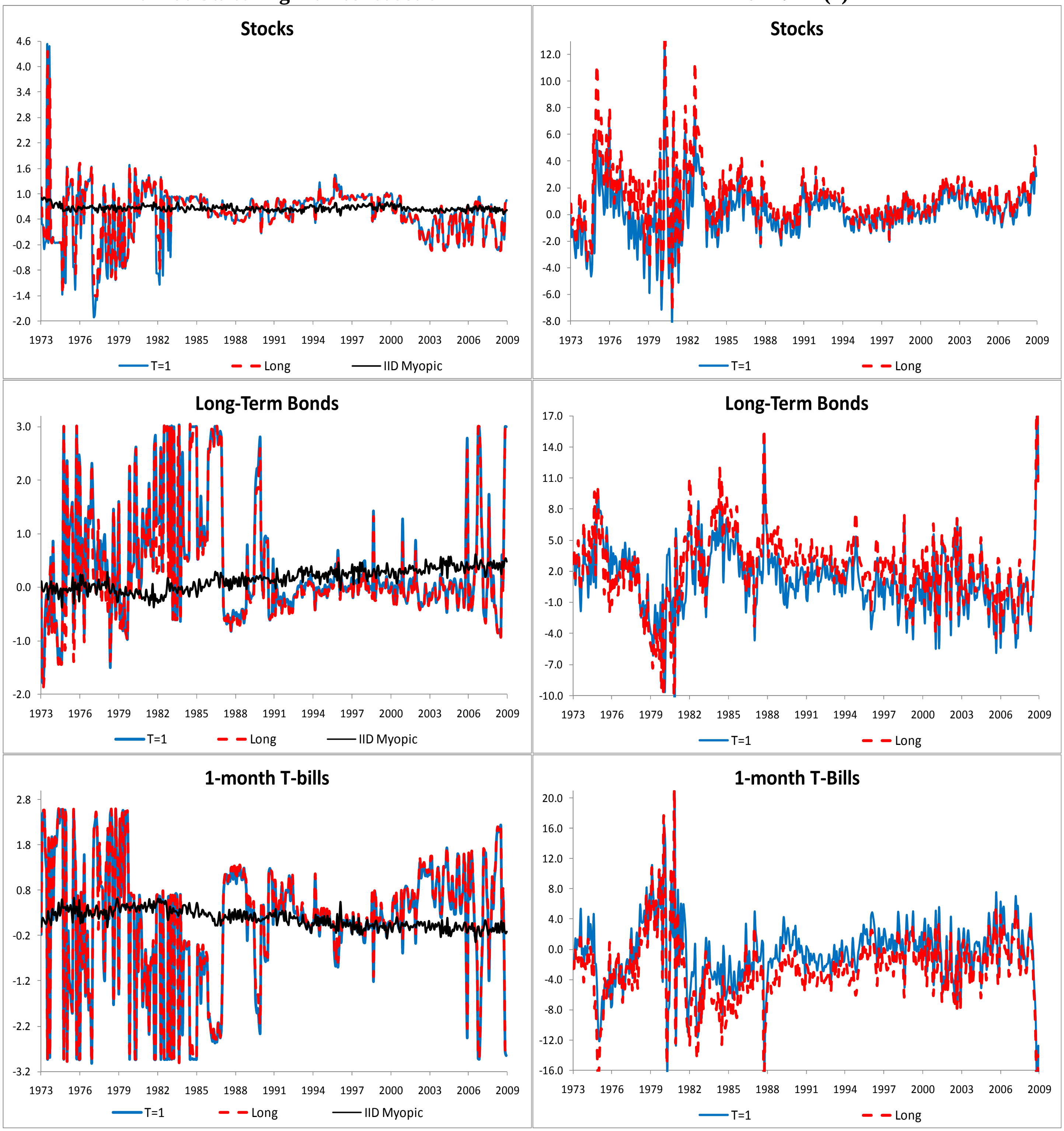

Figure 9

Dynamics of Portfolio Weights under MS VAR(1) vs. Best Performing VAR(1), γ = 5

Markov Switching VAR with