Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Can VAR Models Capture Regime Shifts in Asset Returns?

A Long-Horizon Strategic Asset Allocation Perspective

Massimo Guidolin, Stuart Hyde

Working Paper n. 414

This Version: August 10, 2011

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Can VAR Models Capture Regime Shifts in Asset Returns?

A Long-Horizon Strategic Asset Allocation Perspective*

Massimo Guidolin
Bocconi University, IGIER & CAIR, MBS

Stuart Hyde†
Manchester Business School

August 10, 2011

Abstract

It is often suggested that through a judicious choice of predictors that track business cycles and market sentiment, simple Vector Autoregressive (VAR) models could produce optimal strategic portfolio allocations that hedge against the bull and bear dynamics typical of financial markets. However, a distinct literature exists that shows that nonlinear econometric frameworks, such as Markov switching (MS), are also natural tools to compute optimal portfolios in the presence of stochastic good and bad market states. In this paper we examine whether simple VARs can produce portfolio rules similar to those obtained under MS, 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, 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 nonlinear models for a long-horizon investor with constant relative risk aversion. We conclude that most VARs cannot produce portfolio rules, hedging demands, or (net of transaction costs) out-of-sample performances that approximate those obtained from equally simple nonlinear frameworks. We also compute the improvement in realized performance that may be achieved adopting more complex MS models and report this may be substantial in the case of regime switching ARCH.

JEL codes: G11, C53.

.

*We would like to thank Carol Alexander, "Paul" Moon Sub Choi (a discussant), Paolo Colla, Carlo Favero, René Garcia, Abraham Lioui, Patrick Minford, Francesco Saita, Andrea Sironi, Peter 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, Essex School of Business, and University Reading, ICMA Centre.

†Correspondence to: Prof. Stuart Hyde, Manchester Business School, University of Manchester, MBS Crawford House, Booth Street East, Manchester, UK M13 9PL. Tel: 44 (0) 161 275 4017. Fax: 44 (0) 161 275 4023. E-mail: stuart.hyde@mbs.ac.uk

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 should a risk-averse investor weight each available asset) has almost exclusively devoted its attention to linear predictability models. In a linear predictability model, asset returns are simply forecast by past values of predictor variables (such as the dividend yield and the term spread) within a vector autoregressive (VAR) framework. The linearity consists of the fact that 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 nonlinear dynamics (such as regimes, thresholds, self-exciting mean reversion, 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 exercises has relied on two often-implicit premises. First, that although most normative papers have to be taken only as examples of how practical portfolio choice ought to proceed, 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 the literature (see e.g. Barberis, 2000, and Lynch, 2001), *some* more complicated VAR must exist that is of practical use in terms of consistently improving realized performances. Hence some VARs can be found that can efficiently summarize the overall balance of predictability in asset returns and making the modeling of any residual nonlinear 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, simple, small-scale VAR(1) models must be illustrative already of the first-order effects of linear predictability on dynamic portfolio selection. 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 may deliver dynamic portfolio choices that are: (i) able to approximate the portfolio choices of an investor that exploits both linear and nonlinear predictability patterns in the data, and (ii) competitive in terms of realized portfolio performance.

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") nonlinear benchmark, represented by a plain vanilla 3-state Markov switching (MS) model. This is of course only an ex-ante perspective on the problem: "different" does not imply "worse" in the view of an applied portfolio manager and what could be misspecified and practically not useful is not the VAR family, but the proposed nonlinear benchmark. More importantly, VARs systematically fail to perform better than nonlinear models in recursive (pseudo) out-of-sample tests, in the sense that VARs generally produce lower realized certainty equivalent

[1] The literature on non-linearities in finance is rather voluminous. A few basic elements are discussed in the books of Campbell et al. (1997) and in Granger and Teräsvirta (1993). Fewer papers have 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), Frauendorfer et al. (2007), Gomes (2007), Guidolin and Timmermann (2008a).

returns (i.e., risk-adjusted performances that take into account the curvature of the utility function under which the portfolio choice program has been solved) than multi-state models. This means that VARs provide no approximation tool for more complicated, nonlinear dynamics either ex-ante or ex-post.

We stress that this result that simple VARs cannot capture all predictability patterns typical of well-known, standard financial U.S. data even when all possible combinations of predictors and lag choices are allowed, is obtained with reference to a very simple MS model. Such a choice is motivated by the search for a "lower bound": if, in the presence of nonlinearities in commonly used data, VARs are not even capable to achieve a "tie" (say, in terms of realized out-of-sample performance) with a "naive"MS model that may itself not perfectly capture the dynamic features of the data, then it may be safe to conclude that the role of nonlinear models (not only MS, of course) ought to be considerably larger than the one they have played so far in the empirical finance literature. To illustrate this fact, we try to tease out from our empirical exercise a measure of the realized utility (certainty equivalent return, CER) gains that a risk-averse investor may derive from adopting realistic MS models that may able to provide a good fit to the dynamics of real asset returns. Here we consider a variety of MS models that include MS VARs, MSVAR ARCH models, and MS models with time-varying transition probabilities. Our key result is that the risk-adjusted performance gains may somewhat exceed the plain vanilla MS findings commented early on. In particular, MS ARCH models may yield substantial CER improvements.

These results are obtained with reference to a strategic asset allocation (SAA) application that appears to have played a key role in the literature (see, e.g., Brennan et al., 1997): a risk-averse (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 1953-2009 which also includes the recent financial crisis.[2] We focus on long-horizon portfolio choices (up to a 5-year horizon) of an investor that recursively solves a portfolio problem in which utility derives from real consumption (i.e., cash flows obtained from dividend and coupon payments and from selling securities in the portfolio) and rebalancing is admitted at the same frequency as the data. This means that even when the problem solved is characterized by a 60-period ahead horizon, 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 become available at all these future points. Finally, our investor selects optimal portfolio weights taking into account the presence of both fixed and variable transaction costs. This means that – because a given vector of optimal weights at time t may implicitly imply a need to trade in all assets between time t and $t+1$ – our investor will also take into account the trading needs of her portfolio choices and especially the impact of the transaction costs incurred on expected utility.

Such a portfolio problem seems to be the 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 – the relative economic value of different types of models. 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

[2]Section 6.3 extends our results to U.K. data collected over a similar sample period.

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 choice in all subsequent periods. We are not aware of previous papers that have jointly solved consumption and portfolio choice problems under MS dynamics. Third, as previously stressed by Balduzzi and Lynch (1999) and Lynch and Balduzzi (2000), all SAA problems under predictability and active portfolio management ought to carefully consider whether the forecastable variation in investment opportunity sets offers enough welfare gains to exceed the often large trading costs.

The thrust of our exercise does not consist of investigating the different portfolio implications recursive portfolio weights implications of linear vs. nonlinear models, as this operation has already appeared in the literature for specific linear and nonlinear frameworks (see e.g., Ang and Bekaert, 2004, Detemple et al., 2003, Guidolin and Timmermann, 2007).[3] These papers try and measure the economic loss from model misspecification in (density) forecasting applications by resorting to portfolio choice metric. 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, nonlinear 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). We investigate 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 1,024 different VARs), and experimenting with 4 alternative lag orders throughout, $p =$ 1, 2, 4, and 12 (with restrictions). The seven predictors are typical in the finance literature and consist of widely employed 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.

The rest of the paper is structured as follows. Section 2 describes the research design. Section 3 describes the data, the 3-state Markov switching benchmark, and some features of linear predictability. Section 4 computes and presents optimal portfolio weights and hedging demands under the two classes of models. Section 5 computes realized, recursive out-of-sample portfolio performances. Section 6 performs a few robustness checks. Section 7 extends the set of MS models to encompass MS VAR models, MSVAR ARCH models (with leverage), and MS models with time-varying transition probabilities. Section 8 concludes.

2. **Methodology**

2.1. *Econometric Models*

We perform recursive estimation, portfolio weight calculation and performance evaluation for three groups of models. First and foremost, we entertain a large class of VAR(p) models. These VARs consist of a linear

[3]However, the number of genuine recursive out-of-sample horse races between linear vs. nonlinear models is still limited is limited to a few experiments in which comparisons are performed only with single-state benchmarks (i.e., when ignoring regimes means to ignore predictability altogether, as in Ang and Bekaert, 2004) or with hand-picked VAR(1) models (as in Guidolin and Timmermann, 2007, 2008b).

relationship linking $\mathbf{r}_{t+1}$, a $N \times 1$ vector of risky *real* asset returns 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

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

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). In this paper we consider multiple values of p, $p = 1$, 2, 4, and 12. Note that—conditioning on a choice to always include the lagged values of real asset returns in (1), as in Campbell et al. (2003)—for a given value of p there are 2^M different VARs one can obtain according to which of the M predictors are included in $[\mathbf{r}'_{t+1}\ \mathbf{y}'_{t+1}]'$. In our case, because we shall be using $M = 7$ alternative predictors and accounting for the restrictions on estimation when $p = 12$, this gives a total of 392 different VAR models.

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

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

where the latent, unobservable Markov state is $S_{t+1} = 1, ..., k$ and $\boldsymbol{\mu}$ is a $N \times 1$ vector of state-dependent intercepts. We also allow for the $N \times N$ covariance matrix of residuals $\boldsymbol{\Omega}$ to be state dependent, implying that the variance of the asset returns is also state-dependent, i.e., $Var\left[\mathbf{r}_{t+1}|S_{t+1}\right] = \boldsymbol{\Omega}_{S_{t+1}}$. Even though (2) is a very simple MS model, it serves us well in view of a goal to place a lower bound on the economic value of MS models, because it should be possible to find more complex and better specified nonlinear frameworks that may fit and forecast our asset return data better than (2). However, Section 7 explicitly examines the realized portfolio performance of a few additional, but inevitably more complex MS models, such as time-heterogeneous MS models in which the transition matrix $\mathbf{P}$ is allowed to change over time, and MS-ARCH models, in which given the basic structure of (2), $\boldsymbol{\Omega}_{t+1}$ becomes a function of both the time t information set $\mathcal{F}_t$ and of the regime variable S_{t+1}.

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

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

which is obviously the single-state restriction of (2). Of course, (3) implies that real asset returns are not predictable, while (2) implies that real asset returns may follow nonlinear predictability patterns, driven by the fact that the latent state S_{t+1} may display a predictable Markov structure.

2.2. *The Portfolio Choice Problem*

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

$$\max_{\{C_{t+\tau}, \omega_{i,t+\tau}\}_{\tau=0}^{H-1}\ {}_{i=1}^{N}} \sum_{\tau=0}^{H-1} \beta^{\tau} E\left[\frac{C_{t+\tau}^{1-\gamma}}{1-\gamma} |\mathbf{Z}_t\right] \qquad \beta \in (0,1)\,,\ \gamma > 1, \tag{4}$$

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%, approximately the mean of the real 1-month T-bill rate in our sample), the coefficient γ measures relative risk aversion, C_t is the investor's (real) consumption at time t and $\mathbf{Z}_t$ is the relevant vector of state variables at time t.[4] 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 given by

$$W_{t+1} = (W_t - C_t)(1 + r_{p,t+1}) = W_t (1 - \kappa_t) R_{p,t+1}, \tag{5}$$

where $R_{p,t+1}$ is the gross real 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 (4) and the budget constraint (5) is:

$$\frac{a(\mathbf{Z}_t, t) W_t^{1-\gamma}}{1-\gamma} = \max_{\kappa_t, \boldsymbol{\omega}_t} \left\{ \frac{\kappa_t^{1-\gamma} W_t^{1-\gamma}}{1-\gamma} + \frac{\beta (1-\kappa_t)^{1-\gamma} W_t^{1-\gamma}}{1-\gamma} E\left[a(\mathbf{Z}_{t+1}, t+1) R_{p,t+1}^{1-\gamma} | \mathbf{Z}_t \right] \right\}, \tag{6}$$

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:

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

Equation (7) can be solved by backward iteration, starting with $t = H-1$ and setting $a(\mathbf{Z}_H, H) = 1$ and then computing $a(\mathbf{Z}_H, t)$ by solving the optimization problem in (7) 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 they 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 have been proposed in the literature. We employ Monte Carlo methods for integral (expected utility) approximation.[5]

2.3. *Transaction Costs*

We assume that the consumer faces transaction costs that are proportional to wealth, so that her law of motion for wealth is

$$W_{t+1} = (W_t - C_t)(1 - f_t)(1 + r_{p,t+1}), \tag{8}$$

where f_t is transaction cost per dollar of wealth. The law of motion for wealth in (8) implicitly assumes that consumption at time t and any transaction costs to be paid at time t are obtained by liquidating costlessly

[4] In the case of a VAR(p), $\mathbf{Z}_t \equiv [\mathbf{r}_t' \ \mathbf{y}_t' \ \mathbf{r}_{t-1}' \ \mathbf{y}_{t-1}' \ \dots \ \mathbf{r}_{t-p+1}' \ \mathbf{y}_{t-p+1}']'$ so that the state vectors consists of a combination of lagged values of asset returns and predictor variables. In a Markov switching framework $\mathbf{Z}_t$ consists instead of the vector of state probabilities estimated at time t.

[5] An appendix not for publication provides additional details on the numerical methods used in the solution of the portfolio problem. See also Guidolin and Timmermann (2007, 2008b).

the risky and the riskless assets in the proportions $\{\omega_{i,t}\}_{i=1}^{N}$. This assumption is sensible for liquid assets, especially when they pay coupons or dividends that can be readily used to pay for transaction costs. In particular, we assume that there is both a fixed and a variable component to transaction costs. Therefore f_t is modeled as a function of the difference between the end- and the beginning-of-period wealth allocation to the assets, $\{\omega_{i,t} - \omega_{i,t-1}\}_{i=1}^{N}$

$$f_t = \tau_f I_{\{\exists i \ni \omega_{i,t} \neq \omega_{i,t-1}\}} + \tau_v \sum_{i=1}^{N} |\omega_{i,t} - \omega_{i,t-1}|, \tag{9}$$

where $I_{\{\exists i \ni \omega_{i,t} \neq \omega_{i,t-1}\}} = 1$ when the condition $\omega_{i,t} \neq \omega_{i,t-1}$ is satisfied for at least one $i = 1, ..., N$ (i.e., there is trading in asset i between $t-1$ and t), and 0 otherwise. The first term is a fixed fraction of the total value of the portfolio that represents the fixed cost of rebalancing the portfolio, regardless of the size of the rebalancing. The second term is proportional to the change in the value of the asset holdings. Interestingly, under the new dynamic budget constraint (8), the inherited portfolio allocation from the previous period, $\{\omega_{i,t-1}\}_{i=1}^{N}$, becomes a state variable when either τ_c or τ_v (or both) is greater than zero, since its value determines the transaction costs to be paid at time t. Similarly to Balduzzi and Lynch (1999) we initially set $\tau_f = 0.1\%$, $\tau_v = 0.5\%$.

Under (8), the Bellman equation of the problem becomes:

$$\frac{a(\mathbf{Z}_t, \boldsymbol{\omega}_{t-1}, t)}{1-\gamma} = \max_{\kappa_t, \boldsymbol{\omega}_t} \left\{ \frac{\kappa_t^{1-\gamma}}{1-\gamma} + \frac{\beta(1-\kappa_t)^{1-\gamma}}{1-\gamma} E\left[a(\mathbf{Z}_{t+1}, \boldsymbol{\omega}_t, t+1) R_{p,t+1}^{1-\gamma} |\mathbf{Z}_t \right] \right\}. \tag{10}$$

where $\boldsymbol{\omega}_{t-1}$ is a $N \times 1$ vector that collects the starting portfolio weights. Also in this case, the Bellman equation may be solved by backward recursion, using Monte Carlo methods. The only difference with respect to the case of $\tau_f = \tau_v = 0$ is that a Monte Carlo approximation of the expectation $E[a(\mathbf{Z}_{t+1}, \boldsymbol{\omega}_t) R_{p,t+1}^{1-\gamma} |\mathbf{Z}_t]$ requires now that we draw G random samples for both asset returns $\{\mathbf{R}_{t+1,g}\}_{g=1}^{G}$ and the predictors $\{\mathbf{Z}_{t+1,g}\}_{g=1}^{G}$ and recognizing that the choice of $\boldsymbol{\omega}_t$ also affects the term $E[a(\mathbf{Z}_{t+1}, \boldsymbol{\omega}_t) R_{p,t+1}^{1-\gamma} |\mathbf{Z}_t]$. This turns the maximization in a fixed-point problem that can be easily solved on a $(N-1) \times (N-1)$ grid for $\boldsymbol{\omega}_t$.

2.4. *Performance Measurement*

Our (pseudo) out-of-sample (OOS) experiment has a recursive structure.[6] Within this structure, we entertain both expanding and rolling window estimation schemes. In the rolling window case, we use a window of 10 years of data, which is advised by the need to estimate relatively complex and richly parameterized VAR(4) and VAR(12) models.[7] For instance, in the expanding window case, at the first iteration we estimate all

[6] The experiment has a pseudo out-of-sample structure because of a few choices concerning the general structure of the models examined—such as the number of regimes, the sensible number of lags to be employed, and especially the predictor variables in the VAR case—have been made using the full-sample of data and more generally resorting to typical findings in the literature. However, because we entertain several alternatives concerning the specific structure of the MS model (see Section 7), the number of lags, and a rich set of predictors that spans a range of earlier papers in the literature, we feel that in our experiments the genuine out-of-sample nature of the exercise has been preserved.

[7] In this case, and especially with only 10 years of monthly observations, $p = 12$ could be estimated only in the case of $M = 1$: a VAR(12) estimated on 4 series (3 real asset returns and 1 predictor) requires specifying 206 paramters to be estimated on 480 observations, which is at the boundary of feasibility and implies large standard errors. Also because of this, VAR(12) (fitted on rolling windows as well as on expanding windows) are hardly ever among the best (or even decently) performing VARs.

models (in the case of VARs, these are 392 different linear frameworks) using data for the period 1953:01-1973:01 and then proceed to compute (i) portfolio weights at horizons $H = 1$ and 60 months, in the latter case with the investor that takes into account that continuous (i.e., monthly, at the same frequency as the data) rebalancing will be allowed; (ii) similarly, portfolio weights at $H = 1$ and 60 months under alternative configurations for transaction costs. Therefore the portfolio shares will be indexed as 1973:01 and will refer to the holding period 1973:01-1973:02 in the case of $H = 1$, and to 1973:02-1978:01 for $H = 60$, where 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 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, 1953:01-2009:12. In the rolling window case, the exercise performed as of 1973:01 is identical to the one described for the expanding case, apart from the use of a shorter 10-year sample, 1963:02-1973:01. However, on the next iteration, the estimation sample is simply rolled forward by one month, to the period 1963:03-1973:02, producing again 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, 2000:01-2009:12. In the case of the rolling window OOS scheme, we are able to consistently estimate over time 381 linear models. This implies that the set of VAR models examined in this paper includes a total of 773 VAR models, over alternative values of p, alternative choices of predictors, and rolling vs. expanding window estimation choices.

