

Discussion Paper: 2011/12

Modeling and Estimation of Synchronization in Multistate Markov-Switching Models

Cem Çakmakli, Richard Paap and Dick van Dijk

www.feb.uva.nl/ke/UvA-Econometrics

Amsterdam School of Economics

Department of Quantitative Economics
Valckenierstraat 65-67
1018 XE AMSTERDAM
The Netherlands

UvA  UNIVERSITEIT VAN AMSTERDAM



Modeling and Estimation of Synchronization in Multistate Markov-Switching Models *

Cem Çakmaklı ^{†,1}, Richard Paap ^{‡,2}, and Dick van Dijk ^{§,2}

¹*Department of Quantitative Economics, University of Amsterdam*

²*Econometric Institute, Erasmus University Rotterdam*

December 2011

Abstract

This paper develops a Markov-Switching vector autoregressive model that allows for imperfect synchronization of cyclical regimes in multiple variables, due to phase shifts of a single common cycle. The model has three key features: (i) the amount of phase shift can be different across regimes (as well as across variables), (ii) it allows the cycle to consist of any number of regimes $J \geq 2$, and (iii) it allows for regime-dependent volatilities and correlations. In an empirical application to monthly returns on size-based stock portfolios, a three-regime model with asymmetric phase shifts and regime-dependent heteroscedasticity is found to characterize the joint distribution of returns most adequately. While large- and small-cap portfolios switch contemporaneously into boom and crash regimes, the large-cap portfolio leads the small-cap portfolio for switches to a moderate regime by a month.

Keywords: imperfect synchronization, phase shifts, regime-switching models, Bayesian analysis

JEL Classification: C11, C32, C51, C52

*We thank conference participants at the 4th CSDA International Conference on Computational and Financial Econometrics (CFE'10 London), ZEW 2nd Conference on Recent Developments in Macroeconomics (Mannheim), the Rimini Conference in Economics and Finance RCEF'10, the Workshop on Applied Finance and Financial Econometrics (Berlin) and seminar participants at the Tinbergen Institute, Helsinki Center of Economic Research for useful comments and suggestions. Any remaining errors are our own.

[†]Correspondence to: Cem Çakmaklı, Valckenierstraat 65-67, 1018 XE Amsterdam, The Netherlands, e-mail: C.Cakmakli@uva.nl.

[‡]P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: paap@ese.eur.nl.

[§]P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: djvandijk@ese.eur.nl.

1 Introduction

Synchronization of cyclical behavior of economic and financial time series variables has generated considerable interest in recent years. Typical examples include the synchronization of bull and bear phases in financial markets (see Edwards *et al.*, 2003; Bekaert *et al.*, 2005, among others) and the synchronization of business cycles in different countries or regions (see Bordo and Helbling, 2003; Artis *et al.*, 2004; Canova *et al.*, 2007; Kose *et al.*, 2003, 2008, among others). Much attention in this research area has been focused on the extreme cases of independence and perfect synchronization of cycles. In the first case, the cycles in two variables are purely idiosyncratic, while in the second case the two variables are driven by a single common cycle such that regime shifts occur contemporaneously. A wide range of econometric techniques has been developed to examine these polar cases, ranging from (non-)parametric testing procedures (see Artis *et al.*, 1997; Harding and Pagan, 2006; Pesaran and Timmermann, 2009), to unobserved components models (see Koopman and Harvey, 1997), to Markov-Switching Vector AutoRegressive (MS-VAR) models (see Krolzig, 1997; Artis *et al.*, 2004).

Perhaps not surprisingly, it is often found that neither independence nor perfect synchronization are adequate representations of the cyclical dynamics in economic and financial variables. A variety of mechanisms may lead to the intermediate case of ‘imperfect synchronization’. For example, it may be that cycles in individual variables are partly due to common components and partly due to idiosyncratic factors. This seems relevant when considering, for example, bull and bear markets for country specific asset returns, (see Kole *et al.*, 2006, for example) or business cycles in different countries or regions, (see Kose *et al.*, 2003, 2008; Del Negro and Otrok, 2008, for example). A second possibility is that the different variables in fact share a single common cyclical component, but subject to different phase shifts. This approach is appropriate in the context of comovement in asset returns in the

sense that in financial markets, shifts from bull to bear phases and vice versa may occur earlier in some assets than in others due to differences in liquidity or gradual information diffusion, among other reasons. Similarly, leading indicator variables of the economy are expected to enter to different business cycle phases earlier than coincident indicator variables. Imperfect synchronization due to different phase shifts of a single common cycle has been incorporated by Koopman and Azevedo (2008) in unobserved components models, and by Hamilton and Perez-Quiros (1996) and Paap *et al.* (2009) in MS-VAR models.

In this paper, we extend the framework of MS-VAR models for describing imperfect synchronization due to different phase shifts of a single common cycle. We generalize the models of Hamilton and Perez-Quiros (1996) and Paap *et al.* (2009) in two important directions. First, we allow for multiple regimes and consider the possibility of different phase shifts for the different regimes. Second, we allow for regime dependent-heteroscedasticity and correlations, in addition to regime switching in the means. In this way we obtain a general framework where we can simultaneously estimate the degree of synchronization together with the unobserved regimes.