PC1 | **Best CER VAR (VAR(1)-DY)**

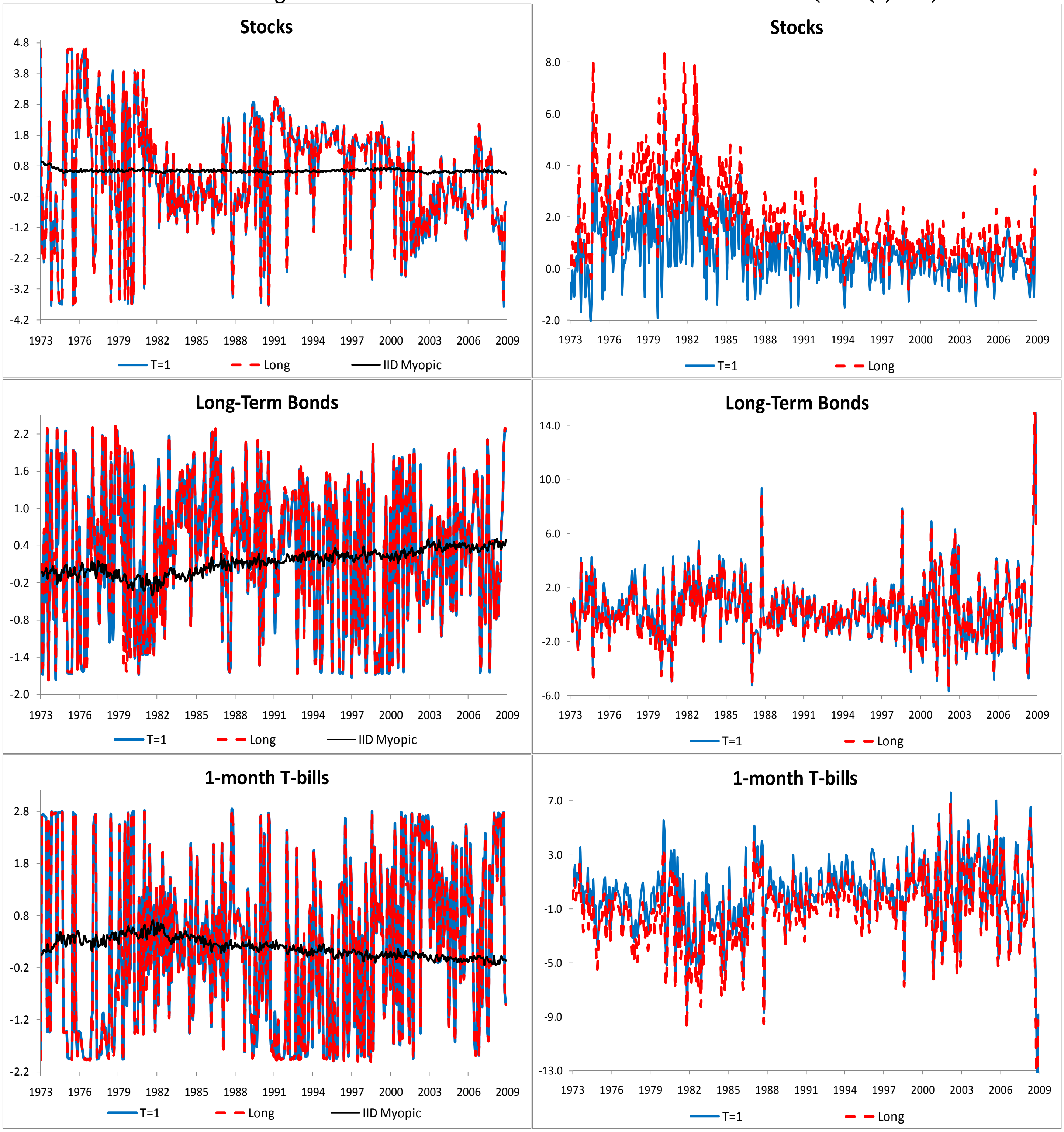

Figure 10

Dynamics of Hedging Demands under Markov Switching vs. Full VAR(1), $\gamma = 5$