To evaluate recursive OOS performance, we focus on two 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 compute/solve for CER as

$$\sum_{\tau=0}^{H-1} \beta^{\tau} E_t \left[\frac{\hat{C}_{t+\tau}^{1-\gamma}(\hat{\boldsymbol{\omega}}_t)}{1-\gamma} \right] = \sum_{\tau=0}^{H-1} \beta^{\tau} E_t \left[\frac{\tilde{C}_{t+\tau}^{1-\gamma}}{1-\gamma} \right] \quad \text{and} \quad \tilde{C}_{t+\tau} = \frac{1-\beta CER^{1-\gamma}}{1-(\beta CER^{1-\gamma})^{(H-\tau+1)/\gamma}}, \tag{11}$$

where $\tilde{C}_{t+\tau}$ 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. Given this definition, transaction costs are ignored for the purposes of computing CER even when the optimal weights $\hat{\boldsymbol{\omega}}_t$ reflect transaction costs. Second, we compute the out-of-sample Sharpe Ratio for each portfolio strategy, defined as

$$SR_t \equiv \frac{\prod_{h=1}^{T}(1+r_{p,t+h}^{H}) - \prod_{h=1}^{T}(1+r_{t+h}^{f})}{\sqrt{\frac{1}{T-H}\sum_{t=1}^{T}(R_{p,t+1}^{H} - T^{-1}\sum_{t=1}^{T} R_{p,t+1}^{H})^2}}, \tag{12}$$

where r_t^f is the real 1-month T-bill and $r_{p,t+h}^{H}$ is real portfolio return on a H-horizon strategy.

3. Data and Preliminary Evidence

Our early tests are based on monthly U.S. data on real asset returns and a standard set of predictive variables sampled over the period 1953:01-2009:12. 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 CRSP/Ibbotson 10-year bond return and the 30-day Fama-Bliss 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, short-term interest rates, inflation, the term and default spreads, industrial production, and the unemployment rate (see e.g., Henkel et al., 2011, and references therein).

Descriptive statistics of the asset returns and predictor variables are reported in Table 1. Mean real stock returns are close to 0.62% per month with mean real long-term bond returns around 0.22% implying annualized returns of 7.4% and 2.6% respectively. Estimates of volatility imply annualized values of around 15% for real stock returns and 7.7% for real bond returns, yielding unconditional monthly Sharpe ratios of 0.12 and 0.06 respectively. Real asset returns are characterized by relative large (in absolute value) 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.

Section 6.3 extends a portion of our tests to an alternative data set, monthly U.K. data on real asset returns and predictive variables sampled over the period 1957:03-2009:12. The data are obtained from Datastream and Global Financial Database. Real stock returns concern the FTSE All Share equity index, the 10-year Government bond return and the short term Treasury bill return, all deflated by the retail inflation rate. In addition to four predictors also used in exercises concerning the U.S. (the equity dividend yield, the term spread, industrial production growth and inflation), in the case of U.K. data we examine two predictors that have been strongly advocated in the literature as capable of yielding accurate linear forecasts of U.K. equity returns, the gilt-equity yield—the ratio of the yield on irredeemable gilts (consols) to the equity dividend yield—(see e.g., Clare et al., 1994, Harris and Sanchez-Valle, 2000) and rate of growth in oil prices (see e.g., Pesaran and Timmermann, 2000). Even though, tabulated summary statistics similar to Table 1 are available upon request, descriptive statistics of the asset returns and predictor variables fail to yield any big surprises: mean real stock returns are 0.65% per month with mean real long-term bond returns around 0.25% implying annualized real returns of 7.8% and 3.0% respectively. Estimates of volatility imply annualized values of around 19% for real stock returns and 5.6% for real bond returns, yielding unconditional monthly Sharpe ratios of 0.09 and 0.06 respectively. Also in this case, all real asset returns are characterized by large kurtosis coefficients and are clearly non-Gaussian.

3.1. *Regimes in U.S. Real Asset Returns*

Following common practice in the Markov switching literature, as an initial step we estimate and compare a range of MS models as distinguished by the number of regimes they require, k. In particular, the Bayes-Schwartz information criterion (BIC, a standard information criterion that trades off in-sample fit for parsimony, where the latter is considered as an indicator of likely predictive accuracy) takes a value of -16.88 for $k = 1$, of -17.13 for $k = 2$, and -17.18 for $k = 3$. Since a lower BIC signals a better model and

BIC is known to maximize the penalty inflicted to large, possibly over-parameterized models, this leads to choose $k = 3$.[8] Davies (1977)-corrected likelihood ratio tests that take into account nuisance parameter issues in standard LR tests applied to MSH confirm this choice (its value is 389.5 which, even under the 24 restrictions implied by a MSH model, is massive).[9] 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 Ono (2006).

Table 2, 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) 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 U.S. 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 events" regime, as it characterizes more than 37% of all long samples one could simulate from the estimated MSH. 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 almost half of the time, the U.S. 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 label 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. 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 returns. This third regime is also highly persistent, with an average duration of 21 months. Insofar as the ability to identify persistent states in the data is a priori indication of good potential forecasting and

[8] The Akaike and Hannan-Quinn information criteria gave identical results and favor $k = 3$ over $k = 2$. A three-state model implies the estimation of 33 parameters, which gives an acceptable saturation ratio (between total number of available observations and total number of parameters) in excess of 60. Moreover, estimates of simpler two-state MSH models reveal that one of the two regimes is scarcely persistent and hard to interpret.

[9] However, it must be recognized the Davies' (1977) correction may give only a rough guidance to the significance level associated to a given LRT. As a result we have also performed a small parametric bootstrap experiment under the null of a Gaussian IID model, and obtained that out of 500 simulation trials, a LRT of 389.5 exceeds the boostrapped simulation in 499 cases, i.e., that the p-value associated with this LRT value may be as low as 0.002.

portfolio properties of a Markov switching model, it is clear that the MSH in Table 2 shows considerable promise. Finally, the estimated transition matrix in Table 2 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 1 completes our description of the MSH 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 U.S. 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 U.S. 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 U.S. financial history, such as most of the 1960s, 1989-1997, and 1999-2000 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.

3.2. *Linear Predictability*

In a MSH model, what is predictable is the state that characterizes (in terms of means, volatilities, and correlations) financial markets. The predictability derives from the Markov chain nature of the latent state variable that drives the regime switching process as the Markov chain is generally forecastable. Since the state is a complicated nonlinear function of all past data before time t, such a predictability pattern is best thought of as a *nonlinear* one. There is another sense in which MSH implies nonlinear 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" (shifts) in the joint density of the data and as such jumps are best described as nonlinear phenomena. In a VAR model, predictability takes a considerably simpler linear nature, in which as specially selected variables, the predictors, move real asset returns with a fixed impulse-response coefficient, as captured by the corresponding VAR coefficient estimate. Therefore, before proceeding to our task we briefly discuss the VAR linear predictability patterns in U.S. data.

Because we entertain hundreds of different VARs by experimenting with all estimable combinations given by the selection of the number of lags and of the underlying predictors, to save space we do not report detailed estimates of any specific VAR model.[10] However, we examine plots of the own- and cross-correlogram functions for real stock, bond, and T-bill returns (up to lag 24), where the cross-correlograms

[10] Detailed estimates for the 773 different VAR models (both expanding and rolling window models) are available upon request. The total of 773 is less than 512×2=1,024 because in the case of rolling window VAR and high values of p, there were cases in which the estimation was impossible because of low saturation ratios and ill-conditioning of the design matrix. Experimenting with value of p between 1 and 12 ought to shields us against the dangers of using either under-parameterized models (too small a p) or over-parameterized ones (too large a p).

are computed with reference not only to lagged real asset returns but also to lagged values of the 7 predictors. We find that number of combinations of assets/predictors for which the cross-correlations are statistically significant at conventional levels are few, 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.

There is stronger evidence of linear predictability in real bond returns: past values of the term spread, the default spread, the short nominal rate, and 1-month real T-bill returns predict future, higher real returns on long-term government bonds. In many cases, these linear patterns are 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 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 subsequent 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. Especially for real bond and bill returns there is in fact little doubt ex-ante that a VAR approach to predictability *must* work *in-sample* and should have *some chances* to also prove useful in *out-of-sample* experiments.

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

4.1. *Recursive Portfolio Weights*

Figure 2 compares recursive optimal portfolio weights (for $T = 1$ month and 5 years) for two models, MSH and an expanding window VAR(1) in which all predictors are included. The right-most column of plots should be taken as an example of the type of qualitative dynamics commonly observed in VAR optimal weights. The left-hand plots also report optimal weights under the Gaussian IID benchmark. The recursive exercise is performed over the period 1973:01 - 2009:12 therefore including the deep financial crisis of 2008-2009. These weights are computed under the assumption of $\gamma = 5$. Initially, we exclude transaction costs ($\tau_f = \tau_v = 0$). Clearly, both models imply rich dynamics in optimal weights for both short- and long-run investors. However, the variability is stronger in the case of linear models than under MSH (notice that the two columns of plots have different scaling, to maximize visibility). The patterns in portfolio weights are also different. Even though there is no easy and one-to-one mapping between Figure 1 and Figure 2, in the plots to the left of Figure 2, one can recognize typical MSH-style regime dynamics. Interestingly, under MSH dynamics the variability in the weights tends to be moderate but also rather uniform over time. On the contrary, all the VAR schedules are characterized by enormous variation in the 1970s and early 1980s, which later attenuates to a narrower range of oscillation (that yet remains wider than the typical range for MSH).[11] Of course, none of the models produces weights that are as smooth as an expanding window Gaussian IID model. In particular, both linear and nonlinear predictability patterns induce strong time

[11] Plots analogous to those in Figure 2 for a rolling window VAR(1) reveal pervasive instability in the structure of the optimal asset allocation. These plots are available from the Authors upon request.

variation in optimal weights for a long-horizon (5-year) investor. Here, it is evident that while under MSH 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: in spite of the large scale for the plots to the right of Figure 2, it is easy to detect periods in which the differences between short- and long-run weights are in the order of 200-300%.

These plots lead to rather obvious remarks on the substantial differences between optimal weights under MSH and VAR(1): as one would expect, the dynamics are completely 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 MSH model naturally expresses.[12] For instance, while MSH implies average weights to stocks that are high by historical norms (around 150%) between 1992 and 1997 and average bond weights that are above average (between 90 and 120%) in the period 1980-1986, 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, while a MSH 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. Moreover, while MSH implies a demand for long-term bonds that is generally positive (even though it is modest for most of the sample) for long-run investors, a VAR has odd and counter-factual (i.e., inconsistent with equilibrium) implications by which the demand for bonds ought to have been strongly trending after 1987 and characterized by an embarrassingly negative average throughout the 1970s and more recently in the early 2000s. As in Guidolin and Timmermann (2007) the reason for these more stable, less extreme long-run asset allocations under MSH comes from the tendency of MSH 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.

Table 3 translates these visual impressions for the case $\gamma = 5$ into summary statistics for our overall sample period. Also in this case, we comment first on results that exclude transaction costs. 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 MSH model, the Gaussian IID benchmark, and a variety of VAR models that are selected in consideration of their pseudo-out-of sample portfolio performance at a 60-month horizon.[13] Table 3 makes one point clear: different VAR models – depending on the predictors they include, on the number of lags, and on whether they are estimated using either a rolling window or an expanding scheme

[12] These implications hold also for more complex (i.e., with longer lags) VAR models, such as a VAR(4) that includes all seven predictors. We present results for the VAR(1) because this model is typical, as commented in the Introduction.

[13] Section 4.1 provides further details on these CER rankings across models. Our goal is to allow a Reader to form her judgment on the differences in SAA across models and therefore we have selected a reasonable range of VAR models to give some impressions for the general type of results obtained throughout. Table 3 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.

– fail to imply homogeneous summary statistics for portfolio weights. Under different VARs, everything is possible. This is of course an exciting discovery for our main goal: because VARs seem to flexible enough to generate many alternative patterns for dynamic asset allocation, can any of these VARs approximate (or out-perform) the OOS performance of MSH? However, Table 3 also illustrates the existence of structural differences across MSH, the no predictability benchmark, and at least *the best performing VARs*, according to all types of summary statistics. In the case of recursive mean weights, the differences mostly concern stocks and 1-month T-bills: while the Gaussian IID benchmark a relatively moderate (plausible) demand for stocks (56%) and MSH a similar demand (61%), the best performing VARs estimated in this paper imply a very high and hardly plausible average weight for stocks, between 80 and 112 percent, depending on the VAR specification examined.[14]

This finding echoes the common complaint 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, nonlinear predictability. On the contrary, while all VAR models yield a negative, large average demand for T-bills (i.e., a VAR investor ought to leverage her portfolio to be able to invest more than 100% in bonds and especially stocks), MSH and the no predictability model deliver portfolios that are only modestly short in T-bills. In spite of these differences in average weights to stocks and nominal riskless cash, there are no systematic patterns to report as far as the demand of long-term bonds is concerned: while MSH implies an average weight of 50%, one can find VAR models that yield similar weights as well as VARs that imply a slightly higher average weight (up to 101%).

Table 3 also reports sample measures of dispersion of recursive portfolio weights. Given its structure, MSH delivers short-term stock and 1-month T-bill weights which are between 3 and 6 times less volatile than VAR weights are. Furthermore, these differences persist when it comes to the sample standard deviations of long-horizon weights and/or all weights computed for long-term bonds. These findings also apply to the 90% empirical range of optimal weights.[15] These results are crucial because they show that the (a priori sensible and) widespread belief that regime switching asset allocation frameworks may imply "excessively" volatile portfolio weights may be misleading when applied to long-run SAA under rebalancing. The intuition here is that also a long-horizon investor who bases her decision on a simple VAR framework, may impress remarkable volatility to her optimal weights, when linear predictability patterns are taken into account.

4.2. *Hedging Demands*

The lower panel of Figure 2 shows the implied recursive hedging demands for the period 1973:01-2009:12. Also in this case, we use a "full" expanding window VAR(1) (in which all predictors appear) as representative of the type of hedging demands that may be typically obtained under linear predictability. However, we also examine VAR(4)-implied hedging demands to substantiate our claims that the order of the VARs and the exact predictors included do not seem to be of first-order importance. The hedging demand for stocks is large (generally in excess of 80% over the entire 1973-2009 sample period) and stable (except for some spikes

[14]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.

[15]As one should expect, the recursive Gaussian IID weights are always the least volatile for all assets.

in the 1970s), 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 some changing drifts. The VAR hedging demand for T-bills trends up turning from negative and large (often below -200%) during the 1970s and early 1980s to negative but lower levels (around -100%) in the rest of the sample period. On the opposite, the VAR hedging demand for long-term bonds trends down from levels of 200% in the early 1980s to less than 100% in more recent periods.

MSH 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, MSH hedging demands are stable over time and tend to simply fluctuate around zero. However, this does not imply that MSH hedging demands are zero such that MSH nonlinear predictability is irrelevant: 1-month T-bills generally command a positive hedging demand with spikes up to 25%, while stocks imply negative hedging demands most of the time. These differences between MSH and VAR hedging demands are made more explicit in Table 2.[16] MSH delivers a positive, +8% hedging demand for 1-month T-bills, i.e., the presence of Markov regimes ends up making an investor more cautious in the long-run than in the short-run, which skews her demand towards bills; the MSH hedging demand for stocks is instead negative on average (-5%) and is basically zero for long-term bonds. 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 tends to be large and positive (easily in excess of 100% for a few of the best performing VARs). Additionally, while the variability of hedging demands for stocks tends to be modest for all VARs, there is more heterogeneity over time for hedging demands involving T-bills and bonds.

4.3. *Taking Transaction Costs into Account*

Figure 2 and Table 3 suggest that an important difference across MSH and VARs strategies may be represented by the much larger portfolio turnover implied – even at long horizons – by the latter. In a way, this structural difference may be used to provide, already at this point, a definitive answer to our motivating question: yes, nonlinear and linear predictability SAA choices are different and the latter can hardly be claimed to be able to approximate the former. Additionally, because in reality money managers do pay transaction costs when they dynamically apply SAA techniques to their portfolios, Table 3 is *prima facie* evidence in favor of MSH and against VARs, as variability indices between 3 and 9-10 times larger under VAR strategies are likely to be steeply penalized in practice. These are for instance the results in Guidolin and Na (2008) when transaction costs are applied ex-post, when computing realized portfolio performances. However, one may object that this logic is incorrect and may end up biasing the results against VARs: a rational portfolio optimizer ought to recognize the presence of fixed and variable transaction costs beforehand, when computing optimal SAA weights, and any OOS recursive exercise computing realized performances should be based on these transaction cost-adjusted weights, obtained as in Section 2.3. It is then possible

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

that, once corrected to account for the ex-ante, projected impact of transaction costs on expected utility, VARs may turn out to be better SAA tools than MSH.

This is exactly what we do: For MSH and the VAR(1) (all-predictors) respectively, Figures 3 and 4 compare raw vs. transaction cost-adjusted portfolio weights. To save space, we focus on $H = 60$ months only. For the same reason, we only plot hedging demands for stocks. Figure 3 shows that – with a few exceptions – the differences between raw and transaction cost-adjusted portfolio weights are modest. This means that when the goal is to maximize expected power utility with $\gamma = 5$, incorporating moderate transaction costs does not strongly affect weights. The intuition is that when regime switches are predicted between time $t+1$ and $t+60$, these are important enough in expected utility terms that, even after taking into account the welfare-reducing effects of bearing the costs of trading to change portfolio structure, the differences in policy functions are usually minimal. Moreover, these effects may go in any direction, sometimes increasing and sometimes decreasing the weight for a particular asset. Over the OOS period 1973-2009, in fact we have that the absolute value of the mean effects of taking transaction costs into account are 0.3, -0.5, and +0.2 percent for T-bills, bonds, and stocks, respectively. As one would expect, the standard deviations of portfolio 60-month weights uniformly decline, by 1.1, 0.3, and 0.6 percent, respectively. However, as Figure 3 shows, taking transaction costs into account has the effect of making hedging demands larger in absolute value and more erratic over time, with increases in the standard deviations of hedging demands of 12, 8.3, and 9.4 percent, for T-bills, bonds, and stocks, respectively. The intuition is that transaction costs have a much larger impact on optimal 1-month portfolio weights than they do on long-run weights and this makes the resulting hedging demands more variable, and in some sense "erratic". However, MSH hedging demands are hardly affected in their mean levels, with the largest change being a +1 percent over our sample period for stocks.