Allowing for multiple regimes is relevant in the context of financial markets as well as of business cycle analysis. For example, when modeling stock returns using Markov-switching models, existing evidence suggests that three or four regimes are required to characterize the dynamics adequately. Next to regimes representing bull and bear markets, additional regimes may be necessary to capture ‘crashes’ and ‘recoveries’ (Guidolin and Timmermann, 2006), or ‘bull market corrections’ and ‘bear market rallies’ (Maheu *et al.*, 2000). Similarly, Boldin (1996) and Clements and Krolzig (2003), among others, show that a three regime model provides a more accurate and robust characterization of the US business cycle. While the specification in Hamilton and Perez-Quiros (1996) can be applied to MS-VAR models with multiple regimes, their formulation allows only for ‘symmetric’ phase shifts, that is,

the regime switches in one variable lead/lag transitions of the other variable by a fixed number of periods. This may be too restrictive given the empirical evidence that, for example, leading indicator variables have a longer lead time at business cycle peaks than at troughs, (see Board, 2001; Paap *et al.*, 2009, for example). Paap *et al.* (2009) in fact provide an MS-VAR specification that allows the phase shifts to be different depending on the type of transition. However, the formulation they use can only be applied to the MS-VAR model with two regimes.

Allowing for regime-specific volatilities may be of crucial importance for reliable and accurate identification of the regimes, especially in the context of asset returns. The main identifying characteristic of bull and bear markets is their different levels of volatilities (together with the sign of the returns), see, among others, Maheu and McCurdy (2000); Guidolin and Timmermann (2006). This gives rise to additional complications in case of imperfectly synchronized cycles, however, as there is no unique way to define a proper covariance matrix. We demonstrate that this complication may be overcome by using the decomposition of the covariance matrix in volatilities and correlations. It is fairly straightforward to guarantee that the resulting covariance matrix is positive definite, for example, while still allowing for a fair amount of flexibility in specifying the regime dynamics of the volatilities and correlations.

We adopt a Bayesian approach for estimation and inference, based upon the aim to estimate the degree of synchronization by estimating the phase shifts in the cycle of one variable relative to the cycle of another variable. As these phase shifts are represented by means of discrete valued parameters in the model, frequentist estimation procedures are not feasible.

We illustrate the potential of our framework for modeling imperfect synchronization using an empirical application for size-based US stock portfolios. There is considerable evidence that returns on (portfolios of) large stocks lead returns on

(portfolios of) small stocks, see Lo and MacKinlay (1990); McQueen *et al.* (1996); Richardson and Peterson (1999); Hou (2007), among others. Using monthly US large-cap and small-cap portfolio returns over the period 1963-2007, we find that a three-regime model with regime-dependent heteroscedasticity captures the dynamics of the portfolio returns most adequately. Allowing for different volatilities and correlations is of crucial importance for identification of the regimes. The regimes can be characterized as ‘boom’ (high average returns with moderate volatility and low correlation), ‘moderate’ (mean returns close to zero, low volatility and high correlations), and ‘bear/crash’ (negative mean return, high volatility and high correlation). We find heterogeneity in the phase shifts across different regimes, in the sense that the large-cap portfolio leads the small-cap portfolio for switches to the moderate regime by a single month, while the two portfolios switch contemporaneously into the boom and crash regimes. This suggests that during turmoil periods, with either good or bad news, information diffusion occurs rapidly and lead-lag effects are not present. News ‘travels slowly’ during more moderate times, in the sense that it is incorporated only gradually into small stocks, such that their prices adjust with a delay compared to large stocks.

Several alternatives to MS-VAR models for analyzing synchronization of cycles are available. One possibility is to use non-parametric ‘dating algorithms’ for identifying ‘turning points’ or regime switches. Once these turning points are determined, independence and perfect synchronization may be tested in a straightforward manner (see Artis *et al.*, 1997; Harding and Pagan, 2006). However, this approach typically is limited to only two regimes (corresponding with bull and bear markets in financial markets or recession and expansion periods in business cycle analysis). Another alternative is provided by multivariate unobserved components models, where trend and (common or idiosyncratic) cycle components of the series are modeled separately, as in the similar cycle model of Koopman and Harvey (1997). Our motiva-

tion for adopting MS-VAR models is that the conventional turning points approach is intrinsically embedded in these models without defining content-specific decision rules. This obviously provides a general framework for analyzing synchronization, reconciling traditional dating algorithms with modern parametric econometric tools.

The remainder of the paper is organized as follows. In Section 2 we describe the general Markov-switching VAR model with different phase shifts for multiple regimes and with regime-dependent heteroskedasticity. We discuss the Bayesian estimation methodology with prior and likelihood specifications in Section 3, with full details being provided in the Appendix. In Section 4 we consider the empirical application to size-based portfolios. Finally, we conclude in Section 5.

2 Model specification

In this section we put forward the MS-VAR model with imperfect synchronization due to phase shifts of a single common cycle. The key feature of the model is that the amount of phase shift can be different across regimes. Furthermore, it allows the cycle to consist of any number of regimes $J \geq 2$. To simplify the exposition, here we describe the model for the bivariate case, but it can be extended to include multiple variables in a straightforward manner.

Let $y_{1,t}$ and $y_{2,t}$ denote the observations of the variables of interest in period t . We assume that the J regimes are characterized by different means of $y_{1,t}$ and $y_{2,t}$. We also allow their variances and contemporaneous correlation to be regime dependent, as described in detail below, but we assume that autoregressive coefficients are constant across regimes. In case of first order autoregressive dynamics, our assumptions imply the model specification

$$\begin{aligned} y_{1,t} - \mu_{1,S_{1,t}} &= \phi_{1,1}(y_{1,t-1} - \mu_{1,S_{1,t-1}}) + \phi_{1,2}(y_{2,t-1} - \mu_{2,S_{2,t-1}}) + \varepsilon_{1,t}, \\ y_{2,t} - \mu_{2,S_{2,t}} &= \phi_{2,1}(y_{1,t-1} - \mu_{1,S_{1,t-1}}) + \phi_{2,2}(y_{2,t-1} - \mu_{2,S_{2,t-1}}) + \varepsilon_{2,t}, \end{aligned} \tag{1}$$

where $S_{l,t}$, $l = 1, 2$ are latent multinomial variables taking the value j if $y_{l,t}$ is in regime j at time t , $\mu_{l,S_{l,t}} = \mathbf{E}[y_{l,t}|S_{l,t}]$ denotes the mean of $y_{l,t}$ in the regime indicated by $S_{l,t}$, and

$$\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim NID(0, \Sigma_t). \quad (2)$$

For now, we do not specify the dynamics of the covariance matrix Σ_t , but we return to this issue in detail below.