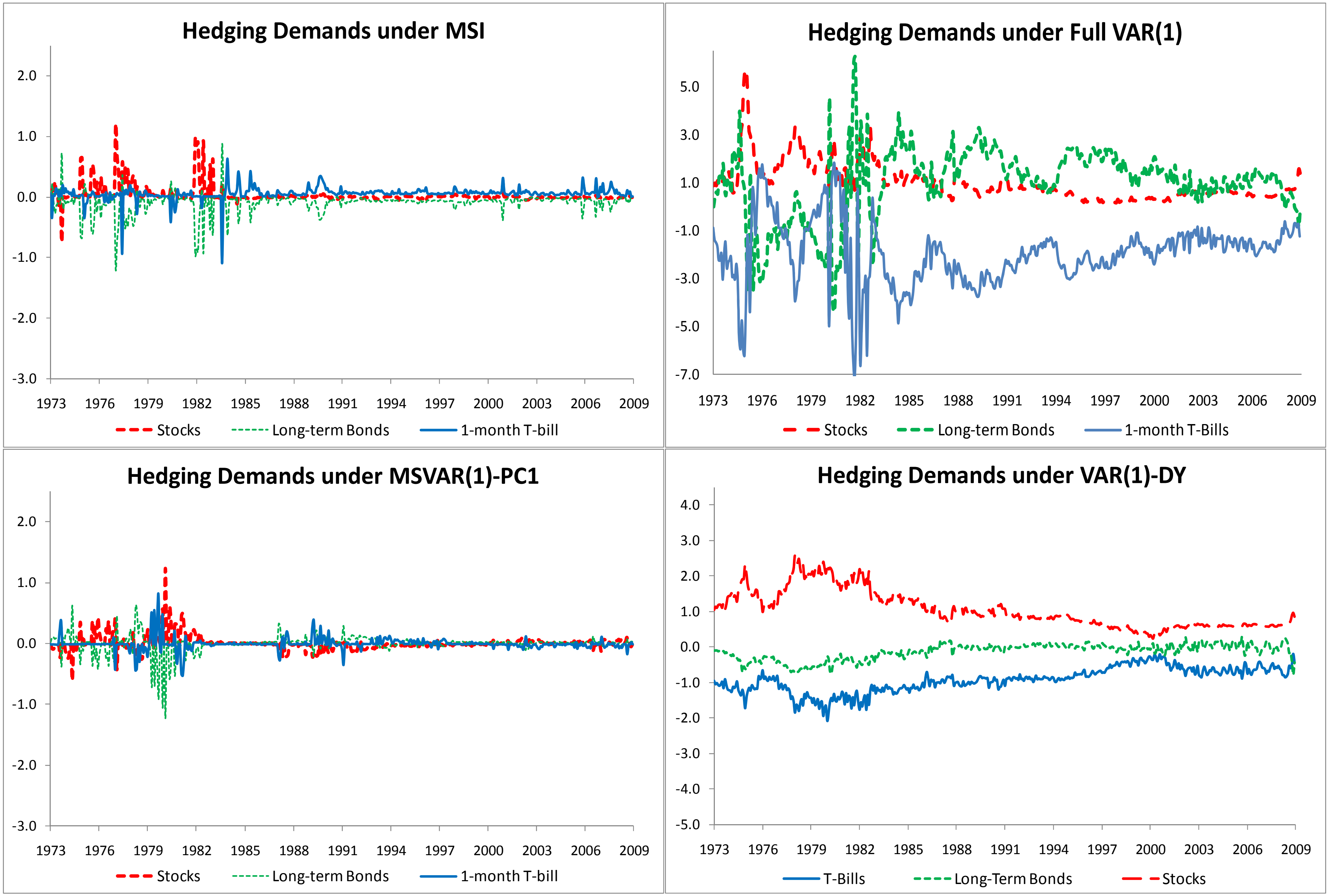

Figure 11

Smoothed Regime Probabilities from Three-State Heteroskedastic Markov Switching Model