Figure 4 performs the same operation with reference to VAR(1) optimal recursive long-run weights. Although the general dynamics and trends of optimal weights are not affected by transaction costs, the plots show that their impact is non-negligible in the linear case, with transaction cost-adjusted weights that sometimes become persistently different from raw weights by as much 60-70 percent. Our intuition is that, as documented by a large body of empirical finance literature, most linear predictability patterns tend to be weak and to only marginally affect optimal expected utility. In this case, even when modest transaction costs are modeled, the impact on the dynamics of optimal portfolio weights may be of first-order importance (see Balduzzi and Lynch, 2000). Interestingly, average portfolio structures are also affected: the long-run demand for T-bills declines by 51%, the one for bonds declines by 15%, and demand for stocks further increases by 66%. As one would expect, transaction costs do make SAA positions less variable, with declines in OOS weight standard deviations of 14% for T-bills, 21% for bonds, and 7% for stocks. The effects on hedging demands follow patterns that are qualitatively consistent with our earlier comments, with modest effects on averages but considerable effects on volatility. Finally, according to standard intuition, the (unreported) effects of transaction costs on Gaussian IID portfolio weights are negligible, although inspecting the weights shows that these now tend to sometimes remain unchanged over intervals of up to 5-6 months.

5. **Realized Recursive Portfolio Performance**

Our finding that VAR models produce dynamic (short- and long-run) SAA weights and hedging demands that depart from the implications of a model that accounts for nonlinear patterns is suggestive that naive linear frameworks may be too simple to pick up predictability patterns that are in the data and that may be important in applications. However, these results are suggestive at best. Since 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 need to examine evidence on the recursive, OOS performance of both models before selecting one. This is what we set out to do in this section: use the recursive experiment outlined in Section 2.4 to assess whether VAR models can yield realized OOS performance that is equivalent (or superior) to MSH.

Before proceeding further to 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 773 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—the MSH model ought in principle to be the best performing among all models: the existence of even a few models that might out-perform the MSH would imply that at least *some* (even if few) VARs could deliver portfolio choices similar or better than MSH, 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, because MSH was rather loosely selected among a number of alternative regime-switching and nonlinear models that have appeared in the literature, one may also consider in a light unfavorable to VARs the finding that the MSH 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 $(1-0.95)\times 773 \simeq 38$ and $(1-0.99)\times 773 \simeq 8$ – we should find that MSH is a "top 40" or even a "top 10" model among all the ones we have tried in our portfolio experiment. Second, it must be stressed that even though in what follows we present realized portfolio performances for both 1-month and 5-year horizons, 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 on our empirical results.

Table 4 reports the key results of the paper.[17] The table refers to the baseline case of $\tau_f = 0.1\%$, $\tau_v = 0.5\%$. For the case of $\gamma = 5$, we report the best 8 performing models (plus a few additional benchmarks) when all models are ranked according to their real CER. The top panel concerns the 60-month horizon, while the lower panel the 1-month horizon. In the view of a long-horizon investor, MSH ranks first out of 773 alternative models with an annualized CER of 8.7%; the attached 95% confidence interval is relatively

[17]In Table 4, the reported 95% confidence bands have been bootstrapped by applying a block bootstrap to each series of recursive, realized performance statistics.

tight, [6.9%, 8.9%], 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 7% to perform SAA using the MSH model. This means that our set of VARs never includes any model that is capable to produce CERs which exceed the CER of MSH. In particular, a rather simple expanding VAR that includes only lagged real asset returns and the dividend yield as predictors of future real returns turns out to be best among all VARs and produces a higher CER of 6.8% with a bootstrapped 95% confidence interval of [2.7%, 10.8%]. On the one hand, it is clear that the two confidence intervals for MSH and the best VAR 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. Even the bootstrapped confidence interval for the 6th best VAR (an expanding window VAR(2) that only includes the dividend yield) is characterized by a 95% confidence interval of [3.2%, 8.4%]. On the other hand, Table 4 also shows the median performance statistics for all 392 expanding window VAR models and the 381 rolling window VAR models entertained in our paper, which implies highly disappointing CERs of 3.1 and 2.9 percent, which both fall outside the bootstrapped confidence interval for the MSH CER. In fact, for roughly a third of all the VARs in the horse-race, we find negative real CERs, an indication that a $\gamma = 5$ investor would have required *to be paid* in order to accept to perform her SAA using these VAR models. Clearly, MSH performs considerably better than the median, representative VAR SAA strategy. Therefore, it is not easy for VARs to beat MSH by generating a higher real CER to match realized average utility, although the best 1-5% among all VARs produce confidence intervals for realized, recursive OOS CERs that include the CER generated by MSH. Interestingly, the no predictability benchmark turns out to be the second best among all models in a long-run perspective, with a real CER of 7.5%.[18]

There is clear structure in the VARs that deliver good recursive portfolio performance and end up in our top CER ranking: these are very parsimonious models with few lags ($p = 2$ at most, but the majority of the top 40 performers we examine are of the $p = 1$ type) and in which four predictors appear in a variety of combinations: the dividend yield, the default spread, the rate of growth of IP, and the unemployment rate. Moreover, it seems that 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. The table also shows statistics for the best large scale expanding VAR(4) (69^th in the CER ranking), for the best among all rolling window VARs (80^th in the ranking), and for best among all VARs with at least 400 parameters (here it is an expanding VAR(4) that includes all predictors but the nominal short-term rate, 408^th in the ranking), to emphasize how poor their performance is, often with CERs that are close to the annualized real rate of 1.2% that a buy-and-hold investment in 1-month T-Bills would have yielded over our OOS period. We stress that these features of the best performing VARs are not exogenously imposed: the empirical finding is simply that richly parameterized VARs that include many predictors simply provide poor support to portfolio decisions.

[18]However, because the Gaussian IID is characterized by a 95% confidence interval of [5.2%, 8.9%] that overlaps with the intervals for the best VARs and MSH, there is also evidence that ignoring predictability may not be harmful to long-horizon investors. However, the ranking between MSH and the Gaussian IID model and the 1.2% differential in yearly, real terms remains unequivocal.

There are also some notable differences in the way in which good realized real CERs are obtained across models, and especially from MSH vs. VARs. In particular, MSH gives a lower mean than all other top-performing VARs (e.g., an annualized real 14.5% vs. 21.0% per annum for the best VAR) but also a sensibly lower volatility (e.g., 13.4% per year vs. 74.2% for the best performing VAR). It should not be really surprising to stumble across these good performances subject to very high volatility for VARs strategies, as we have seen in Section 4.1 that these tend to be considerably leveraged. However, these lower means with higher variances do not prevent MSH from yielding a high Sharpe ratio (0.29 in monthly terms vs. 0.08 for the best VAR), which is however below the Sharpe ratio that a Gaussian IID investor would have been able to secure (0.31) by ignoring predictability. However, the fact that a power utility investor may attach a higher real CER to MSH even though it delivers a Sharpe ratio that is inferior to Gaussian IID is not surprising: especially with a long-horizon, a power utility investor is different from a mean-variance investor who simply maximizes her portfolio Sharpe ratio (see, e.g., Zakamouline and Koekebakker, 2009). 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 H. What can then account for the difference between the real CERs for MSH vs. the no predictability benchmark? 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 mean and variance. In fact, Table 4 shows that while MSH has positive excess kurtosis that is rather close to the asymmetry exhibited by the best VAR models, MSH also has a high positive skewness that the no predictability model does not generate. Because positive skewness (right asymmetries) in realized portfolio returns benefits a power utility investor, the implication is that MSH is rewarded by a relatively high CER not because of its pure mean-variance reward ratio, but also because MSH is a way for a long-run investor to enjoy the potential benefits of large and positive realized performances.[19]

The lower panel of Table 4 reports on model performance for the best 8 models when the horizon is short. Although this is admittedly less interesting for our paper, MSH still comes in first in the ranking, with a real CER of 7% per annum. Once more, the second best realized recursive performance is obtained when all predictability patterns (linear and nonlinear) are simply ignored. The Gaussian IID real annualized CER is 5.8%, although its bootstrapped 95% confidence interval ([0.4%, 10.8%]) is the only interval that fails to include zero. We can summarize this finding as follows: a short-term investor with $\gamma = 5$ that takes into account the uncertainty surrounding our estimates of real CER, 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 MSH nor obtain a comparable recursive OOS performance. The best among all VARs yields a rather disappointing 4.3% real CER. It is of some interest to also stress that the acceptable real CER performance of MSH is now generated by properties of portfolio returns which are quite different from those commented for the

[19]In general, VAR models tend to produce positive skewness, but also relatively 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. Of course, this is a product of the high leverage that most VAR strategies require.

$H = 60$ months case. Now MSH gives the best annualized mean performance among all models (12%), although its volatility is comparable to those typical of VARs (e.g., 11.5% vs. 15.7% for the best VAR). This delivers MSH Sharpe ratios that are higher than the typical VAR Sharpe ratios (0.25 vs. 0.11). Although the results in the lower panel of Table 4 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 and how ignoring predictability may actually lead to superior 1-month recursive performance.

6. **Additional Performance Results**

6.1. *Risk Aversion*

One potential concern may be that our results are driven by a special (even though, rather typical) assumption on the coefficient of risk aversion. To address this concern we expand the range of portfolio performance results to the cases of $\gamma = 2$ and 10. First, inspection of dynamic portfolio weights and hedging demands reveal qualitatively similar structure to Figure 2 that expand the range of portfolio performance results to the cases of $\gamma = 2$ and 10. Interestingly, the general path followed by optimal portfolio weights is very similar across different coefficients of risk aversion: only the general level at which the variation occurs and the range of such variation differs (shrinks) as γ increases. This phenomenon is visible also in the case of IID Gaussian benchmark. However, another phenomenon is of interest: as one would expect from first principles, hedging demands (i.e., the differences between long-horizon and myopic portfolio demands under predictable investment opportunities) increase in absolute value as γ increases.

With respect to the SAA performance, in the case of a low risk aversion long-horizon investor, MSH is ranked third on the basis of the annualized real CER (6.2% vs. 11.3% for the best VAR). However, once more the bootstrapped 95% confidence band for MSH ([5.5%, 6.8%]) overlaps with the real CER confidence band for the best VARs (e.g., [6.2%, 16.4%] for the best performing VAR), so that it is hard to actually distinguish MSH from the top 2 models.[20] The MSH has the second highest Sharpe ratio (0.24), which indicates that for a low risk aversion investor, MSH performs well both in a mean-variance space and in a power utility space in which all moments matter. Many other comments expressed with reference to Table 4 apply also in this case. For instance, the best performing VARs are relatively parsimonious models. In the case of a high risk-aversion investor with $\gamma = 10$, MSH is the best performing model for long-horizon SAA purposes, with an annualized real CER of 4.5%. However, now the no predictability benchmark becomes a serious competitor for a long-horizon investor, with a 4.4% real CER. None of these results obtain for $\gamma = 10$ when an investor uses VARs, as their real CERs are all below 2.1% and generally negative.[21]

6.2. *Transaction Costs*

One further concern is that our choices $\tau_f = 0.1\%$ $\tau_v = 0.5\%$ may be too conservative in terms of how expensive the trading implied by the computed SAA weights may be. We address this by considering

[20] On the contrary, to tell MSH apart from the median VAR is easy, as the latter generates a disappointing -2.5% annualized real CER (this is the grand median across expanding and rolling window VARs). Also for $\gamma = 2$ the Gaussian IID never performs well and it is the second-best model. A tabulation of complete results is available from the authors upon request.

[21] Results unreported and available from the authors on request.

recursive OOS realized portfolio performances for two alternative scenarios for τ_f and τ_v. First, high variable transaction costs, i.e., $\tau_f = 0.1\%$, $\tau_v = 1\%$. This may be the case of a financial institution that benefits of a rather low level of fixed costs, but can rarely trade inside the bid-ask spread. Second, high absolute transaction costs, i.e., $\tau_f = 0.5\%$, $\tau_v = 1\%$, typical of a retail investor which pays high commissions and brokerage fees for each transaction and never trades inside the bid-ask spread. To save computational time, these two alternative experiments have been performed for the set of 392 expanding VAR models only and fixing $\gamma = 5$. The results provide both bad and good news.[22] The good news are that, independently of the transaction cost configuration assumed, MSH always outperforms – both in CER and monthly Sharpe ratio terms – even the best among the VARs. For instance, in the high transaction cost configuration case, MSH yields a real annualized CER of 6.6% and a Sharpe ratio of 0.16, which is to be compared with the matching statistics of 5.3% of 0.05 for the best VAR model (a VAR(1) in which predictors are the default spread and the unemployment rate). The more intriguing news concern instead the highest real CER model, which in all cases of high transaction costs comes to coincide with the no predictability, Gaussian IID benchmark. For instance, in our experiments depicting a retail investor, refraining from exploiting predictability altogether would have delivered a realized, recursive CER of 8.3% and a Sharpe ratio 0.23. The intuition is clear: if forced to profit from predictability, an investor that pays high transaction costs should model and exploit nonlinearities, here in the form of a three-state MSH, to be preferred over a large class of linear predictability models; however, unconditionally (when not forced to adopt an active SAA managing stance), the same investor ought to base her choices on a simple Gaussian IID model.[23] This result is in line with the early evidence in Balduzzi and Lynch (1999) who have reported that the utility costs of behaving myopically and of ignoring predictability can be substantial, but when intermediate consumption is allowed – as we do in this paper – this reduces the utility costs.

6.3. *U.K. Data: Are Results Sample-Specific?*

It is reasonable to wonder whether the results reported are specific to the U.S. sample employed. Even though both our SAA application and the sample period appear to be typical of the empirical portfolio choice literature at least since Brennan et al.'s (1997) seminal paper, and despite most of the applied literature on Markov switching in asset returns employing identical or similar data (see e.g., Ang and Bekaert, 2004, and Guidolin and Timmermann, 2007) the chances of a sample selection bias that may favor MS is a reason for concern. We therefore replicate a substantial portion of the results in Sections 3 to 5 using the data set on U.K. real asset returns and predictors presented in Section 3.1. Here we provide only a short summary of the key findings for the case of $\gamma = 5$. Unreported analyses confirm that also for U.K. real asset returns, a three-state MSH with regime-dependent covariance matrix provides an excellent fit to the data, outperforming both in terms of information criteria and of (nuisance parameter-adjusted) likelihood

[22] Results unreported and available from the authors on request.

[23] Higher transaction costs actually increase the realized, net-of-trading costs realized CER. This is only apparently contradictory: an investor that considers transaction costs ex-ante when optimizing would actually benefit from avoiding bad trades that only expose her to sure trading costs in exchange of modest (and uncertain) expected utility gains. Notice that also under high transaction costs, the bootstrapped 95% confidence intervals for CERs are wide enough to cast doubts on most pairwise comparisons.

ratio tests both single-state and simpler two-state MS models. Interestingly, the three regimes have the same interpretation as the three-state model for the U.S. in Table 2, i.e., a bear regime, an equity bull regime characterized by low volatility, and a bond bull state with high volatility.[24] A table similar to Table 3 (omitted to save space) emphasizes that also in the U.K. case, MSH implies average portfolio weights and a variability over time that are different from both the Gaussian IID benchmark and VAR models.[25] In this case, MSH implies higher (lower) weights on stocks and 1-month T-bill (bonds) than VAR models do. In essence, while the majority of VAR models suggest that an investor should—both over short and long horizons—leverage her portfolio to invest in stocks and especially long-term bonds, under MSH an investor should basically enter a portfolio long in stocks and short in bonds. Moreover, MSH implies weights that are approximately 10-15 times less volatile than VAR models do, although the key difference is that while most VARs impress drifts in the dynamics of portfolio weights, MS weights are instead characterized by typical regime shifting behavior.

However, what is most important for our purposes is the realized, OOS performance of MSH vs. VAR models. For the case of $\tau_f = 0.1\%$, $\tau_v = 0.5\%$, Table 5 reports results in the same format as Table 4 did for U.S. data. The qualitative conclusions of Table 4 go through intact: MSH is the best performing model both at a 1- and at a 60-month horizon; even the best performing VARs appear to be substantially inferior to both MSH and the single-state benchmark. However, Table 5 is in no way trivial in its implications for applied portfolio management. Focussing on long-run asset allocation, the results in Table 5 are considerably stronger than those in Table 4: the annualized CER from MSH is in this case 12% and it exceeds not only the 9.9% CER from the single-state model, but is almost double the 6.5% real CER produced by the best ranking VAR, in this case a rolling-window VAR(1) that includes the dividend yield and oil price inflation as predictors. The bootstrapped 90% confidence intervals of even the best performing VARs, fail to overlap with those of the MSH model, which points to differential performances that are sufficiently strong to be statistically significant. Also in this case, the higher CER performance of MSH is due to their superior higher-moment properties, in particular their high realized skewness: in fact, the monthly Sharpe ratio of MSH (0.12) is not much higher than the monthly Sharpe ratios for VAR models (0.10-0.11 even in the strongest performing cases).[26]

[24]However, U.K. data yield lower persistence for all regimes, and in particular for regimes 1 and 3, and considerably lower durations. Detailed results are available from the Authors upon request.

[25]In the U.K. application, we investigate four values for the number of lags p (1, 2, 4 and 12) and use two alternative estimation schemes (again, expanding and 10-year rolling). With 6 predictors, this gives a total of $3 \times 2 \times 2^6 + 2 \times 7 = 398$ models.

[26]Although this may be less interesting for our purposes, the 1-month realized OOS performances in the lower panel of Table 5 have stark implications that do not appear in Table 4: although MSH yields the highest real CER, this is very low (2.2%) and its 90% bootstrapped confidence interval includes zero. Such a low level is dangerously close to the 1.8% real performance that an investor might have simply obtained by buying and holding a portfolio invested 100% in 1-month short-term bonds. VAR models produce even worse CERs, always negative. This finding of very weak economic value of any form of predictability—both linear and nonlinear—is mostly due to the high realized kurtosis of the resulting optimal portfolios. This is emphasized by the fact that the best realized Sharpe ratios are never negligible (e.g., 0.28 for MSH and 0.25 for the best VAR) and actually higher than the 60-month ratios in the upper panel of the table.