The model is completed by specifying the dynamic properties of the unobserved regime indicators $S_{1,t}$ and $S_{2,t}$. We assume that $S_{1,t}$ and $S_{2,t}$ are first-order J -state homogenous Markov processes with transition probabilities

$$\Pr(S_{l,t} = j | S_{l,t-1} = i) = p_{ij,l}, \quad \text{for } i, j = 1, \dots, J \text{ and } l = 1, 2.$$

Without loss of generality, we assume that $y_{1,t}$ is the ‘reference series’ and we define the properties of $S_{2,t}$, the regime indicator of $y_{2,t}$, relative to $S_{1,t}$. Different specifications of the relation between the two Markov processes $S_{1,t}$ and $S_{2,t}$ imply different types of relations between the cycles of the two variables. Two extreme cases can be distinguished. First, we may assume that $S_{1,t}$ and $S_{2,t}$ are independent, that is, each variable has its own idiosyncratic cycle. Second, we can assume that $S_{1,t}$ and $S_{2,t}$ are identical, that is,

$$S_{2,t} = S_{1,t}, \quad (3)$$

or, put differently, there is a single cycle governing both variables. Following Harding and Pagan (2006), we refer to the latter case as ‘perfect synchronization’ (PS). Note that the specification with independent Markov processes $S_{1,t}$ and $S_{2,t}$ does not imply that the cycles in the two variables do not exhibit any synchronization at all. In fact, letting $p_{l,j}$ denote the unconditional probability that $S_{l,t} = j$ for $l = 1, 2$ and

$j = 1, \dots, J$, the two cycles are in the same regime with probability

$$\Pr(S_{2,t} = S_{1,t}) = \sum_{j=1}^J p_{1,j} p_{2,j} > 0.$$

In practice, the degree of synchronization between cycles in different financial and macroeconomic variables may not be perfect but still higher than the expected level under independence as given above, see the empirical evidence in Harding and Pagan (2006) and Candelon *et al.* (2008, 2009), among others. Hence, neither perfect synchronization nor independence may be adequate representations of the relation between cycles. A natural and elegant approach to accommodate this possibility is to assume that the cycle in $y_{1,t}$ leads/lags the cycle in $y_{2,t}$ by κ periods, as suggested by Hamilton and Perez-Quiros (1996). In other words, there is a common cycle but it affects the different variables with a certain phase shift. We refer to this case as imperfect synchronization with ‘symmetric’ phase shifts (SPS), where symmetry refers to the fact that all possible regime transitions in the two variables differ by the same number of time periods, κ . This specification can be formulated as

$$S_{2,t-\kappa} = S_{1,t}. \tag{4}$$

Obviously, if κ equals to zero, it reduces to the perfect synchronization case in (3).

The specification in (4) may still be too restrictive, in the sense that the phase shift of the cycle may not be the same for all possible regime transitions. For example, leading indicator variables typically have a considerably longer lead time at business cycle peaks than at troughs, see Board (2001). For this purpose, Paap *et al.* (2009) consider a two-regime model with possibly different phase shifts κ_{12} and κ_{21} for these two types of turning points.

In case of multiple regimes $J > 2$, several extensions of the specification in (4) are possible to allow for the possibility of ‘asymmetric’ phase shifts. First, following Paap

et al. (2009) we can assume a specific lead/lag structure for every possible transition from regime i to regime j . While this specification obviously provides a great deal of flexibility, it is not without problems as the number of phase shift parameters increases rapidly as the number of regimes J becomes larger. For a model with J regimes, it would result in $J(J-1)$ phase shift parameters κ_{ij} , $i, j = 1, \dots, J$; $i \neq j$. In addition, the κ_{ij} parameters are identified only by specific transitions from regime i to regime j . Hence, in order to obtain accurate estimates of these parameters we need a fair number of all possible regime switches to occur.

If the regimes are ordered in a logical way (for example according to the magnitude of the mean of the reference series $y_{1,t}$), another possible extension of (4) could be to assume identical phase shifts for all possible upward and downward regime changes. Independent of the number of regimes J , this would require estimating only two phase shifts parameters, one for transitions to regimes with a higher mean and one for transitions to regimes with a lower mean. However, as all regime changes to the same direction are assumed to be subject to identical phase shifts, the asymmetry in the lead/lag times of the cycles is highly restricted.

Here we propose an alternative solution, which aims to strike a balance between flexibility and parsimony. Specifically, instead of specifying the phase shifts for each possible type of regime switch, we assume that the regime indicator $S_{1,t}$ itself is shifted but allowing the amount of phase shift to be different across regimes. That is, we generalize (4) to the case of imperfect synchronization with ‘asymmetric’ phase shifts (APS) as

$$S_{2,t-\kappa_{S_{1,t}}} = S_{1,t}, \quad (5)$$

where the subscript $S_{1,t}$ to κ indicates that the regime indicator is shifted by a possibly different number of time periods for each regime. Hence, this specification involves a separate regime shift parameter κ_j for each regime $j = 1, \dots, J$. To put things differently, we assume that the lead/lag time is different per regime, such that

each regime in the second series starts later or earlier by κ_j periods. The advantage of this specification is that it provides a parsimonious solution (as the number of phase shift parameters grows only linearly with the number of regimes), while at the same time it is quite flexible and allows for considerable heterogeneity in the amount of phase shifts of the common cycle.