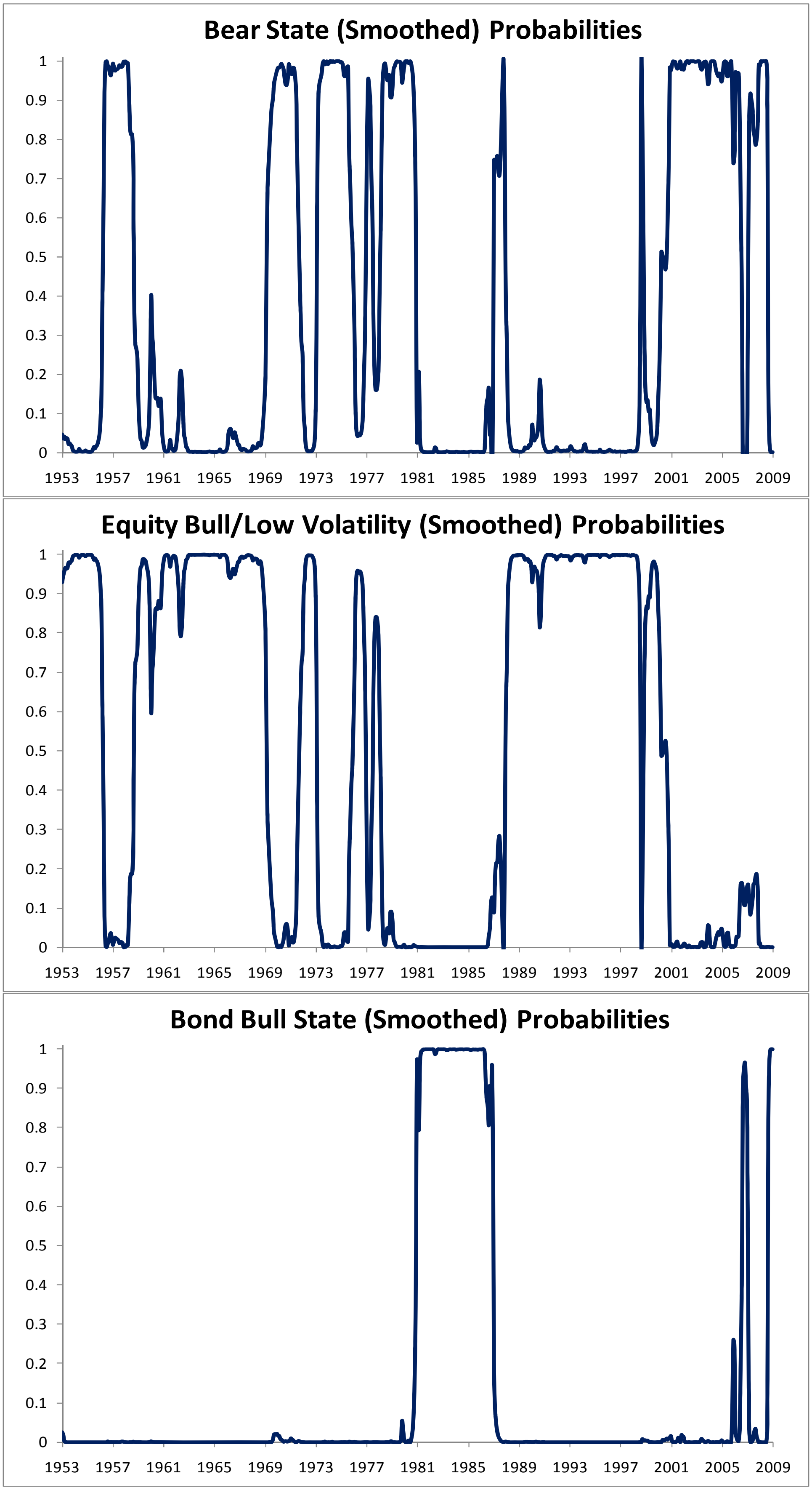

Figure 12

Dynamics of Portfolio Weights under Heteroskedastic Markov Switching Model, $\gamma = 5$