7. **How Well Can Markov Switching Models Fare?**

So far our research design has been exclusively targeted to establish a simple result: that no VAR model—even when the search is extended to encompass a variety of predictors, lag structures, and estimation strategies—may produce either in-sample portfolio weights or realized OOS performances that come close to those of a very simple, and yet powerful three-state MSH. On the one hand, in the light of some claims implicit in portions of the empirical finance literature, this is an important result to establish, even in the presence of all the limitations of our efforts. [27] On the other hand, one may object that—given the overwhelming evidence of regimes and nonlinearities in financial data (see e.g., Ang and Timmermann, 2011)—such a result is relatively unsurprising: the limitations of linear models for predicting financial returns are indeed well-known and diffusely debated in the literature (see e.g., Goyal and Welch, 2008). In this final Section, we therefore extend our tests to ask a related but slightly different question: how much better than VARs can MS models be, in terms of their realized OOS performance? Of course, this is a tough question because, differently from VAR models that are simply "indexed" by a choice of predictors and of the number of lags, MS models may come in a number of different fashions. Therefore, we resort to a relatively large number of additional, competing but sufficiently heterogeneous MS frameworks to tease out an answer to this question. For simplicity, the results in this Section concern U.S. data and the case of $\gamma = 5$, similar to the bulk of the findings in Sections 3-6.

7.1. *Constant Variances and Correlations*

First we attempt to place a lower bound on the earlier results concerning the "distance" between MS models and VARs by applying our OOS design to simpler three-state models in which only real expected returns are allowed to depend on the Markov regime. We estimate a simpler, plain Markov switching-in-mean model (MSM) in which volatilities and correlations are held constant over time. Although differences obviously exist, the qualitative patterns and dynamics of the optimal weights are similar to the ones presented in Figure 2. For instance, while Table 3 implies that under MSH the 60-month demand for stocks is on average 55% with a sample standard deviation of 53%, the matching statistics for MSM at the same horizon are 54% and 58%, which are different but not radically so. The other relevant long-run statistics for the bond and T-bill weights implied by MSM are 21% (sample standard deviation 105%) and 25% (sample standard deviation 137%) implying that a MSM investor demands more stocks and bills and less bonds than a MSH investor does, and that her demands of bonds and bills are more volatile than under MSH. The OOS performance of MS confirms the conclusion that no VAR model may produce optimal SAA that may approximate – in terms of either weight patterns or the resulting realized performance – what an investor would get when taking into account the presence of regimes. In fact, assuming $\tau_f = 0.1\%$, $\tau_v = 0.5\%$ and in the perspective of a 5-year horizon investor (see Table 6, first panel), MSM becomes now the sixth best performing model with an annualized real CER of 3.4% which outperforms the CER of the best performing VAR (identical to Table

[27] Let us stress that not many papers have systematically investigated the realized, OOS performance of MS models as compared to standard benchmarks such as VAR models, as we do here. Partial exceptions are Ang and Bekaert (2004, limited a simple mean-variance framework and with single-state benchmarks) and Guidolin and Timmermann (2007, 2008b) that however perform only casual horse-races with hand-picked VAR(1) models.

4), 2.2%. However, it must be recognized that MSM is outperformed by the Gaussian IID, no-predictability benchmark, with a real CER of 4.4%. These figures give us the desired lower bound for the performance of MS models: it is positive, a 1.2% per year in excess of the best performing VAR when the comparison is with linear predictability frameworks; it is negative, a -1% per year, when compared with the single-state model. Interestingly, these conclusions are not completely robust to the assumed size of transaction costs. In particular, Table 6 shows that the relative performance of MSM may benefit from increasing fixed costs to $\tau_f = 0.5\%$: assuming $\tau_f = 0.5\%$, $\tau_v = 1\%$ as in third panel, the ranking of CER is fourth, and its CER moves up to 5.6% per year; assuming $\tau_f = 0.1\%$, $\tau_v = 1\%$ (higher variable costs), leads to opposite effects. The intuition is that MSM models imply scarcely persistent inferred regimes (details on smoothed and filtered probabilities confirm this claim and are available upon request) and to such a high turnover that may often lead to poor trades. It is then only a relatively high fixed brokerage cost τ_f that may weaken this tendency and improve realized OOS performance of MSM.

7.2. *Markov Switching Factor VARs*

So far our research design has simply opposed (1) to (2). Yet, a well-known class of hybrid models, called MSVARs (see Henkel et al., 2011) exist that possess a mixture of the properties of MS models and of VARs, e.g.:

$$\begin{bmatrix} \mathbf{r}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = \boldsymbol{\mu}_{S_{t+1}} + \sum_{j=1}^{p} \mathbf{A}_{j,S_{t+1}} \begin{bmatrix} \mathbf{r}_{t+1-j} \\ \mathbf{y}_{t+1-j} \end{bmatrix} + \boldsymbol{\epsilon}_{t+1} \qquad \boldsymbol{\epsilon}_{t+1} \sim N(\mathbf{0}, \boldsymbol{\Omega}),$$

where the latent, unobservable Markov state is $S_{t+1} = 1, ..., k$ and $\mathbf{A}_{S_{t+1}}$ is a $(N+M) \times (N+M)$ matrix of state-dependent vector autoregressive coefficients. We know from Section 3.2 that two-state models perform poorly because our data actually demand MS models with three or more states. However, any three- or four-state MSVAR that employs any number of the original predictors ends up being richly parameterized. In practice, with our data, it is not even feasible to estimate any three-state MSVAR(p) models with $p \geq 2$ when two or more predictors were included because of insurmountable numerical difficulties. We therefore resort to a class of models—factor VARs—that has recently received considerable attention in the empirical finance literature (see Ludvigson and Ng, 2007). Consider a MSVAR model that uses not some sub-set of the original predictors, but instead Q principal components that distill the 7 U.S. predictors in a relatively smaller number of "factors",

$$\begin{bmatrix} \mathbf{r}_{t+1} \\ \mathbf{pc}_{t+1}^{(Q)} \end{bmatrix} = \boldsymbol{\mu}_{S_{t+1}} + \sum_{j=1}^{p} \mathbf{A}_{j,S_{t+1}} \begin{bmatrix} \mathbf{r}_{t+1-j} \\ \mathbf{pc}_{t+1-j}^{(Q)} \end{bmatrix} + \boldsymbol{\epsilon}_{t+1} \qquad \boldsymbol{\epsilon}_{t+1} \sim N(\mathbf{0}, \boldsymbol{\Omega}), \tag{13}$$

where $\mathbf{pc}_t^{(Q)}$ is a $Q \times 1$ vector of Q principal components extracted from the full set of M predictors $\mathbf{y}_t$, with $Q \leq M$. By resorting to $Q < M$ principal components to replace the 7 predictors, we aim at shrinking the number of parameters while minimizing the information loss, see, e.g., Heij et al. (2008).

We apply standard (based on correlation matrix decompositions) principal component (PC) methods to $\mathbf{y}_t$, producing the first three components which summarize more than 74% of the total variability of $\mathbf{y}_t$. In particular, the first PC accounts for 35%, 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 the original predictors though they have a rather straightforward structure (detailed results are available upon request from the Authors).

PC1 loads positively with approximately equal weights on four of the seven predictors, the dividend yield, the nominal 3-month 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 the 3-month bill and the inflation rates. Finally PC3 can be basically identified with the IP growth rate. We verify that MSVAR(p) models that use PC1-PC3 perform better than all MSVAR models that include any sub-set of the predictors and that we could reliably estimate. AIC and H-Q information criteria indicate that MSVAR(1) models using PC1 and PC2 are quite competitive in terms of trade-off between fit and parsimony. While a MSVAR(1) that uses PC1 as its only predictor is the model selected by BIC over any other competing model in spite of its relatively medium-scale size (76 parameters). We therefore focus our attention on this MSVAR benchmark, labeled MSVAR-PC1.

In unreported QMLE estimation of (13) when $Q = 1$ and PC1 is the selected summary of the original predictors, we find that the three regimes carry the same interpretations as the regimes in Table 2. However, the regimes are now more persistent. Focussing on unconditional means computed as within-regime moments ($\hat{E}[\mathbf{r}_t|S = s] = (\mathbf{I} - \hat{\mathbf{A}}_s)^{-1}\hat{\boldsymbol{\mu}}_s,\ s = 1, 2, 3$), Regime 1 is a bear state in which real T-bill and stock returns are negative (-0.07% and -1.44% per month), while real bond returns are essentially zero. In this regime, PC1 predicts all asset returns with coefficients that are statistically significant. Additionally, past real stock returns predict their own future and also real bond and T-bill returns. Linear predictability is pervasive and the associated VAR coefficients are estimated with precision. This regime has an average duration of almost 22 months and it characterizes 23% of the sample. Regime 2 is a normal state with average duration of around 40 months that occurs with a high 57% frequency in any long sample because of its extreme persistence. In this regime, unconditional real mean returns are positive for all the assets, although they are modest in the case of bonds (0.10% per month, against 0.12% and 1.06% for T-bills and stocks, respectively). 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 1.42% in the case of T-bills and stocks), although the bull characterization is strong in the case of bonds (0.92% per month). Also this regime is persistent, with an average duration of 22 months, so to characterize 20% of the sample. Interestingly, in this state hardly any linear predictability remains, with the minor exception of real 1-month T-bill returns being forecastable using past real returns on other assets. Finally, unreported smoothed probability plots simply turn out to be a less jagged, smoother version of Figure 1 that convey the same basic regime classification. In fact the correlations between smoothed probabilities series of the MSH model in Sections 3-5 and the MSVAR-PC1 model in this section are all positive and statistically significant (they range from 0.42 to 0.51).

In terms of realized OOS, for intermediate and high transaction costs, MSVAR-PC1 turns out to be: (i) inferior to MSH analyzed in Sections 3-6 (e.g., its real CER is 2.8% vs. 6.6% in Table 6, second panel); (ii) superior to MSM both in terms of CER and of Sharpe ratios, which is the correct comparison to perform because both MSM and MSVAR-PC1 feature constant variances and covariances; (iii) outperforming all VAR models, including the best model that is reported in Table 6 (e.g., the real CER is 5.7% vs. 5.3% in the case of high transaction costs, last panel of Table 6). However, for low transaction costs, MSVAR-PC1 is rather disappointing (real CER of 2.3% only), although it remains superior to VARs. In a way, MSVAR-PC1 yields realized performances that have hybrid features that fall in between those of MSH and those typical of (the best among) VARs: average monthly turnover tends to be large, often larger than in VARs because the market timing optimally advised by linear predictability patterns is supplemented with timing deriving from regime switching; realized mean returns are often higher than in other MS models, but also volatility climbs up. As a result, the realized utility performance of MSVAR-PC1 is particularly penalized under low (or no) transaction costs, when the volume of trades performed is massive generating high volatility and kurtosis. However, CERs somewhat improve when adequate transaction costs are introduced and act as a filter preventing the execution of poor trades. All in all, although MSVAR-PC1 may easily be geared towards producing performances that are superior to MSM, hence this hybrid model does not represent the lower bound of realized MS performance, it is also obvious that here MSVAR-PC1 fails to maximize the potential of realized portfolio performance within the MS family.

7.3. *Markov Switching VAR-ARCH*

MSH in (2) has restricted heteroskedasticity to assume a specific form, $\boldsymbol{\epsilon}_{t+1} \sim N(\mathbf{0}, \boldsymbol{\Omega}_{S_{t+1}})$. However, a rapidly expanding class of MS models exist that allow $\boldsymbol{\Omega}_{t+1}$ to follow a regime switching ARCH model, in which variances and covariances vary over time continuously as a function of past shocks to (real) asset returns, but in which the parameters that govern such a process are also subject to infrequent shifts, as captured by a Markov chain process with constant transition probabilities (see e.g., Cai, 1994, and Dueker, 1997). In general terms, we can write the time series process of a vector of real asset returns $\mathbf{r}_t$ as stating that the conditional density of returns at time t will depend on at most of finite number of lags, $q \geq 1$, of the Markov state variable S_t.[28] In the empirical finance literature, it is typical to replace this general specification with simple VAR frameworks that allow for ARCH(q) effects in which all the matrices collecting parameters become a function of the Markov state S_t:[29]

$$\begin{aligned} \mathbf{r}_{t+1} &= \boldsymbol{\mu}_{S_{t+1}} + \sum_{j=1}^{p} \mathbf{A}_{j,S_{t+1}} \mathbf{r}_{t+1-j} + \boldsymbol{\epsilon}_{t+1} \quad \boldsymbol{\epsilon}_{t+1} \sim F(\mathbf{0}, \boldsymbol{\Omega}_{t+1}; \mathbf{v}_{S_{t+1}}),\ E_t[\boldsymbol{\epsilon}_{t+1}] = \mathbf{0},\ Var_t[\boldsymbol{\epsilon}_{t+1}] = \boldsymbol{\Omega}_{t+1} \\ \boldsymbol{\Omega}_{t+1} &= \mathbf{B}_{0,S_{t+1}} \mathbf{B}'_{0,S_{t+1}} + \sum_{j=1}^{q_1} (\mathbf{B}_{j,S_{t+1}} \mathbf{B}'_{j,S_{t+1}}) \boldsymbol{\epsilon}_{t+1-j} \boldsymbol{\epsilon}'_{t+1-j} + \sum_{j=1}^{q_2} \mathbf{N}_{t-j} \odot (\boldsymbol{\Upsilon}_{j,S_{t+1}} \boldsymbol{\Upsilon}'_{j,S_{t+1}}) \boldsymbol{\epsilon}'_{t+1-j} \boldsymbol{\epsilon}_{t+1-j} \end{aligned} \tag{14}$$

where $\boldsymbol{\Omega}_{t+1}$ is the time t conditional covariance matrix of dimension $N \times N$, $F(\cdot; \mathbf{v})$ denotes a generic density function parameterized by the vector $\mathbf{v}_{S_{t+1}}$, the (regime-dependent) matrices $\mathbf{A}_{0,S_{t+1}}$, $\mathbf{A}_{j,S_{t+1}}$, and $\boldsymbol{\Upsilon}_{j,S_{t+1}}$ are matrices of rank up to N, and the matrices $\mathbf{N}_{t-j}$ are selector matrices with generic element $\mathbf{e}'_i \mathbf{N}_{t-j} \mathbf{e}_l = I_{\{\epsilon_{i,t-j}<0,\ \epsilon_{l,t-j}<0\}}$, where $I_{\{\epsilon_{i,t-j}<0,\ \epsilon_{l,t-j}<0\}} = 1$ if $\epsilon_{i,t-j} < 0$, $\epsilon_{l,t-j} < 0$ and zero otherwise. That

[28] The assumption of that q is finite provides a rather general framework but constrains the type of conditional heteroskedastic models to be embedded within the MS framework to the classical ARCH(q) type and prevents modeling regime shifts as richer GARCH processes. Haas et al. (2004) show that direct transposition of a GARCH model to a MS setting creates technical difficulties. In particular, the variance at time t will depend on the entire history (from time 0) of regimes. Therefore the evaluation of likelihood function for a sample of length T will require the integration over all K^T possible regime paths, rendering the estimation of MS GARCH processes infeasible unless approximations are used. However, a number of recent papers (e.g., Haas et al., 2004, or Henry, 2009) have strived to apply MS GARCH models to financial data.

[29] In the expression that follows, $\odot$ denotes the element-by-element (Hadamard) product. Given two conformable matrices $\mathbf{A}$ and $\mathbf{B}$, the generic element $[i, j]$ of $\mathbf{A} \odot \mathbf{B}$ is $a_{ij} \times b_{ij}$. In (14) all the parameter matrices appear in a outer product format (e.g., $\mathbf{B}_{0,S_{t+1}} \mathbf{B}'_{0,S_{t+1}}$ instead of $\mathbf{B}_{0,S_{t+1}}$) to ensure that the resulting covariance matrix $\boldsymbol{\Omega}_{t+1}$ is positive definite within each regime.

is, at each lag $j = 1, ..., q_2$, the matrices $\mathbf{N}_{t-j}$ select elements of $\boldsymbol{\Upsilon}_{j,S_{t+1}}\boldsymbol{\Upsilon}'_{j,S_{t+1}}\boldsymbol{\varepsilon}_{t-j}\boldsymbol{\varepsilon}'_{t-j}$ that are associated to pairs of negative return shocks only. This effect ought to capture the existence of "leverage" (asymmetries) in asset returns, i.e., the fact that negative return shocks (or interactions of negative shocks) ought to increase variance and covariances more than positive return news (see Engle and Ng, 1993). Essentially (14) is a regime switching version of a multivariate VECH GARCH model, except that it fails to include a GARCH component (or, equivalently, it is of the GARCH$(q, 0)$ type), and is extended to model leverage effects, of order q_2. Since p, q_1, and q_2 are finite, the condition that the conditional density $\mathbf{r}_t$ ought at most to depend on a finite number of lags of the history of $\mathbf{r}_t$ itself is satisfied. In most applications, $F(\cdot; \mathbf{v}_{S_{t+1}})$ is either a multivariate Gaussian density (in which case, $\mathbf{v}_{S_t}$ is empty, can be set to any constant and is irrelevant in the estimation) or a multivariate Student t, in which case the scalar $v_{S_{t+1}}$ corresponds to the degrees of freedom parameters. The t-Student case has been very popular in the empirical finance literature, because asset returns are generally prone to outliers (e.g., Dueker, 1997).