The specification in (5) is not sufficient though, as it may lead to situations where for some time periods $S_{2,t}$ is assigned multiple values or is not defined at all. To clarify the problem more explicitly, without loss of generality, consider a model with only two regimes and suppose the reference series $y_{1,t}$ is in regime 1 in period $\tau - 1$ and switches to regime 2 in period τ . Using the formulation in (5) for determining the regime of the second series implies a phase shift of κ_1 periods at time $\tau - 1$ and κ_2 periods at time τ , that is

$$S_{2,\tau-1-\kappa_1} = S_{1,\tau-1},$$

$$S_{2,\tau-\kappa_2} = S_{1,\tau}.$$

If $\kappa_1 > \kappa_2$, the shift at $\tau - 1$ is greater than the shift at τ . Hence, by applying this dating procedure we do not assign any of the regimes to the second variable for periods $[\tau - \kappa_1, \tau - \kappa_2 - 1]$. In the opposite case, we assign both regimes 1 and 2 simultaneously for periods $[\tau - \kappa_2, \tau - 1 - \kappa_1]$. This example demonstrates that we need a further decision rule to resolve these possible conflicts in determining the relevant regime for $y_{2,t}$. We suggest to use the decision rule that the regime with the larger amount of phase shift is assigned to such conflicting periods. Hence, in the example above, this rule implies assigning regime 1 to the second variable in case $\kappa_1 > \kappa_2$ for time periods that are left undetermined by (5), that is, setting $S_{2,t} = 1$ for $t \in [\tau - \kappa_1, \tau - \kappa_2 - 1]$. Similarly, we set $S_{2,t} = 2$ for $t \in [\tau - \kappa_2, \tau - 1 - \kappa_1]$ in the opposite case when $\kappa_1 < \kappa_2$.

The main advantage of the proposed decision rule to assign the regime with the

larger amount of phase shift to conflicting periods is that it implies that the phase shift is equal to κ_j when entering regime j . Possible alternatives would not achieve this. For example, assigning the regime with the smaller amount of phase shift to conflicting periods would lead to a specification where the the phase shift is equal to κ_j when leaving regime j .

In sum, the specification in (5) together with the censoring rule to assign the regime with the larger κ_j parameter in case of conflicted time periods constitutes a general framework for imperfect synchronization with multiple regimes, which encompasses perfect synchronization and imperfect synchronization with symmetric phase shifts.

Finally, we return to the specification of the structure of the covariance matrix to finalize the general model framework. As discussed before, in many applications assuming homoskedasticity for the error terms may not be appropriate assumption. In fact, regime-dependent variances and co-variances may be crucial for accurate identification of the regimes. To facilitate the specification of a regime-dependent covariance matrix we decompose it into the variances and correlations, that is

$$\Sigma_t = D_t R_t D_t, \tag{6}$$

where $D_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t})$ is a diagonal matrix with standard deviations of the error terms as diagonal elements and R_t is a matrix with ones on the diagonal and the correlation ρ_t as the off diagonal element, (see, for example, Barnard *et al.*, 2000).

In case of perfect synchronization, it is straightforward to specify a regime-dependent covariance matrix. The value of the variances and correlation just depend on the value of $S_{1,t}$. Imperfect synchronization (either with symmetric or asymmetric phase shifts) allows for the possibility that the different series are in different regimes in a given period. Consequently, there are several ways of relating the variances and correlations to the regimes. The easiest solution is to maintain the same

assumption as in the perfect synchronization case, namely that the variance and correlation regimes are completely determined by the regime indicator of the first series. This results in

$$\begin{aligned}\sigma_{l,t} &= \sigma_{l,S_{1,t}} \quad \text{for } l = 1, 2, \\ \rho_t &= \rho_{S_{1,t}},\end{aligned}\tag{7}$$

where we also allow for regime switching in the correlation parameter (see, for example, Pelletier, 2006).

Specification (7) implies perfect synchronization in the variance regimes of both series. In case phase shifts lead to imperfect synchronization in the means μ_l , $l = 1, 2$, this assumption may not be realistic. Instead, it seems more natural to relate the variances of two series to their individual regime indicator $S_{1,t}$ and $S_{2,t}$. This leads to the following specification

$$\begin{aligned}\sigma_{l,t} &= \sigma_{l,S_{l,t}} \quad \text{for } l = 1, 2, \\ \rho_t &= \rho_{S_{1,t}},\end{aligned}\tag{8}$$

where we use the regime indicator of the first series $S_{1,t}$ to describe the regime of the correlation parameter.

3 Estimation

We use a Bayesian approach for estimation and inference in the MS-VAR model with imperfect synchronization due to asymmetric phase shifts using Markov Chain Monte Carlo (MCMC) techniques. Specifically, we use Gibbs sampling together with data augmentation (see Geman and Geman, 1984; Tanner and Wong, 1987) to obtain posterior results. The main reason for adopting a Bayesian approach is that we treat the lead/lag times κ_j , $j = 1, \dots, J$, as unknown parameters to be estimated. Given that the κ_j 's are discrete, a frequentist approach (such as maximum likelihood combined with the EM algorithm) is infeasible. In Section 3.1

we derive the likelihood function of the model, while we discuss the specifications of the prior distributions in Section 3.2. In Section 3.3 we outline the Gibbs sampling algorithm for simulating from the posterior distribution. Finally, in Section 3.4 we discuss how to select a preferred specification from several competing models with possibly different numbers of regimes, different types of synchronization, and different specifications of the covariance matrix.

3.1 Likelihood function

We follow the common practice in the literature to treat the unobserved regimes in Markov-Switching models as parameters to be estimated. For this purpose we derive the complete data likelihood function. We do so for the bivariate MS-VAR model as given in (1) with (2), with imperfect synchronization due to asymmetric phase shifts as specified in (5) and covariance matrix as in (6)–(8). In vector notation we can write the model as

$$Y_t - \mathcal{M}_{S_t} = \Phi(Y_{t-1} - \mathcal{M}_{S_{t-1}}) + \mathcal{E}_t, \quad \text{with } \mathcal{E}_t \sim N(0, \Sigma_t), \quad (9)$$

where

$$Y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}, \quad S_t = \begin{pmatrix} S_{1,t} \\ S_{2,t} \end{pmatrix}, \quad \mathcal{E}_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix},$$

and

$$\mathcal{M}_{S_t} = \begin{pmatrix} \mu_{1,S_{1,t}} \\ \mu_{2,S_{2,t}} \end{pmatrix}, \quad \text{and } \Phi = \begin{pmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{pmatrix}.$$

Given the assumption of multivariate normality for the shocks \mathcal{E}_t , the conditional density of Y_t given the past observations $Y^{t-1} = \{Y_1, \dots, Y_{t-1}\}$ and given the past

and current states $S^t = \{S_1, \dots, S_t\}$ is given by

$$f(Y_t|Y^{t-1}, S^t, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \Phi, \kappa) = \left(\frac{1}{\sqrt{2\pi}}\right)^2 |\Sigma_t|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathcal{E}'_t \Sigma_t^{-1} \mathcal{E}_t\right), \quad (10)$$

where $\mu_l = (\mu_{l,1}, \dots, \mu_{l,J})'$ and $\sigma_l^2 = (\sigma_{l,1}^2, \dots, \sigma_{l,J}^2)'$ for $l = 1, 2$, $\rho = (\rho_1, \dots, \rho_J)'$ and $\kappa = (\kappa_1, \dots, \kappa_J)'$. The complete data likelihood function for model (9) conditional on the first observation equals

$$f(Y^T, S^T|Y_1, \theta) = \left(\prod_{i=1}^J \prod_{j=1}^J (p_{ij}^{T_{ij}})\right) \prod_{t=2}^T f(Y_t|Y^{t-1}, S^t, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \Phi, \kappa), \quad (11)$$

where T denotes the sample size, T_{ij} is the number of transitions from regime i to regime j , and $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \text{vec}(\Phi), \kappa, \text{vec}(P))$ represents all the model parameters, with $P = \{p_{ij}\}_{i,j=1}^J$ the matrix with transition probabilities. The likelihood function conditional only on the model parameters can be obtained by summing (11) over all the possible states as

$$f(Y^T|Y_1, \theta) = \sum_{S_{1,1}=1}^J \sum_{S_{1,2}=1}^J \dots \sum_{S_{1,T}=1}^J f(Y^T, S^T|Y_1, \theta). \quad (12)$$

3.2 Prior distributions

To estimate the models parameters together with the unobserved regimes, we would like to obtain posterior results that are driven by the data rather than by the prior distributions. Therefore, we impose rather diffuse prior specifications for the model parameters.

The parameters of key interest are the lead/lag times or phase shifts κ_j , $j = 1, \dots, J$. For those, we use a discrete uniform prior where we assign equal probability

to each possible value of κ in a predefined set \mathcal{C} , that is

$$f(\kappa) \propto \begin{cases} 1 & \text{for all } \kappa \in \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

The specification of the set \mathcal{C} determines the type of synchronization that is assumed. For example, when $\mathcal{C} = \{\kappa \in \mathbb{Z}^J \mid -c_j \leq \kappa_j \leq c_j, j = 1, \dots, J\}$ for certain positive valued $c_j, j = 1, \dots, J$, we allow for imperfect synchronization with the phase shifts being restricted only by the bounds c_j . Additional restrictions may be imposed such that, for example, the difference between the phase shifts of the distinct regimes cannot exceed a certain threshold d . In this case, \mathcal{C} can be specified as $\mathcal{C} = \{\kappa \in \mathbb{Z}^J \mid -c_j \leq \kappa_j \leq c_j, |\kappa_i - \kappa_j| \leq d, i, j = 1, \dots, J\}$. Note that setting d equal to zero (while $c_j > 0$ for $j = 1, \dots, J$) results in a model where phase shift parameters are identical across regimes, that is, corresponding with the specification of Hamilton and Perez-Quiros (1996) given in (4). Setting $c_j = 0$ we obtain an MS-VAR model with perfect synchronization as in (3).

For the transition probabilities we use a uniform distribution on the unit interval $(0,1)$, that is

$$f(p_{ij}) = \mathbb{I}[0 < p_{ij} < 1] \quad \text{for } i, j = 1, \dots, J. \quad (14)$$

When this non-informative prior for the transition probabilities is used special attention must be paid to the prior specifications of the regime-dependent parameters. This follows from the fact that the value of the complete data likelihood function is the same if we switch all regime dependent parameters together with the corresponding transition probabilities. This ‘label switching problem’ complicates proper posterior analysis as the posterior distributions of the regime dependent parameters become multimodal, see Frühwirth-Schnatter (2001) and Geweke (2007) for discussion. To circumvent this problem, we define the prior for the regime dependent mean

parameters $\mu = (\mu'_1, \mu'_2)'$ in such a way that it identifies the regimes. Specifically, we set the prior specification for μ as

$$f(\mu) \propto \begin{cases} 1 & \text{if } \mu_1 \in \{\mu_1 \in \mathbb{R}^J \mid \mu_{1,1} < \dots < \mu_{1,j} < \dots < \mu_{1,J}\} \\ 0 & \text{elsewhere.} \end{cases} \quad (15)$$

For the autoregressive coefficients we use flat priors

$$f(\Phi) \propto 1, \quad (16)$$

when the characteristic roots of Φ lie outside the unit circle and 0 otherwise.