We perform a brief specification search concerning the number of regimes k, the number of lags p, q_1 and q_2 as well as the need of a t-Student distribution (i.e., $v_S < \infty$ for some S), to ensure we propose an adequate model for applied portfolio management purposes. Notice that $p > 0$ implies specifying a partial VAR concerning asset returns only, which is different from the factor MSVARs in Section 7.2. Also in the MSVAR (T)ARCH (here T stands for threshold effects) in (14), there is overwhelming evidence of regimes. For instance, the log-likelihood function approximately increases from a level of 5,175-5,267 in the $k = 1$ case to a range of 5,893-5,910 in the case $k = 2$. This corresponds to a likelihood ratio test statistic in excess of 500 which – even with a number of restrictions ranging from 8 to 27 – commands a p-value which is essentially zero. This is confirmed by all information criteria dropping by at least 10% from $k = 1$ to $k = 2$. However, it is now obvious that the number of parameters to be estimated grows more than proportionally relative to k, reaching levels easily in excess of 60 for $k = 3$, 4. In fact, despite the log-likelihood function continuing to decline as the number of regimes is increased, the information criteria all indicate that a three-state MSVAR(2) *diagonal* (T)ARCH(3) model with leverage and t-Student shocks achieves the best possible trade-off between fit and parsimony. We report the estimates obtained for this model in Table 7.[30] Visibly, in a diagonal MS (T)ARCH model, only lagged values of squared shocks on real returns on asset n affect the conditional variance of asset n; similarly, only products of lagged values of squared shocks on real returns on asset n and l affect the conditional covariance between real returns on assets n and l. Again we find that all regimes are rather persistent. Regime 1 is a bear state (the implied real expected returns, $\hat{E}[\mathbf{r}_t|S = s] = (\mathbf{I} - \hat{\mathbf{A}}_{1S} - \hat{\mathbf{A}}_{2S})^{-1}\hat{\boldsymbol{\mu}}_s$ are -1.5%, -0.1%, and -0.04% for stocks, bonds, and T-bills, respectively) with high (unconditional, regime specific) volatilities (23%, 16%, and 7%, in annualized terms); in this regime, linear predictability is weak and mostly driven by the effects of real T-bill rates, while ARCH effects are modest and there is no evidence of asymmetric effects of positive vs. negative shocks in either conditional variances or covariances. In regime 1, the shape of the tails are close to a Gaussian benchmark. Regime 2 is a bull regime characterized by high real short-term rates (the implied real expected returns are 1.6%, 0.5%, and 0.8%) and volatile real bond returns (9.8% in annualized terms), while real stock returns are stable (15%). In this regime, linear predictability is rather strong and presents two distinctive features:

[30] To improve readability, in this table we have boldfaced all parameters that gave p-values of 0.05 or lower.

it goes from real bill and bond rates to all real asset returns, and it tends to link time t real returns to $t-2$ lagged returns, which is consistent with the specification of a VAR(2) model. Also ARCH effects are strong and imply high persistence of conditional variance; interestingly, the asymmetric effects seem to concern more conditional covariances than conditional variances. Interestingly, this regime proposes fat t-Student tails over and beyond what is implied by the ARCH components, with $\hat{v}_2 = 9.8$. Finally, regime 3 is once more a bond bull state (0.9% per month) characterized by low unconditional, within-regime volatility (13%, 7%, and 4% in annualized terms). ARCH effects, especially in conditional variances, remain strong.

Table 6 presents the realized OOS portfolio performance of MSVAR(2) (T)ARCH(3). In this case, we seem to have found the upper bound for the potential performance of MS models, at least within the set of models entertained in this paper: regardless of the structure assumed for transaction costs, MSVAR(2) (T)ARCH(3) always yields the highest real CER among all models—both linear and nonlinear. For instance, in the baseline case of $\tau_f = 0.1\%$, $\tau_v = 0.5\%$, the annualized CER is 9.2% and exceeds by more than 0.5% the utility-equivalent rate for MSH. Significantly, the 90% bootstrapped confidence interval for MSVAR(2) (T)ARCH(3) no longer overlaps with the CER confidence bands for the no predictability benchmark as do the weaker MS models, such as MSM and MSVAR. Such a high real CER derives not from exceptionally high Sharpe ratios, but from the good higher-moment performance of realized portfolio returns. In Table 6, as we lower the transaction costs applied, the CER performance of MSVAR(2) (T)ARCH(3) improves—as it happens for many other MS models examined before—but it is remarkable to see that as the worst trades (in an ex-ante perspective) are blocked by higher transaction costs, the distance between MSVAR(2) (T)ARCH(3) and MSH widens and in fact the second model in the CER ranking becomes the Gaussian IID benchmark. These are strong indications that modeling regime shifts in the dynamic process of conditional variances and covariances may carry a remarkable importance in SAA applications, a question that seems to have escaped the attention of the applied finance literature and that probably deserves additional work.

7.4. *Time-Varying Transition Probabilities*

To this point, all the three-regime MS models we experiment with—MSH, MSM, MSVAR(1)-PC1, and MSVAR(2) (T)ARCH(3)—maintain the plain assumption that the dynamics of regime shifts is simply governed by a homogeneous Markov chain, i.e., in which transition probabilities are constant over time and become themselves estimable parameters. However, it would be unwise to rule out the possibility that the extra power given by heterogenous Markov chains may help the performance of MS models in SAA applications. Diebold et al. (1994) extend standard MS models to a class of models with endogenously time-varying transition probabilities. Consider again the case of a three-state Markov chain with the same structure as (2) but in which the elements of the transition matrix, $\mathbf{P}_t$, are modeled as logistic functions of the vector of real stock and bond returns, $\mathbf{r}_t$:

$$\begin{aligned}
\Pr(S_t &= i|S_{t-1} = i, \mathbf{r}_{t-1}) = \frac{\exp(a_{ii} + \mathbf{b}_{ii}'\mathbf{r}_{t-1})}{1+\exp(a_{ii} + \mathbf{b}_{ii}' r\mathbf{R}_{t-1})} \in (0,1) \ (i = s, b, tb) \\
\Pr(S_t &= 2|S_{t-1} = i, \mathbf{r}_{t-1}) = \frac{\exp(a_{i2} + \mathbf{b}_{i2}'\mathbf{r}_{t-1})}{1+\exp(a_{ii} + \mathbf{b}_{ii}'\mathbf{r}_{t-1}) + \exp(a_{i2} + \mathbf{b}_{i2}'\mathbf{r}_{t-1})} \in (0,1) \\
\Pr(S_t &= 3|S_{t-1} = i, \mathbf{r}_{t-1}) = 1 - \frac{\exp(a_{ii} + \mathbf{b}_{ii}'\mathbf{r}_{t-1})}{1+\exp(a_{ii}+\mathbf{b}_{ii}'\mathbf{r}_{t-1})} - \frac{\exp(a_{i1} + \mathbf{b}_{i2}'\mathbf{r}_{t-1})}{1+\exp(a_{ii}+\mathbf{b}_{ii}'\mathbf{r}_{t-1}) + \exp(a_{i2}+\mathbf{b}_{i2}'\mathbf{r}_{t-1})} \in (0,1).
\end{aligned}$$

Clearly, when $\mathbf{b}_{ii} = \mathbf{b}_{i1} = \mathbf{0}$ for $i = 1, 2, 3$ the transition probabilities are constant. The choice of modeling the dynamics in transition probabilities as a function of asset returns follows similar choices in the literature that link $\mathbf{P}_t$ to endogenous variables that enter the MS model, e.g., Ang and Bekaert (2004).[31] With our U.S. data this choice leads to a maximized log-likelihood of 5731.11 to be compared to a log-likelihood of 5729.75 obtained under a constant transition probability model. Considering that our model with time-varying transition probabilities implies the need to estimate 12 additional parameters, a likelihood ratio test gives a test statistic of 2.72 which—under a χ^2_{12}—yields a p-value of 0.997, which fails to reject the null of constant transition probabilities.[32] Even though statistical tests do not reject our earlier assumption of constant transition probabilities, Table 6 reports the OOS performance of portfolios based on MSH with logistic transition probabilities that depend on lagged real returns.

A few papers have also considered the possibility that the dynamics in transition probabilities may be driven by aggregate macroeconomic factors. For instance, Bae, Kim, and Nelson (2007) entertain the possibility that lagged changes in the index of economic leading indicators may represent the information variables that affect the likelihood of stock volatility regime shifts since the variables are useful to time business cycle turning points, and the high (low) volatility regime tends to coincide with business cycle contractions (expansions). Therefore we repeat the estimation when the explanatory variable for the time-varying transition probabilities is the same first principal component of our vector of standard predictors for U.S. asset returns (see Section 7.2). This model leads to a maximized log-likelihood of 5739.11 to be compared to a log-likelihood of 5729.75 obtained under a constant transition probability model. Interestingly, the properties for the three regimes that are not qualitatively different from those that appear in Table 2. However, even in average terms (as $\mathbf{P}_t$ now changes over time), the three regimes now have rather modest duration. Using the same logic used above, a likelihood ratio test gives a test statistic of 18.72 which—under a χ^2_{12}—yields a p-value of 0.096, which fails to reject the null of constant transition probabilities at standard 1 and 5 percent levels. Also the associated information criteria are not substantially lower (BIC is in fact higher) than those obtained under constant transition probabilities. All in all, our data fail to strongly support the need to let transition probabilities change over time, although this claim is conditioned on our specific choices of exogenous variables allowed to affect $\mathbf{P}_t$.

Not surprisingly, Table 6 confirms that this mediocre (or definitely poor, when lagged returns are used in the logistic specification) performance of time-varying transition probability (TVTP) models fails to lead to realized portfolios that are much better than simpler MSM models. For instance, in the benchmark configuration for transaction costs, even TVTP-PC1 yields a real CER of 5.5% that is 3 full annualized percentage points below MSH and almost half the realized performance of MSVAR(2) (T)ARCH(3). The associated bootstrapped confidence interval is very wide and includes zero so that TVTP-PC1 appears to be

[31] Estimation proceeds via the EM algorithm as in the standard case, although—as suggested by Diebold et al. (1994) and given the choice to model transition probabilities as following a logit process—the first-order conditions related to the elements of $\mathbf{P}_t$ are approximated through a first-order Taylor series expansion. For reasons of parsimony, our logistic model that employs lagged real asset returns is limited to stock and bond returns. In fact, we are unable to obtain convergence for a model that also projects lagged real T-bill rates on transition probabilities. Detailed estimation results are available upon request.

[32] The associated AIC, H-Q, and BIC information criteria are -17.1904, -17.1058, and -16.9719, respectively. These can be compared to values of -17.2420, -17.1574, -17.0235 obtained assuming constant transition probabilities and point to the superiority of the simpler model.

inferior to the single-state no predictability benchmark. Moreover, the realized OOS performance of TVTP-R is worse and further degrades as transaction costs are increased. In this case, the poor performance derives not only from adverse higher-order moment properties of realized returns, but also from unsatisfactory Sharpe ratios.

8. **Conclusion**

This paper asks whether it is possible for a large class of VAR models that forecast real asset returns on stocks, bonds, and T-bills to imply dynamic strategic asset allocation choices and realized, net of transaction costs performances similar (or superior) to portfolio choices and realized performances typical of slightly more complicated nonlinear frameworks in which the existence of regimes is accounted for. After identifying the nonlinear framework with a simple three-state Markov switching model of the type recently employed by Ang and Bekaert (2002, 2004) and Guidolin and Timmermann (2007, 2008b), we obtain a rather clear negative answer to our main research hypothesis – simple VAR predictability models are not "sufficient" in either an economic or a statistical sense to summarize all the relevant predictability present in well-known U.S. data for the period 1953-2009. Our key result is that in a simple, recursive portfolio experiment we find that only a small portion of all the VARs estimated is able to produce SAA choices for long-horizon investors that compete with those obtainable under a three-state MSH model. This result does not depend on the assumed level of constant relative risk aversion, the level and structure of transaction costs, the details of the Markov model considered, or even the specific data set employed as we report qualitatively similar findings in a similar SAA application to U.K. data.

We then push forward and ask how much better than standard VARs can MS model do, in terms of realized OOS long-horizon portfolio real, risk-adjusted performances. Interestingly, here the result is two-fold. On the one hand, we find that it is possible to specify, estimate, and use Markov switching models that are sensibly superior to our plain vanilla MSH model from Sections 3-6 to support dynamic asset allocation decisions. In particular, we uncover evidence in favor of the use of MS ARCH in portfolio applications. On the other hand, even after exploring a relatively wide set of MS models, we conclude that it is difficult to outperform a simple MSH model. Our exercises provide examples of relatively complex MSVAR and MS models with time-varying transition probabilities that either do not substantially improve over the realized performance of MSH, or fail to do so outright.

There are a number of aspects of the recursive portfolio experiment that could be different and equally realistic when compared to the design we propose. To mention a few, our investor may be specified as caring for the utility of final wealth only (i.e., the problem may have no interim consumption, as in Avramov, 2002); her preferences could be different and not simply of the power type (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 want to impose constraints when solving the portfolio problem, such as short-sale constraints and upper bounds on their overall leverage. Moreover, although our paper proposes two independent empirical applications, it remains possible that a number of our conclusions—for instance on the fact that MSH systematically outperforms all VAR models—may be data- and sample-specific. It would be clearly informative to perform carefully designed simulations to assess whether and why rich VAR specifications may be useful in portfolio applications when the data are generated by nonlinear process, for instance affected by regime shifts. However, while simple designs would probably give trivial answers (apart from small sample problems, that are however not typical of empirical finance work)—for instance, that VARs can hardly outperform a k-state MSH when data are simulated from a k-state MSH—complex and realistic experiments are likely to run against problems related to the fine details of the underlying, assumed data generating process. Even though these are admittedly interesting, we keep our efforts in this paper purely empirical. Finally, Normandin and St-Amour (2008) recently use a dynamic portfolio choice framework in which predictability is captured by a large set of VAR models to ask a positive question, i.e., whether the optimal portfolio shares generated by these models are consistent with the observed investment decisions of U.S. households. Although our goals in this paper are purely normative, it would be interesting to examine whether nonlinear models may help reconcile the data with observed patterns in actual portfolio decisions.

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., A., Timmermann, 2011, "Regime Changes and Financial Markets", NBER Working paper No. 17182.

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

[6] Bae, J., C., J. Kim, C., Nelson, 2007. Why are stock returns and volatility negatively correlated? Journal of Empirical Finance 14, 41-58.

[7] Balduzzi, P., Lynch, A., 1999. Transaction costs and predictability: some utility cost calculations. Journal of Financial Economics 52, 47-78.

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

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

[10] Cai, J., 1994. A Markov model of switching-regime ARCH. Journal of Business and Economic Statistics 12, 309-316.

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

[12] Campbell, J.Y., Lo, A., MacKinlay, C., 1997. The Econometrics of Financial Markets. Princeton University Press: Princeton.

[13] Campbell, J.Y., Viceira, L., 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. Oxford University Press: Oxford.

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

[15] Clare, A.D., Thomas, S.H., Wickens, M.R., 1994. Is the gilt-equity yield ratio useful for predicting UK stock returns? The Economic Journal 104, 303-315.

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

[17] Diebold, F., J-H., Lee, G., Weinbach 1994. Regime switching with time-varying transition probabilities. In: C., Hargreaves (ed.), Nonstationary Time Series Analysis and Cointegration, Oxford University Press, Oxford, 283-302.

[18] Dueker, M., 1997. Markov switching in GARCH processes and mean-reverting stock market volatility. Journal of Business and Economic Statistics 15, 26-34.

[19] Engle, R., F., V., Ng, 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48, 1749-1778.

[20] Frauendorfer, K., Jachoby, U., Schwendener, A., 2007. Regime switching based portfolio selection for pension funds, Journal of Banking and Finance 31, 2265-2280.

[21] Gomes, F., 2007. Exploiting short-run predictability. Journal of Banking and Finance 31, 1427-1440.

[22] Goyal A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455-1508.

[23] Granger, C., Teräsvirta, T., 1993. Modelling Nonlinear Economic Relationships. Oxford University Press: Oxford.

[24] Guidolin, M., Na, F.-Z., 2008. The economic and statistical value of forecast combinations: regime switching: an application to predictable U.S. returns. In M., Wohar and D., Rapach, eds., Forecasting in the Presence of Structural Breaks and Model Uncertainty, pp. 595-657. Emerald Publishing Ltd. & Elsevier Press.

[25] 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.

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

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

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

[29] Haas, M., S., Mittnik, M., Paolella, 2004. A new approach to markov-switching GARCH models. Journal of Financial Econometrics 2, 493-530.

[30] Hamilton, J.D., 1994. Time-Series Analysis. Princeton University Press: Princeton.

[31] Harris, R.D.F., Sanchez-Valle, R., 2000. The gilt-equity yield ratio and the predictability of UK and US equity returns. Journal of Business Finance and Accounting 27, 333-357.

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

[33] Henry, O., 2009. Regime switching in the relationship between equity returns and short-term interest rates in the UK. Journal of Banking and Finance 33, 405-414.

[34] Henkel, S., Martin, J., Nardari, F., 2011, Time-varying short-horizon predictability, Journal of Financial Economics 99, 560-580.

[35] Ingersoll, J., 1987. Theory of Financial Decision Making. Rowman & Littlefield: Totowa.

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

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

[38] Lynch, A., Balduzzi, P., 2000. Predictability and Transaction Costs: The Impact on Rebalancing Rules and Behavior. Journal of Finance 55, 2285-2309.

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

[40] Normandin, M., St-Amour, P., 2008. An empirical analysis of aggregate household portfolios. Journal of Banking and Finance 32, 1583-1597.

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

[42] Pesaran, M.H., Timmermann, A., 2000. A recursive modelling approach to predicting UK stock returns. The Economic Journal 110, 159-191.

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

[44] Zakamouline,V., S., Koekebakker, 2009. Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. Journal of Banking and Finance 33, 1242-1254.