Finally, for the regime-dependent variance parameters as well as for the correlation parameters we take uninformative priors

$$f(\sigma_{l,j}^2) \propto \frac{1}{\sigma_{l,j}^2} \quad \text{for } l = 1, 2 \text{ and } j = 1, \dots, J, \quad (17)$$

$$f(\rho_j) \propto (1 - \rho_j^2)^{-\frac{3}{2}}, \quad j = 1, \dots, J. \quad (18)$$

These priors can be derived by decomposing a covariance matrix with inverted Wishart distribution where the degrees of freedom approaches to zero into the conditional distributions of variances and correlations.

3.3 Posterior simulation

The posterior distribution for the model parameters is proportional to the product of the likelihood function (12) and the joint prior $f(\theta)$, with the latter obtained as the product of (14)–(18). Gibbs sampling together with data augmentation leads to the following algorithm to draw from the joint posterior distribution:

1. Sample μ from $f(\mu \mid \Phi, \sigma^2, \rho, S^T)$

2. Sample Φ from $f(\Phi|\mu, \sigma^2, \rho, S^T)$
3. Sample σ_l^2 from $f(\sigma_l^2|\mu, \Phi, \rho, S^T)$ for $l = 1, 2$.
4. Sample ρ_j from $f(\rho_j|\mu, \Phi, \sigma_{i,j}^2, S^T)$ for $j = 1, \dots, J$
5. Sample p_{ij} from $f(p_{ij}|S_1^T)$, for $i, j = 1, \dots, J$.
6. Sample κ_j from $f(\kappa_j|\mu, \Phi, \sigma^2, \rho, S_1^T)$ for $j = 1, \dots, J$.
7. Sample S_1^T from $f(S_1^T|\mu, \Phi, \sigma^2, \rho, \kappa)$

Details on the derivation of the conditional posterior distributions resulting from the likelihood function and prior distributions are given in Appendix A.

3.4 Model selection

The multistate MS-VAR model with imperfect synchronization with asymmetric phase shifts provides a flexible and intuitive framework for characterizing partial synchronization of cycles in different macroeconomic and financial variables. However, moving away from more restricted but more parsimonious forms of synchronization may also result in overfitting. Therefore, we want to establish the potential of the model by comparing it with more restricted (nested) alternative models, in particular MS-VAR models with symmetric phase shifts and with perfect synchronization.

Selecting the best model is not an easy task when the competing alternatives embody regime switching dynamics. In most cases standard testing procedures apply only when the number of regimes is the same in the models being compared. In this case marginal likelihood based comparisons, e.g. Bayes factors, can be implemented. When the task is determination of the number of the regimes, however, a complication arises when improper priors are used for the regime dependent parameters of interest. In this case, Bayes factors are not properly defined and tend to select the more parsimonious model (see Gelfand and Dey, 1994, for details). An alternative

way of model comparison is by means of predictive Bayes factors computed using predictive likelihoods. This approach has the advantage that it is not affected by the choice of prior distributions and overfitting, while it is directly related to the posterior model probabilities. For a given model, the predictive likelihood of the observation at $t_0 + 1$, Y_{t_0+1} , conditional on the previous observations Y^{t_0} , is given by

$$f(Y_{t_0+1}|Y^{t_0}) = \int f(Y_{t_0+1}|\theta)f(\theta|Y^{t_0})d\theta, \quad (19)$$

where $p(\theta|Y^{t_0})$ is the posterior distribution of the model parameters θ given the observations until t_0 , and $p(Y_{t_0+1}|\theta)$ is the density of the observation y_{t_0+1} , which can be written as

$$f(Y_{t_0+1}|\theta) = \sum_{j=1}^J f(Y_{t_0+1}|S_{1,t_0+1} = j, \theta)f(S_{1,t_0+1} = j|\theta, Y^{t_0}). \quad (20)$$

We can use the posterior simulator to obtain the distribution of the model parameters and estimate the predictive likelihood by $G^{-1} \sum_{g=1}^G f(Y_{t_0+1}|Y^{t_0}, \theta^{(g)})$ where G is a large number of draws from the posterior distribution. This can be extended to compute the predictive likelihood of a sequence $Y^{t_0+1, t_0+h} = (Y'_{t_0+1}, \dots, Y'_{t_0+h})'$ where

$$f(Y^{t_0+1, t_0+h}|Y^{t_0}) = \int \prod_{l=1}^h f(Y_{t_0+l}|Y^{T_0+l-1}, \theta)f(\theta|Y^{t_0})d\theta, \quad (21)$$

see Frühwirth-Schnatter (2006), for details.

4 Lead-lag effects in size-based portfolio returns

In this section, we use the MS-VAR models to analyze the degree of synchronization in monthly returns of stock portfolios with different levels of market capitalization. Empirical evidence convincingly has demonstrated that asset returns are not normally distributed. Markov-switching models provide an interesting approach for

describing this non-normality, by means of a mixture of underlying normal distributions, see Timmermann (2000) and Perez-Quiros and Timmermann (2001), among others. A further attractive property of Markov-switching models is that they formalize the idea of cyclical variation in financial markets, with phases corresponding with periods of bull and bear sentiment, for example, during which asset prices have quite different characteristics. Not surprisingly then, Markov-switching models have become increasingly popular in modeling and forecasting asset prices. Previous research focuses mostly on univariate modeling of asset returns. This generally leads to the finding that a model with two or three regimes captures the univariate dynamics in aggregate stock returns and identifies bull and bear markets successfully (see Turner *et al.*, 1989; Chauvet and Potter, 2000; Maheu and McCurdy, 2000; Guidolin and Timmermann, 2006, among others). Relatively fewer papers focus on multivariate settings analyzing the dependence between different asset returns (see Perez-Quiros and Timmermann, 2000; Ang and Bekaert, 2002; Guidolin and Timmermann, 2006, for example). More regimes typically seem to be required for an adequate representation of the joint return distribution. For example, using excess returns of size-based stock portfolios and bonds, Guidolin and Timmermann (2006) show that a four regime model representing bear, bull, recovery and crashes performs better than competing models with fewer regimes.

Previous applications of multivariate regime-switching models to asset returns assume perfect synchronization in the cyclical dynamics of the included assets. While it does not seem unreasonable to represent the idea of regimes in financial markets by means of a single common cycle, the assumption that regime switches occur contemporaneously for different assets may not be appropriate. In particular, a substantial body of empirical research has documented the presence of lead-lag effects between different types of stocks. Determinants of the lead-lag effect include size, analyst coverage, institutional ownership and trading volume, see Lo and MacKinlay (1990);

Brennan *et al.* (1993); Badrinath *et al.* (1995); Chordia and Swaminathan (2000), among others. Focusing on size, the typical finding is that returns on (portfolios of) large stocks lead returns on (portfolios of) small stocks, see Lo and MacKinlay (1990); McQueen *et al.* (1996); Richardson and Peterson (1999); Hou (2007), among others. Several explanations have been put forward for this size-based lead-lag effect, including time-varying expected returns and market microstructure biases such as infrequent trading, see Boudoukh *et al.* (1994). Most empirical evidence, however, points in the direction of the tendency of small-caps to adjust more slowly to new (common) information than large caps as the underlying reason. Previous studies typically examine the lead-lag effect in size-based portfolios using cross-autocorrelations within a vector autoregressive framework. Interestingly, McQueen *et al.* (1996) and Hou (2007) document a ‘directional asymmetry’ in the lead-lag effect. Specifically, McQueen *et al.* (1996) find that large caps lead small caps only in response to good news, but not bad news. Hou (2007) reports the opposite effect, namely that small stocks adjust slowly to negative news but not to positive news. Here we examine whether the lead-lag effect between small caps and large caps can also be established when Markov-switching models are used for describing their joint distribution. In particular, we adopt our framework where the regime dynamics are governed by a single common cycle that is subject to phase shifts, which possibly are different across regimes.

We use our MS-VAR modeling framework to examine the presence of lead-lag effects in size-based portfolios based on all NYSE, AMEX, and NASDAQ stocks over the period July 1963 - December 2007. We consider monthly returns on equal-weighted large-cap and small-cap portfolios, consisting of the 20% largest and smallest stocks (sorted according to market capitalization), respectively.¹ We use the

¹The data is obtained from Kenneth French’s website data library <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>

returns in excess of the risk-free rate.²

We estimate MS-VAR models with two, three and four regimes. For the relation between the regime dynamics in the two portfolio returns series, we consider the four possibilities discussed in Section 2, namely (i) independent cycles (IND); (ii) perfect synchronization (PS); (iii) imperfect synchronization with a single phase shift of the common cycle (SPS); and (iv) imperfect synchronization with regime-specific phase shifts (APS). For the specification of the covariance matrix, we examine the two characterizations discussed in the Section 2. On the one hand, we impose perfect synchronization in the variance dynamics, as in (7). Here we assume that the regime dynamics of the covariance matrix are governed by the Markov process that also drives the regime dynamics of the expected excess returns on the large-cap portfolio. Note that the large-cap portfolio resembles the market index much more closely than the small-cap portfolio and, therefore, is a more adequate measure to capture the dynamics of volatility and correlation in the stock market. This model is denoted as Σ -PS, as we assume perfect synchronization of the portfolios in terms of their covariance dynamics. On the other hand, we assume that the expected excess return and the variance of each portfolio have identical regime dynamics, as in (8). In this case, the regime switches of the correlation are still determined by the regime indicator of the large-cap portfolio.³ This covariance matrix specification is indicated by Σ -APS, as it allows for imperfect synchronization with asymmetric phase shifts (APS) in the volatility dynamics of the portfolios.

For the models with imperfect synchronization, we specify the set of admissible phase shift parameters \mathcal{C} such that all κ_j 's, $j = 1, \dots, J$, are in the interval $[-2, 2]$,

²For the risk-free rate we use Fama-Bliss risk free rates from CRSP (Center for Research in Security Prices). This series is discontinued in December 2001. For the remaining part of the sample period we use the 4-week T-Bill rate measured at close on last trading day of each month from the St. Louis Federal Reserve(<http://research.stlouisfed.org/fred2/>).

³We also estimate models with the correlation regime being determined by the regime of the small-cap portfolio's expected return and volatility. Overall these models perform worse in terms of marginal likelihood and predictive likelihood. Detailed results are available upon request.

i.e. $c_j = 2$, as it is unlikely that the delay exceeds two months in financial assets. We also restrict the differences between the phase shift parameters for different regimes to be one month at most, i.e. $d = 1$.