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	0.615**	0.908**	4.368	0.119	-22.623	15.776	-0.515	4.786	121.2**
Long-term Govt. Bonds Real Returns	0.218*	0.096*	2.231	0.056	-7.476	10.453	0.443	4.906*	126.0**
1-month T-bill Real Returns	0.094**	0.105**	0.319	0	-1.120	1.938	0.183	5.243*	147.2**
CPI Inflation rate	0.307**	0.296**	0.357	__	-1.915	1.806	0.049	6.118*	277.7**
Dividend Yield (annual 12-month MA)	3.265**	3.175**	1.173	__	1.100	6.260	0.222*	2.576**	10.74**
Short-Term Nominal Rate (annualized)	4.805**	4.560**	2.886	__	0.005	18.190	1.107**	5.026*	256.6**
Riskless Term Spread (annualized)	1.538**	1.490**	1.328	__	-4.300	6.920	-0.137	4.148	39.68**
Default Spread (Baa-Aaa, annualized)	0.974**	0.840**	0.461	__	0.320	3.380	1.782**	7.186**	861.4**
Industrial production growth (annualized)	2.734*	4.639**	2.861	__	-12.026	10.815	-0.560**	4.081*	69.11**
Unemployment Rate (percentage)	5.776**	5.600**	1.565	__	2.400	11.400	0.666**	3.456	56.45**

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

Table 2

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

Panel A - SINGLE STATE MODEL			
	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns
1. Mean returns	0.0937**	0.2177*	0.6149**
2. Correlations/Volatilities			
Real 1-month T-bill Returns	0.3189**		
Real Long-Term Bond Returns	0.2888*	2.2295**	
Real Stock Returns	0.0821	0.1602*	4.3645**
Panel B - THREE-STATE MODEL			
	Real 1-month T-bill Returns	Real Long-Term Bond Returns	Real Stock Returns
1. Mean returns			
Bear State	-0.0726**	-0.0831	-0.0902
Equity Bull/Low Volatility State	0.1247**	0.2077*	1.1179**
Bond Bull State	0.4710**	1.1863**	0.6067
2. Correlations/Volatilities			
Bear State			
Real 1-month T-bill Returns	0.3248**		
Real Long-Term Bond Returns	0.2059*	2.4346**	
Real Stock Returns	0.0327	-0.0117	5.1965**
Equity Bull/Low Volatility State			
Real 1-month T-bill Returns	0.2084*		
Real Long-Term Bond Returns	0.1796	1.5641**	
Real Stock Returns	0.1449	0.2600*	3.4759**
Bond Bull State			
Real 1-month T-bill Returns	0.3323**		
Real Long-Term Bond Returns	0.3979**	3.4260**	
Real Stock Returns	-0.0568	0.4229**	4.7020**
3. Transition probabilities	Bear State	Equity Bull/Low Volatility State	Bond Bull State
Bear State	0.9464**	0.0383*	0.0153
Equity Bull/Low Volatility State	0.0289	0.9709**	0.0002
Bond Bull State	0.0456	0.0017	0.9518**
Panel C - MARKOV CHAIN PROPERTIES, THREE-STATE MODEL			
	Bear State	Equity Bull/Low Volatility State	Bond Bull State
Ergodic Probabilities	0.3764	0.5019	0.1218
Average Duration (in months)	18.7	34.3	20.7

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

Table 3

Summary Statistics for Realized, Recursive Optimal Portfolio Weights Computed under Power Utility Preferences (γ = 5)

CER Rank	Model	Lags	Predictors Included							1-month T-bills			Long-term Bonds			Stocks		
			DY	Short	Term	Def.	Infl.	IP grw.	Unempl.	T=1	Long	Hedging	T=1	Long	Hedging	T=1	Long	Hedging
										Sample Mean of Portfolio Weights								
1 (2)	MSH	0	N	N	N	N	N	N	N	-0.104	-0.024	0.080	0.499	0.470	-0.029	0.605	0.554	-0.051
2 (3)	Gaussian IID	0	No Predictability							-0.078	__	__	0.518	__	__	0.560	__	__
3 (6)	Exp. VAR	1	Y	N	N	N	N	N	N	-0.114	-1.124	-1.011	0.312	0.254	-0.058	0.802	1.870	1.068
4 (7)	Exp. VAR	1	N	N	N	Y	N	Y	N	-0.850	-0.363	0.487	0.867	0.336	-0.531	0.984	1.027	0.044
5 (8)	Exp. VAR	1	Y	N	N	Y	N	Y	N	-0.963	-1.586	-0.622	1.016	0.718	-0.298	0.947	1.867	0.920
6 (9)	Exp. VAR	1	N	N	N	Y	N	Y	Y	-0.921	-0.399	0.522	0.800	0.213	-0.586	1.121	1.186	0.065
7 (10)	Exp. VAR	1	Y	N	N	N	N	Y	N	-0.167	-1.204	-1.036	0.344	0.353	0.010	0.824	1.850	1.027
8 (11)	Exp. VAR	2	Y	N	N	N	N	N	N	-0.068	-1.355	-1.287	0.262	0.210	-0.052	0.806	2.144	1.339
69 (72)	Exp. VAR	4	N	N	N	N	N	N	Y	-0.441	0.027676	0.469	0.414	-0.018	-0.432	1.027	0.990	-0.037
80 (83)	Rolling VAR	1	N	N	N	N	N	Y	N	-0.749	-0.752	-0.003	1.067	0.998	-0.070	0.681	0.754	0.073
408 (411)	Exp. VAR	4	Y	N	Y	Y	Y	Y	Y	-0.584	-2.904	-2.320	1.056	2.468	1.412	0.528	1.436	0.908
__ (1)	MS-TARCH(2)	0	N	N	N	N	N	N	N	-0.945	-0.413	0.532	0.516	0.527	0.011	1.429	0.886	-0.543
__ (4)	MSH w/TVP (PC1)	0 (1)	N	N	N	N	N	N	N	0.081	0.129	0.048	0.117	0.077	-0.040	0.802	0.794	-0.008
__ (5)	MSH w/TVP (R)	0 (1)	N	N	N	N	N	N	N	0.156	0.171	0.015	0.144	0.108	-0.035	0.700	0.720	0.020
			DY	Short	Term	Def.	Infl.	IP grw.	Unempl.	**Sample Standard Deviation of Portfolio Weights**								
1 (2)	MSH	0	N	N	N	N	N	N	N	0.440	0.419	0.158	0.392	0.362	0.121	0.579	0.534	0.149
2 (3)	Gaussian IID	0	Predictability							0.043	__	__	0.041	__	__	0.028	__	__
3 (6)	Exp. VAR	1	Y	N	N	N	N	N	N	2.563	2.585	1.476	2.183	2.059	1.334	1.193	1.490	0.855
4 (7)	Exp. VAR	1	N	N	N	Y	N	Y	N	2.843	2.470	1.314	2.420	2.152	1.139	1.127	1.129	0.681
5 (8)	Exp. VAR	1	Y	N	N	Y	N	Y	N	2.983	2.889	0.368	2.529	2.156	0.370	1.237	1.555	0.076
6 (9)	Exp. VAR	1	N	N	N	Y	N	Y	Y	2.777	2.455	0.367	2.421	2.151	0.369	1.125	1.128	0.077
7 (10)	Exp. VAR	1	Y	N	N	N	N	Y	N	2.534	2.570	0.370	2.113	2.068	0.371	1.203	1.459	0.077
8 (11)	Exp. VAR	2	Y	N	N	N	N	N	N	3.191	3.166	0.372	2.718	2.596	0.373	1.254	1.609	0.077
69 (72)	Exp. VAR	4	N	N	N	N	N	N	Y	3.598	3.324	0.363	3.314	3.093	0.365	1.451	1.309	0.077
80 (83)	Rolling VAR	1	N	N	N	N	N	Y	N	4.362	4.151	1.262	3.912	3.637	1.020	1.203	1.186	0.342
408 (411)	Exp. VAR	4	Y	N	Y	Y	Y	Y	Y	7.254	7.857	0.363	6.403	6.645	0.364	2.902	2.589	0.078
__ (1)	MS-TARCH(2)	0	N	N	N	N	N	N	N	0.573	0.245	0.529	1.395	0.100	0.382	1.563	0.156	1.545
__ (4)	MSH w/TVP (PC1)	0 (1)	N	N	N	N	N	N	N	1.386	1.388	0.066	1.020	1.025	0.040	0.790	0.772	0.046
__ (5)	MSH w/TVP (R)	0 (1)	N	N	N	N	N	N	N	0.963	0.947	0.065	0.741	0.746	0.047	0.841	0.808	0.056
			DY	Short	Term	Def.	Infl.	IP grw.	Unempl.	**Empirical 90% Range**								
1 (2)	MSH	0	N	N	N	N	N	N	N	1.250	1.196	0.456	1.039	0.950	0.251	1.640	1.524	0.415
2 (3)	Gaussian IID	0	Predictability							0.124	__	__	0.141	__	__	0.090	__	__
3 (6)	Exp. VAR	1	Y	N	N	N	N	N	N	7.996	8.458	4.322	6.640	6.275	3.817	3.703	4.476	2.415
4 (7)	Exp. VAR	1	N	N	N	Y	N	Y	N	9.286	7.961	1.150	7.202	6.232	1.213	3.546	3.601	0.178
5 (8)	Exp. VAR	1	Y	N	N	Y	N	Y	N	8.860	8.739	1.147	6.892	6.083	1.191	3.867	4.671	0.167
6 (9)	Exp. VAR	1	N	N	N	Y	N	Y	Y	9.081	8.226	1.138	6.877	6.068	1.170	3.548	3.483	0.171
7 (10)	Exp. VAR	1	Y	N	N	N	N	Y	N	8.355	8.232	1.149	6.629	6.152	1.176	3.852	4.645	0.167
8 (11)	Exp. VAR	2	Y	N	N	N	N	N	N	9.862	9.916	1.140	8.146	7.801	1.175	3.908	4.978	0.166
69 (72)	Exp. VAR	4	N	N	N	N	N	N	Y	11.011	10.346	1.138	10.681	10.005	1.168	4.535	4.187	0.170
80 (83)	Rolling VAR	1	N	N	N	N	N	Y	N	12.894	12.958	3.645	11.843	10.525	2.913	4.188	3.893	2.913
408 (411)	Exp. VAR	4	Y	N	Y	Y	Y	Y	Y	21.993	23.512	1.129	18.730	19.502	1.165	7.662	7.662	0.171
__ (1)	MS-TARCH(2)	0	N	N	N	N	N	N	N	1.890	0.761	1.849	4.496	0.336	4.245	4.984	0.489	4.922
__ (4)	MSH w/TVP (PC1)	0 (1)	N	N	N	N	N	N	N	4.668	4.650	0.090	3.154	3.189	0.077	2.707	2.655	0.057
__ (5)	MSH w/TVP (R)	0 (1)	N	N	N	N	N	N	N	3.672	3.464	0.116	2.258	2.314	0.090	2.878	2.839	0.120

Table 4

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 = 5)

CER Rank	Model	Lags	Predictors Included: DY	Short	Term	Default	Infl.	IP grw.	Unempl.	Horizon	Annualized mean returns	Annualized Mean: 95% Conf. Int. -- LB	Annualized Mean: 95% Conf. Int. -- UB	Annualized volatility	Annualized Volatility: 95% Conf. Int. -- LB	Annualized Volatility: 95% Conf. Int. -- UB	Sharpe ratio	Sharpe Ratio: 95% Conf. Int. -- LB	Sharpe Ratio: 95% Conf. Int. -- UB	CER (% Annualized)	CER: 95% Conf. Int. -- LB	CER: 95% Conf. Int. -- UB	Skewness	Kurtosis	Average monthly turnover (adjusted)
1	MSH	0	N	N	N	N	N	N	N	60	14.563	9.518	12.513	13.446	11.844	15.137	0.288	0.257	0.351	**8.667**	6.850	8.894	1.769	5.922	0.311
2	Gaussian IID	0	No Predictability							60	10.523	9.672	11.373	**8.856**	8.343	9.370	**0.306**	0.254	0.411	7.476	5.169	8.897	0.566	2.836	**0.229**
3	Exp. VAR	1	Y	N	N	N	N	N	N	60	20.969	19.505	31.344	74.184	65.805	82.564	0.077	0.051	0.129	6.766	2.733	10.799	1.677	4.077	0.607
4	Exp. VAR	1	N	N	N	Y	N	Y	N	60	20.286	16.630	23.942	74.024	67.408	80.639	0.075	0.039	0.147	6.523	3.236	9.809	1.752	4.317	0.531
5	Exp. VAR	1	Y	N	N	Y	N	Y	N	60	21.128	18.678	29.892	74.760	66.778	82.742	0.077	0.049	0.133	6.249	4.552	7.946	1.690	4.081	0.546
6	Exp. VAR	1	N	N	N	Y	N	Y	Y	60	21.013	17.217	24.809	63.425	56.997	69.854	0.091	0.053	0.165	5.980	2.433	9.527	2.002	5.340	0.523
7	Exp. VAR	1	Y	N	N	N	N	Y	N	60	**22.106**	20.069	31.762	71.956	63.674	80.239	0.084	0.058	0.137	5.898	1.861	9.935	1.773	4.456	0.523
8	Exp. VAR	2	Y	N	N	N	N	N	N	60	18.374	12.064	24.684	76.920	69.081	84.759	0.065	0.038	0.119	5.811	3.191	8.430	1.429	3.354	0.544
69	Exp. VAR	4	N	N	N	N	N	N	Y	60	15.151	10.489	19.813	74.372	67.135	81.610	0.054	0.024	0.116	4.806	4.299	5.313	1.710	4.226	0.782
80	Rolling VAR	1	N	N	N	N	N	Y	N	60	17.912	14.735	21.088	66.033	60.598	71.467	0.073	0.034	0.152	4.719	4.301	5.137	1.772	4.413	0.825
408	Exp. VAR	4	Y	N	Y	Y	Y	Y	Y	60	20.875	15.126	26.624	80.329	74.483	86.175	0.071	0.047	0.119	3.157	2.987	3.328	1.335	3.072	2.237
	Median Expanding VAR performance									60	23.950	18.263	29.636	73.784	66.429	81.139	0.067	0.060	0.074	3.065	2.690	3.439	1.641	4.042	0.777
	Median Rolling VAR performance									60	15.885	10.553	21.217	77.544	70.265	84.823	0.067	0.060	0.074	2.921	2.686	3.155	1.441	3.432	1.157
1	MSH	0	N	N	N	N	N	N	N	1	**12.041**	7.557	16.524	11.504	10.611	12.397	**0.254**	-0.048	0.215	**6.952**	-8.915	9.773	-2.201	**27.646**	0.176
2	Gaussian IID	0	No Predictability							1	4.965	1.394	8.804	**9.327**	8.728	9.926	0.109	0.007	0.270	5.770	0.404	10.822	0.102	3.851	**0.034**
3	Exp. VAR	1	Y	N	N	N	N	N	Y	1	7.863	-5.475	21.200	15.737	13.507	17.968	0.108	0.053	0.162	4.275	-27.376	35.926	1.290	7.143	0.881
4	Exp. VAR	1	Y	Y	N	Y	N	Y	Y	1	8.352	-7.716	24.420	21.022	18.185	23.858	0.087	0.035	0.140	4.145	-371.951	380.240	**2.084**	8.955	1.317
5	Exp. VAR	1	N	N	N	Y	N	N	N	1	8.797	-4.435	22.030	14.431	11.761	17.102	0.136	0.082	0.190	4.112	-96.193	104.416	1.107	7.717	0.805
6	Exp. VAR	1	Lagged real asset returns only							1	8.383	-3.456	20.221	12.427	10.219	14.636	0.148	0.093	0.203	4.079	-25.182	33.340	0.939	11.577	0.901
7	Exp. VAR	1	Y	Y	N	N	N	N	N	1	8.130	-4.582	20.842	16.772	14.767	18.778	0.106	0.052	0.159	3.437	-54.159	61.034	1.445	7.174	1.049
8	Exp. VAR	1	N	Y	N	N	N	N	Y	1	3.336	-9.978	16.649	15.165	13.241	17.090	0.025	-0.024	0.075	2.206	-61.938	66.349	1.127	7.383	1.892
47	Rolling VAR	1	Lagged real asset returns only							1	10.683	-3.574	24.940	23.075	20.887	25.262	0.109	0.051	0.166	1.695	-24.469	27.860	1.575	9.765	0.970
55	Rolling VAR	1	N	N	N	Y	N	N	Y	1	9.256	-7.573	26.085	25.760	23.424	28.096	0.081	0.025	0.137	1.590	-167.428	170.609	1.562	7.295	1.104
12	Exp. VAR	4	Y	N	N	Y	N	N	Y	1	9.223	-9.339	27.785	27.743	24.467	31.019	0.075	0.012	0.139	1.520	-339.461	342.501	1.108	5.262	1.401
	Median Expanding VAR performance									1	6.335	-10.454	23.124	27.384	24.676	30.093	0.046	-0.008	0.100	1.332	-258.945	261.609	1.083	7.289	1.002
	Median Rolling VAR performance									1	5.166	-14.977	25.309	32.025	28.305	35.744	0.029	-0.030	0.087	1.372	-328.740	331.484	1.175	6.394	1.744
	Predictors	Total	9	3	1	8	1	7	8																

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

Table 5

Best Models Ranked According to Average Long-Horizon Recursive Certainty Equivalent Return Obtained from Optimal Strategic Asset Allocation Choices Under Power Utility Preferences (γ = 5): U.K. Data

			Predictors Included							Annualized Mean			Annualized Volatility			Sharpe Ratio			Annualized CER					
CER Rank	Model	Lags	Gilt/Equity Yield Ratio	Term Spread	Inflation	IP	Oil Price Inflation	DY	Horizon	Mean returns	95% Conf. Int. -- LB	95% Conf. Int. -- UB	Volatility	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	Skewness	Kurtosis	Average monthly turnover (adjusted)
1	MSIH	0	N	N	N	N	N	N	60	23.190	20.912	25.808	47.213	44.890	55.741	0.121	0.083	0.165	**12.025**	9.076	16.664	**1.447**	4.689	1.012
2	Gaussian IID	0			No Predictability				60	8.298	7.174	8.654	**15.437**	13.898	16.386	**0.134**	0.092	0.194	9.872	1.372	14.770	0.293	3.629	**0.421**
3	VAR Rolling	1	N	N	N	N	Y	Y	60	19.435	15.993	22.403	73.179	67.843	76.792	0.072	-0.001	0.118	6.538	6.258	9.517	0.647	1.915	2.814
4	VAR Rolling	1	N	N	N	N	Y	Y	60	20.073	18.551	23.478	49.570	45.167	54.646	0.110	0.012	0.166	6.538	6.313	8.759	1.002	3.172	3.126
5	VAR Rolling	1	N	N	N	N	Y	N	60	18.904	17.890	22.519	52.780	47.389	57.724	0.097	0.021	0.144	6.538	6.436	8.288	1.115	3.387	2.669
6	VAR Expanding	1	Y	Y	Y	Y	N	Y	60	22.912	16.186	22.017	58.845	52.094	62.921	0.107	0.006	0.157	6.537	6.179	7.128	0.882	2.415	3.036
7	VAR Rolling	1	N	N	N	Y	N	Y	60	25.110	20.398	27.805	69.892	65.291	74.701	0.099	0.022	0.152	6.536	6.272	7.837	0.633	**1.832**	3.549
8	VAR Rolling	1	N	Y	N	N	Y	N	60	21.343	19.037	24.584	53.987	49.685	59.396	0.108	0.016	0.160	6.535	6.295	7.618	0.894	2.790	2.619
101	VAR Expanding	4	N	N	N	Y	N	N	60	17.922	15.238	20.607	60.388	50.524	62.186	0.080	-0.031	0.135	5.171	5.017	5.440	0.956	2.618	5.384
					Median Expanding VAR performance				60	17.513	23.521	13.521	58.433	53.910	65.052	0.156	0.054	0.107	-1.217	-14.609	15.165	0.917	2.626	4.711
					Median Rolling VAR performance				60	17.036	22.176	15.176	62.022	57.992	68.862	0.170	0.072	0.125	-1.225	-14.704	22.583	0.725	2.223	3.871
1	MSIH	0	N	N	N	N	N	N	1	52.904	37.377	70.899	50.364	43.559	52.562	**0.284**	0.227	0.348	**2.217**	-4.650	0.827	0.297	6.058	0.778
2	Gaussian IID	0			No Predictability				1	3.661	-2.276	**4.035**	**9.099**	8.087	10.068	0.080	0.026	0.139	2.131	-0.372	0.170	-0.097	**5.447**	**0.090**
3	VAR Expanding	1	Y	Y	N	N	N	Y	1	**70.827**	-31.615	84.395	81.674	65.812	89.620	0.246	-0.043	0.319	-0.515	-29.970	27.442	4.874	26.806	4.214
4	VAR Expanding	1	N	N	Y	Y	Y	N	1	-6.670	-46.553	1.496	83.204	70.443	96.446	-0.027	-0.252	0.054	-0.889	-30.142	27.011	8.243	87.835	3.380
5	VAR Expanding	1	Y	Y	N	N	N	Y	1	38.267	-26.162	43.793	83.370	66.429	95.872	0.129	-0.153	0.238	-1.050	-30.326	26.741	6.473	46.507	3.419
6	VAR Expanding	1	N	N	N	N	Y	N	1	-26.353	-54.418	-15.450	84.138	62.375	90.922	-0.094	-0.301	-0.011	-2.395	-31.146	25.798	3.212	22.398	3.730
7	VAR Expanding	2	N	N	N	Y	Y	N	1	36.935	-28.966	50.118	89.298	79.235	101.939	0.116	-0.172	0.190	-6.143	-33.645	19.546	5.919	41.155	5.031
8	VAR Rolling	4	N	N	N	N	Y	N	1	29.480	-38.133	45.174	102.107	92.475	115.585	0.080	-0.208	0.166	-8.436	-22.509	3.943	4.855	34.351	5.913
12	VAR Expanding	4	Y	N	Y	N	Y	N	1	-69.701	-71.765	-40.988	99.447	85.218	109.033	-0.206	-0.401	-0.128	-12.771	-38.063	10.968	-0.222	4.875	5.570
					Median Expanding VAR performance				1	-20.629	-54.030	-27.239	99.261	83.729	110.077	-0.014	-0.266	-0.177	-87.588	-87.941	-62.204	5.619	11.168	5.300
					Median Rolling VAR performance				1	-14.494	-53.133	-23.359	103.331	87.920	122.381	-0.009	-0.247	-0.136	-49.931	-56.461	-34.831	5.376	10.270	4.847
	Predictors	Total	4	4	3	5	9	6																

Table 6

Best Models Ranked According to Average Long-Horizon Recursive Certainty Equivalent Return Obtained from Optimal Strategic Asset Allocation Choices Under Power Utility Preferences (γ = 5): Alternative Markov Switching Models and Transaction Cost Configurations

			Predictors Included								Annualized Mean			Annualized Volatility			Sharpe Ratio			CER				
CER Rank	Model	Lags	DY	Short	Term	Default	Infl.	IP grw.	Unempl.	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	Skewness	Kurtosis	Average monthly turnover (adjusted
										Baseline transaction cost configuration: Fixed 0.1%, Variable 0.5%														
1	MSVAR(2) TARCH(2)	2	N	N	N	N	N	N	N	12.098	11.554	13.699	14.034	8.742	10.433	0.226	0.172	0.385	9.204	7.366	10.434	1.135	3.858	0.307
2	MSH	0	N	N	N	N	N	N	N	14.563	9.518	12.513	13.446	11.844	15.137	0.288	0.257	0.351	8.667	6.850	8.894	1.769	5.922	0.311
3	MSH w/TVP (PC1)	0 (1)	N	N	N	N	N	N	N	7.964	6.273	8.976	12.041	10.991	13.095	0.164	0.119	0.300	5.492	-4.159	10.706	0.979	2.828	0.521
4	MSH w/TVP (R)	0 (1)	N	N	N	N	N	N	N	6.789	5.386	7.563	11.526	10.424	12.174	0.142	0.093	0.339	4.927	-7.199	9.811	0.731	3.671	0.276
5	Gaussian IID	0				No Predictability				5.377	4.804	5.611	9.187	8.696	9.678	0.134	0.078	0.245	4.395	3.731	5.059	0.144	2.865	0.027
6	MSM	0	N	N	N	N	N	N	N	11.971	9.921	13.659	21.474	17.655	24.045	0.146	0.060	0.275	3.416	1.023	3.688	2.060	6.480	0.896
7	MSVAR(1)-1PC	1	N	N	N	N	N	N	N	9.786	5.527	14.893	26.880	18.052	35.030	0.093	-0.013	0.326	2.341	1.532	4.792	1.058	4.494	1.139
8	Exp. VAR	1	Y	N	N	N	N	N	Y	13.114	12.352	13.877	23.255	21.369	25.141	0.149	0.110	0.226	2.194	1.389	2.998	1.669	4.268	0.415
	Median Expanding VAR performance									25.575	23.487	27.664	53.344	47.806	58.882	0.132	0.093	0.211	1.335	1.014	1.656	1.982	5.259	0.606
										High variable transaction cost configuration: Fixed 0.1%, Variable 1%														
1	MSVAR(2) TARCH(2)	2	N	N	N	N	N	N	N	12.969	11.429	13.565	24.949	21.769	25.977	0.137	0.094	0.333	10.816	9.776	11.470	1.134	3.854	0.308
2	Gaussian IID	0				No Predictability				8.505	7.783	10.929	8.931	7.865	9.996	0.238	0.165	0.386	8.651	8.068	9.916	0.510	2.828	0.037
3	MSH	0	N	N	N	N	N	N	N	11.488	9.955	13.022	19.711	18.063	21.359	0.152	0.110	0.234	6.640	5.103	6.998	1.828	6.189	0.217
4	MSH w/TVP (PC1)	0 (1)	N	N	N	N	N	N	N	9.009	8.153	11.409	29.793	27.587	32.868	0.076	-0.222	0.244	6.217	-3.217	8.980	0.848	2.831	0.695
5	MSH w/TVP (R)	0 (1)	N	N	N	N	N	N	N	8.717	7.025	9.647	30.328	27.410	32.013	0.072	-0.232	0.558	5.435	-2.735	7.859	0.948	3.675	0.368
6	Exp. VAR	1	N	N	N	Y	N	N	N	20.506	16.933	24.080	62.335	56.526	68.145	0.090	0.042	0.186	5.529	2.520	8.538	1.916	4.973	0.508
7	MSVAR(1)-1PC	1	N	N	N	N	N	N	N	12.620	5.434	14.724	18.197	12.031	24.785	0.177	-0.034	0.587	2.785	1.787	5.882	1.037	4.402	0.876
8	MSM	0	N	N	N	N	N	N	N	9.444	10.376	14.215	31.479	26.926	33.928	0.076	0.025	0.234	2.617	0.762	2.902	2.129	6.773	0.627
	Median Expanding VAR performance									19.319	14.825	23.813	34.282	31.073	37.491	0.153	0.116	0.228	1.725	1.605	1.846	1.473	4.973	0.912
										High transaction cost configuration: Fixed 0.5%, Variable 1%														
1	MSVAR(2) TARCH(2)	2	N	N	N	N	N	N	N	13.098	11.554	13.699	25.085	21.856	26.082	0.138	0.138	0.138	9.891	8.870	11.596	1.135	3.858	0.307
2	Gaussian IID	0				No Predictability				8.164	7.443	10.517	8.728	7.687	9.768	0.233	0.160	0.378	8.325	7.929	9.755	0.510	2.855	0.037
3	MSH	0	N	N	N	N	N	N	N	11.566	10.080	13.053	19.351	17.700	21.002	0.156	0.115	0.238	6.581	4.986	6.887	1.715	5.892	0.206
4	MSVAR(1)-1PC	1	N	N	N	N	N	N	N	16.862	5.248	15.045	19.934	15.638	26.168	0.228	-0.199	1.409	5.664	4.465	9.878	1.049	4.482	0.587
5	MSM	0	N	N	N	N	N	N	N	11.409	12.608	17.097	37.085	31.660	40.035	0.080	0.018	0.272	5.613	3.394	5.927	2.397	7.737	0.714
6	MSH w/TVP (PC1)	0 (1)	N	N	N	N	N	N	N	10.771	8.528	11.771	30.603	27.476	32.736	0.091	0.083	0.109	5.444	-2.599	10.698	0.979	2.828	0.596
7	Exp. VAR	1	N	N	N	Y	N	N	Y	24.030	20.387	27.672	66.907	61.079	72.734	0.099	0.051	0.195	5.311	2.010	8.612	1.625	3.880	0.448
13	MSH w/TVP (R)	0 (1)	N	N	N	N	N	N	N	9.211	7.363	9.976	30.186	27.300	31.885	0.077	0.034	0.124	4.886	-5.220	12.336	0.731	3.671	0.315
	Median Expanding VAR performance									18.806	14.219	23.393	35.920	32.770	39.069	0.101	0.062	0.140	1.675	1.492	1.796	1.481	5.015	0.837

Table 7

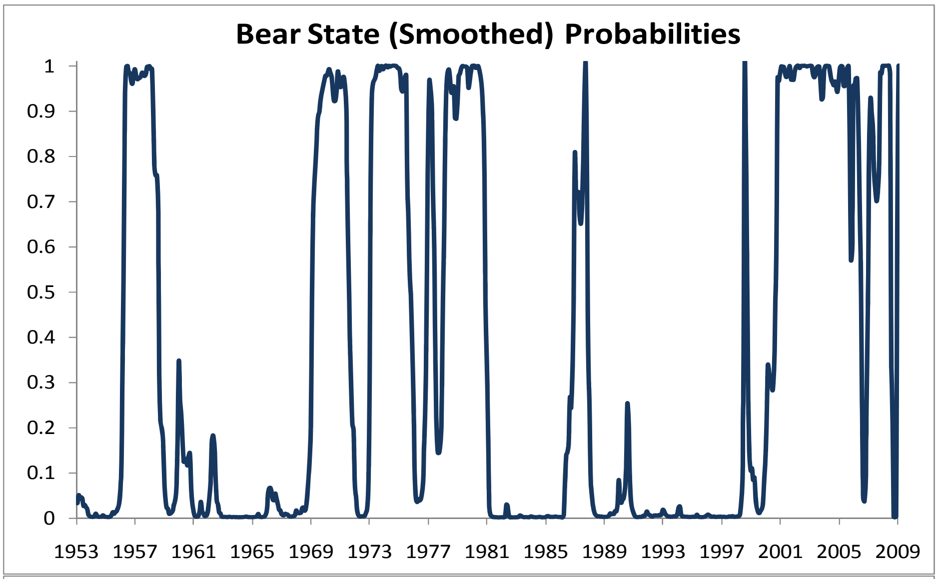
Full-Sample Estimates of Three-State Heteroskedastic MS VAR(2) Threshold ARCH(3) Multivariate Model for Real Stock, Bond, and 1-month T-Bill Returns

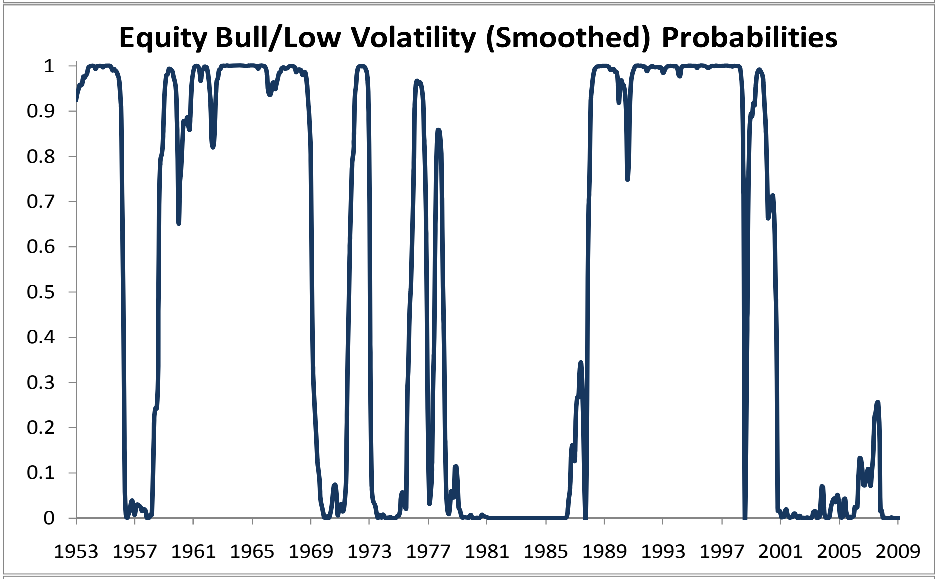
	Regime 1	
	Parameter Estimates (Std. Errors in parenthesis)	**t-Student par. (v)**
Conditional mean function -- stock returns	$r_{s,t} = \mathbf{-1.584} + 0.114r_{s,t-1} - \mathbf{0.188}r_{s,t-2} + \mathbf{0.603}r_{b,t-1} + 0.006r_{b,t-2} - \mathbf{0.453}r_{tb,t-1} + \mathbf{0.035}r_{tb,t-2} + \varepsilon_{s,t}$	
	(0.641) (0.101) (0.098) (0.209) (0.211) (0.143) (0.013)	
Conditional mean function -- bond returns	$r_{b,t} = -0.301 + 0.045r_{s,t-1} - \mathbf{0.122}r_{s,t-2} + 0.110r_{b,t-1} - \mathbf{0.451}r_{b,t-2} - \mathbf{0.594}r_{tb,t-1} - 0.114r_{tb,t-2} + \varepsilon_{b,t}$	
	(0.299) (0.044) (0.044) (0.094) (0.123) (0.183) (0.093)	
Conditional mean function -- T-bill returns	$r_{tb,t} = 0.114 + 0.070r_{s,t-1} - 0.008r_{s,t-2} + 0.121r_{b,t-1} + \mathbf{0.214}r_{b,t-2} + \mathbf{0.663}r_{tb,t-1} - \mathbf{0.158}r_{tb,t-2} + \varepsilon_{tb,t}$	
	(0.098) (0.056) (0.032) (0.134) (0.128) (0.007) (0.076)	
Conditional variance function -- stock returns	$h_{s,t} = \mathbf{23.183} + 0.033\varepsilon^2_{s,t-1} + \mathbf{0.210}\varepsilon^2_{s,t-2} + 0.094\varepsilon^2_{s,t-3} + 0.256\, I_{\{\varepsilon s,t-1<0\}}\varepsilon^2_{s,t-1}$	
	(3.675) (0.117) (0.120) (0.094) (0.221)	
Conditional variance function -- bond returns	$h_{b,t} = \mathbf{3.923} + 0.147\varepsilon^2_{b,t-1} + \mathbf{0.499}\varepsilon^2_{b,t-2} + 0.055\varepsilon^2_{b,t-3} + 0.204\, I_{\{\varepsilon b,t-1<0\}}\varepsilon^2_{b,t-1}$	**13.148**
	(0.789) (0.130) (0.164) (0.143) (0.270)	(6.907)
Conditional variance function -- T-bill returns	$h_{b,t} = \mathbf{0.349} + \mathbf{0.674}\varepsilon^2_{tb,t-1} + \mathbf{0.210}\varepsilon^2_{tb,t-2} - 0.047\varepsilon^2_{tb,t-3} + 0.164\, I_{\{\varepsilon tb,t-1<0\}}\varepsilon^2_{tb,t-1}$	
	(0.115) (0.007) (0.064) (0.102) (0.204)	
Conditional covariance: stock - bonds	$h_{s\|b,t} = \mathbf{2.874} - 0.008\varepsilon_{s,t-1}\varepsilon_{b,t-1} + \mathbf{0.313}\varepsilon_{s,t-2}\varepsilon_{b,t-2} + 0.058\varepsilon_{s,t-3}\varepsilon_{b,t-3} + 0.225\, I_{\{\varepsilon s,t-1<0\}}\varepsilon_{s,t-1}I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}$	
	(1.178) (0.134) (0.123) (0.115) (0.288)	
Conditional covariance: stocks - T-bills	$h_{s\|tb,t} = 0.453 + 0.035\varepsilon_{s,t-1}\varepsilon_{tb,t-1} + \mathbf{0.224}\varepsilon_{s,t-2}\varepsilon_{tb,t-2} + 0.088\varepsilon_{s,t-3}\varepsilon_{tb,t-3} - 0.294\, I_{\{\varepsilon s,t-1<0\}}\varepsilon_{s,t-1}I_{\{\varepsilon tb,t-1<0\}}\varepsilon_{tb,t-1}$	
	(0.339) (0.074) (0.100) (0.248) (0.184)	
Conditional covariance: bonds - T-bills	$h_{b\|tb,t} = \mathbf{1.007} + 0.033\varepsilon_{s,t-1}\varepsilon_{tb,t-1} - \mathbf{0.256}\varepsilon_{s,t-2}\varepsilon_{tb,t-2} - 0.083\varepsilon_{s,t-3}\varepsilon_{tb,t-3} + 0.289\, I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}I_{\{\varepsilon tb,t-1<0\}}\varepsilon_{tb,t-1}$	
	(0.457) (0.083) (0.124) (0.110) (0.234)	
	Regime 2	
Conditional mean function -- stock returns	$r_{s,t} = \mathbf{1.052} + 0.022r_{s,t-1} - 0.161r_{s,t-2} + \mathbf{0.555}r_{b,t-1} + 0.222r_{b,t-2} - \mathbf{0.589}r_{tb,t-1} - 0.134r_{tb,t-2} + \varepsilon_{s,t}$	
	(0.287) (0.089) (0.090) (0.231) (0.330) (0.204) (0.072)	
Conditional mean function -- bond returns	$r_{b,t} = \mathbf{0.489} - 0.030r_{s,t-1} + 0.019r_{s,t-2} + \mathbf{0.199}r_{b,t-1} - 0.181r_{b,t-2} + \mathbf{0.102}r_{tb,t-1} - 0.080rt_{b,t-2} + \varepsilon_{s,t}$	
	(0.093) (0.025) (0.027) (0.095) (0.113) (0.051) (0.143)	
Conditional mean function -- T-bill returns	$r_{tb,t} = 0.003 + \mathbf{0.132}r_{s,t-1} - 0.008r_{s,t-2} + \mathbf{0.129}r_{b,t-1} - 0.076r_{b,t-2} + \mathbf{0.837}r_{tb,t-1} - 0.114rt_{b,t-2} + \varepsilon_{tb,t}$	
	(0.022) (0.009) (0.113) (0.099) (0.230) (0.048) (0.120)	
Conditional variance function -- stock returns	$h_{s,t} = \mathbf{8.775} + 0.147\varepsilon^2_{s,t-1} + \mathbf{0.199}\varepsilon^2_{s,t-2} + 0.054\varepsilon^2_{s,t-3} + 0.311\, I_{\{\varepsilon s,t-1<0\}}\varepsilon^2_{s,t-1}$	
	(1.885) (0.131) (0.100) (0.142) (0.259)	**9.785**
Conditional variance function -- bond returns	$h_{b,t} = \mathbf{0.318} + \mathbf{0.276}\varepsilon^2_{b,t-1} + \mathbf{0.345}\varepsilon^2_{b,t-2} + \mathbf{0.197}\varepsilon^2_{b,t-3} + 0.285\, I_{\{\varepsilon b,t-1<0\}}\varepsilon^2_{b,t-1}$	(4.001)
	(0.058) (0.106) (0.117) (0.095) (0.208)	
Conditional variance function -- T-bill returns	$h_{tb,t} = \mathbf{0.079} + \mathbf{0.289}\varepsilon^2_{tb,t-1} + \mathbf{0.279}\varepsilon^2_{tb,t-2} + \mathbf{0.205}\varepsilon^2_{tb,t-3} + 0.360\, I_{\{\varepsilon tb,t-1<0\}}\varepsilon^2_{tb,t-1}$	
	(0.015) (0.109) (0.052) (0.072) (0.209)	
Conditional covariance: stock - bonds	$h_{s\|b,t} = \mathbf{0.652} - 0.124\varepsilon_{s,t-1}\varepsilon_{b,t-1} + \mathbf{0.303}\varepsilon_{s,t-2}\varepsilon_{b,t-2} - 0.113\varepsilon_{s,t-3}\varepsilon_{b,t-3} + \mathbf{0.412}\, I_{\{\varepsilon s,t-1<0\}}\varepsilon_{s,t-1}I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}$	
	(0.207) (0.087) (0.144) (0.101) (0.124)	
Conditional covariance: stocks - T-bills	$h_{s\|tb,t} = \mathbf{-0.573} - 0.007\varepsilon_{s,t-1}\varepsilon_{b,t-1} + 0.203\varepsilon_{s,t-2}\varepsilon_{b,t-2} - 0.088\varepsilon_{s,t-3}\varepsilon_{b,t-3} + \mathbf{0.536}\, I_{\{\varepsilon s,t-1<0\}}\varepsilon_{s,t-1}I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}$	
	(0.272) (0.110) (0.140) (0.076) (0.197)	
Conditional covariance: bonds - T-bills	$h_{b\|tb,t} = 0.093 + \mathbf{0.246}\varepsilon_{b,t-1}\varepsilon_{tb,t-1} + 0.083\varepsilon_{b,t-2}\varepsilon_{tb,t-2} - 0.076\varepsilon_{b,t-3}\varepsilon_{tb,t-3} + 0.218\, I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}I_{\{\varepsilon tb,t-1<0\}}\varepsilon_{tb,t-1}$	
	(0.132) (0.097) (0.113) (0.088) (0.225)	
	Regime 3	
Conditional mean function -- stock returns	$r_{s,t} = \mathbf{1.005} + 0.003r_{s,t-1} + 0.083r_{s,t-2} + 0.132r_{b,t-1} - 0.189r_{b,t-2} - \mathbf{0.445}r_{tb,t-1} + 0.080r_{tb,t-2} + \varepsilon_{s,t}$	
	(0.221) (0.078) (0.120) (0.249) (0.284) (0.163) (0.083)	
Conditional mean function -- bond returns	$r_{b,t} = \mathbf{0.583} + \mathbf{0.209}r_{s,t-1} + \mathbf{0.088}r_{s,t-2} - 0.063r_{b,t-1} + 0.103r_{b,t-2} + \mathbf{0.177}r_{tb,t-1} - 0.032rt_{b,t-2} + \varepsilon_{s,t}$	
	(0.093) (0.025) (0.027) (0.095) (0.113) (0.051) (0.143)	
Conditional mean function -- T-bill returns	$r_{tb,t} = \mathbf{0.256} + 0.003r_{s,t-1} - 0.034r_{s,t-2} - \mathbf{0.183}r_{b,t-1} - 0.134r_{b,t-2} + \mathbf{0.903}r_{tb,t-1} - \mathbf{0.284}rt_{b,t-2} + \varepsilon_{tb,t}$	
	(0.073) (0.013) (0.145) (0.073) (0.210) (0.035) (0.109)	
Conditional variance function -- stock returns	$h_{s,t} = \mathbf{1.493} + \mathbf{0.456}\varepsilon^2_{s,t-1} + 0.173\varepsilon^2_{s,t-2} + 0.088\varepsilon^2_{s,t-3} + \mathbf{0.356}\, I_{\{\varepsilon s,t-1<0\}}\varepsilon^2_{s,t-1}$	
	(0.494) (0.154) (0.184) (0.104) (0.164)	**12.367**
Conditional variance function -- bond returns	$h_{b,t} = \mathbf{0.569} + \mathbf{0.293}\varepsilon^2_{b,t-1} + \mathbf{0.303}\varepsilon^2_{b,t-2} + 0.083\varepsilon^2_{b,t-3} + 0.374\, I_{\{\varepsilon b,t-1<0\}}\varepsilon^2_{b,t-1}$	(2.560)
	(0.158) (0.103) (0.133) (0.090) (0.173)	
Conditional variance function -- T-bill returns	$h_{tb,t} = \mathbf{0.114} + \mathbf{0.370}\varepsilon^2_{tb,t-1} + \mathbf{0.158}\varepsilon^2_{tb,t-2} + \mathbf{0.184}\varepsilon^2_{tb,t-3} + \mathbf{0.388}\, I_{\{\varepsilon tb,t-1<0\}}\varepsilon^2_{tb,t-1}$	
	(0.045) (0.119) (0.073) (0.043) (0.142)	
Conditional covariance: stock - bonds	$h_{s\|b,t} = \mathbf{4.838} - 0.083\varepsilon_{s,t-1}\varepsilon_{b,t-1} + 0.083\varepsilon_{s,t-2}\varepsilon_{b,t-2} + \mathbf{0.142}\varepsilon_{s,t-3}\varepsilon_{b,t-3} + \mathbf{0.294}\, I_{\{\varepsilon s,t-1<0\}}\varepsilon_{s,t-1}I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}$	
	(0.574) (0.063) (0.092) (0.042) (0.134)	
Conditional covariance: stocks - T-bills	$h_{s\|tb,t} = \mathbf{1.003} + 0.045\varepsilon_{s,t-1}\varepsilon_{b,t-1} + 0.183\varepsilon_{s,t-2}\varepsilon_{b,t-2} - 0.003\varepsilon_{s,t-3}\varepsilon_{b,t-3} + \mathbf{0.463}\, I_{\{\varepsilon s,t-1<0\}}\varepsilon_{s,t-1}I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}$	
	(0.338) (0.149) (0.120) (0.072) (0.174)	
Conditional covariance: bonds - T-bills	$h_{b\|tb,t} = \mathbf{0.638} + \mathbf{0.393}\varepsilon_{b,t-1}\varepsilon_{tb,t-1} + 0.093\varepsilon_{b,t-2}\varepsilon_{tb,t-2} - 0.004\varepsilon_{b,t-3}\varepsilon_{tb,t-3} + 0.173\, I_{\{\varepsilon b,t-1<0\}}\varepsilon_{b,t-1}I_{\{\varepsilon tb,t-1<0\}}\varepsilon_{tb,t-1}$	
	(0.193) (0.149) (0.103) (0.146) (0.203)	

Estimated Transition Matrix	*Regime 1*	*Regime 2*	*Regime 3*
Regime 1	**0.797**	0.013	0.190
	(0.395)	(0.045)	
Regime 2	**0.066**	**0.932**	0.002
	(0.031)	(0.193)	
Regime 3	**0.055**	0.024	0.921
	(0.019)	(0.029)	

Figure 1

Smoothed Regime Probabilities from Three-State Heteroskedastic Markov Switching Model





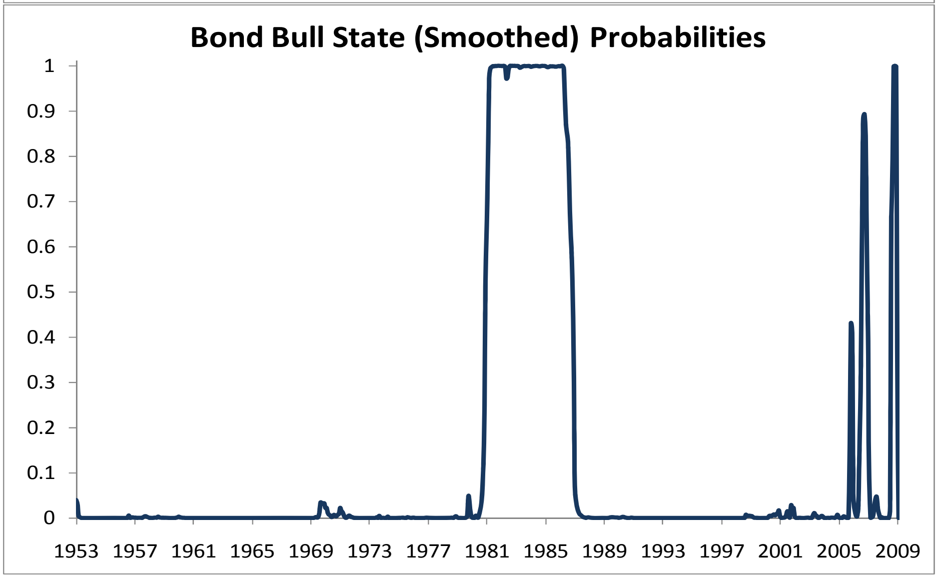


Figure 2

Dynamics of Portfolio Weights under Markov Switching vs. Full VAR(1), γ = 5: Ignoring Transaction Costs

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

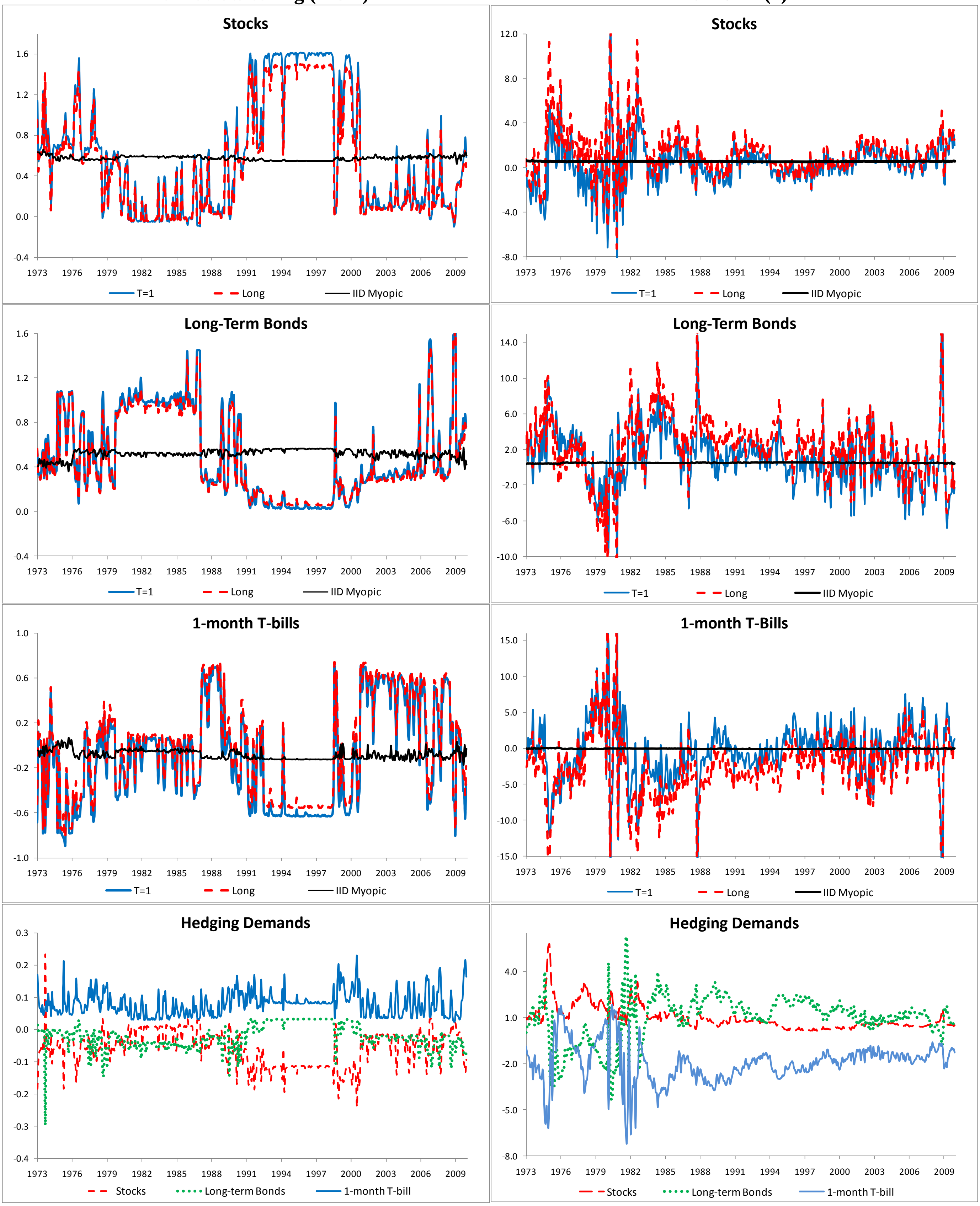


Figure 3

Effects of Modeling Transaction Costs under Markov Switching, $\gamma = 5$

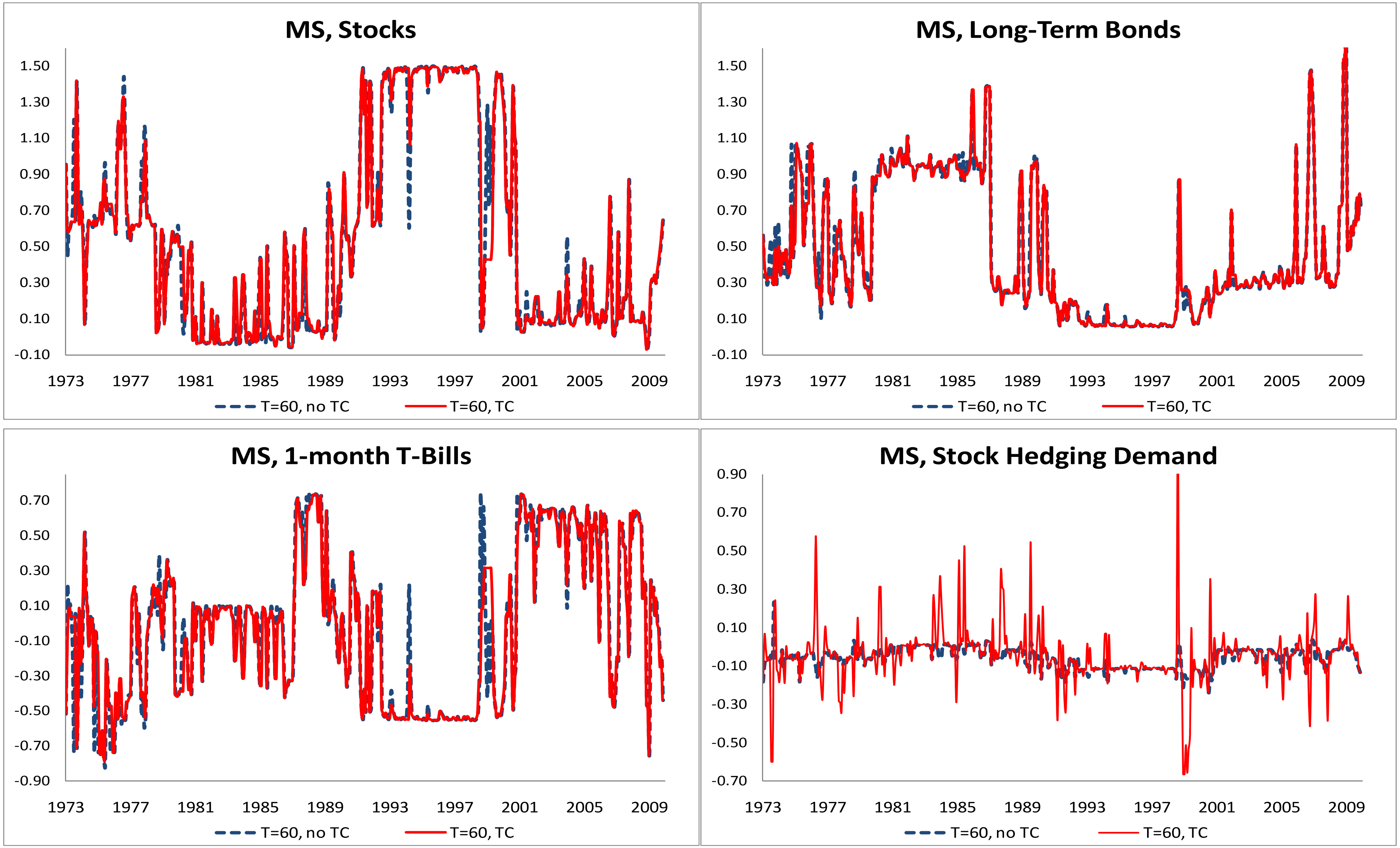


Figure 4

Effects of Modeling Transaction costs under a VAR(1), γ = 5