Finally, we do not include autoregressive dynamics in any of the models. This may seem restrictive, especially in light of Boudoukh *et al.* (1994) who argue that lead-lag effects among size-based portfolios may be spurious as they are in fact, to a large extent, due to contemporaneous correlation and own-autocorrelation of small-cap stocks. This suggests a large stock portfolio does not necessarily lead a small stock portfolio once lagged small-cap portfolio returns are included to account for these autocorrelation effects. With this in mind, we also estimate all models with the first lag of the small-cap returns (as well as with unrestricted first-order autoregressive dynamics) included. The results, especially those related to lead-lag parameters, are very similar to those reported here. Here we do not display these results for the sake of brevity.⁴ In order to achieve a robust model selection we compute both the marginal likelihood using the whole sample period as well as the predictive likelihood. For evaluating the predictive likelihood we use the sample from July 1963 until December 2000 for estimation of the model and the remaining period from January 2001 until December 2007 to compute the predictive likelihood. For the complete out-of-sample period the predictive likelihood is computed using the posterior distribution based on the initial estimation sample, as recursive updating is not feasible in terms of computation time. The posterior distributions are based on 50 000 iterations of the Gibbs sampling algorithm. The first 20 000 draws are discarded as burn-in sample and the final 30 000 draws are used for inference and for predictive likelihood computation. The convergence of the MCMC sampler is checked using statistical and visual inspection and in all the models convergence is assured.

⁴Detailed results are available upon request.

Table 1 displays the log marginal likelihood and log predictive likelihood values of the competing models. Table 1 offers five clear-cut conclusions. First, independent of the number of regimes and the specification of the covariance matrix, the models with a (possibly shifted) common cycle convincingly outperform the model with independent cycles in terms of marginal likelihood. The differences become larger when the number of regimes increases or when perfect synchronization of the covariance dynamics is imposed. The model with independent cycles also performs worse in terms of predictive likelihood, with the exception of the two regime model with imperfect synchronization in the volatilities.

Second, independent of the number of regimes and the specification of the relation between the cycles in the two portfolio returns, the models with imperfect synchronization (Σ -APS) achieve higher marginal likelihood and predictive likelihood values than the corresponding models with perfect synchronization (Σ -PS). The only exception is the two-regime model with symmetric phase shifts (and the two-regime model with asymmetric regime shifts, but only in terms of predictive likelihood). Hence, the assumption of perfect synchronization in the variance dynamics does not seem appropriate.

Third, independent of the number of regimes, for the Σ -APS models we find that the marginal likelihood and predictive likelihood values for the specifications with perfect synchronization and with imperfect synchronization due to symmetric phase shifts are (almost) identical. This occurs because the phase shift parameters in the SPS models are estimated to be equal to zero, which reduces this specification to the PS model.

Fourth, independent of the number of regimes, for the Σ -APS models we find that the marginal likelihood and predictive likelihood values for the specifications with imperfect synchronization due to asymmetric phase shifts are (substantially) higher than those for the PS and SPS specifications. This suggests that allowing for

Table 1: Comparison of the models estimated using LC and SC Data

	<u>Log marg. likelihood</u>		<u>Log pred. likelihood</u>	
	Σ -PS	Σ -APS	Σ -PS	Σ -APS
Panel A: 2 regimes				
IND	-3356.3	-3306.5	-465.1	-461.7
PS	-3220.7	-3218.8	-465.0	-464.8
SPS	-3206.5	-3218.8	-462.8	-464.8
APS	-3211.3	-3210.4	-461.2	-463.9
Panel B: 3 regimes				
IND	-3508.4	-3443.6	-464.7	-470.5
PS	-3283.2	-3257.1	-465.9	-463.2
SPS	-3213.5	-3257.1	-464.9	-463.2
APS	-3225.9	-3206.1	-463.7	-459.6
Panel C: 4 regimes				
IND	-3984.6	-3770.8	-469.8	-475.2
PS	-3322.4	-3276.0	-463.1	-463.3
SPS	-3322.4	-3276.0	-463.1	-462.6
APS	-3326.9	-3230.5	-463.5	-461.6

Note: The table presents log marginal likelihoods and log predictive likelihoods of the competing models with (i) different number of regimes; (ii) different types of synchronization in expected returns, and (iii) different types of regime switching of the variances and correlations. The models are applied to monthly excess returns on large-cap and small-cap portfolios for the sample period July 1963-December 2001. Marginal likelihoods are those obtained when the models are estimated for the complete sample period. Predictive likelihood values are for the period January 2001-December 2007, which are computed conditional on the posterior distributions of the model parameters obtained with an estimation sample ending in December 2000. The regime dynamics in the expected excess returns correspond with (i) independent cycles (IND); (ii) perfect synchronization (PS); (iii) imperfect synchronization with a single phase shift of the common cycle (SPS); and (iv) imperfect synchronization with regime-specific phase shifts (APS). The specifications of the covariance matrix dynamics correspond with (i) perfect synchronization (Σ -PS), according to the regime indicator for the expected excess returns of the large-cap portfolio; and (ii) imperfect synchronization with regime-specific phase shifts (Σ -APS), where the correlation switches regimes according to the regime indicator for the expected excess returns of the large-cap portfolio. Posterior results for computing log marginal (predictive) likelihoods are based on 50 000 (30 000) simulations of which the first 20 000 are discarded for burn-in.

regime-specific lead-lag times is appropriate for the large-cap and small-cap portfolios. While the differences in likelihood values are relatively small for the two-regime models, they become substantial when multiple regimes are considered.

Fifth, focusing on the APS specification with Σ -APS, the model with three regimes offers the best performance in terms of predictive likelihood.⁵

In sum, the results in Table 1 lead us to select the model with three regimes and imperfect synchronization with asymmetric phase shifts in both the expected excess returns and the volatilities as the preferred specification. In the remainder of this section, we therefore focus on the estimation results for this model specification.

Table 2 summarizes the posterior results of the model parameters. These indicate that the three regimes in the model have distinct features. A general pattern is that in each regime the volatility of the small-cap portfolio is much higher than the corresponding volatility of the large-cap portfolio, as expected. However, this higher volatility of the small-cap portfolio comes with a higher mean return (indicating a risk-return trade-off) only for the first regime. Regime 1 is characterized by a large positive mean return for both portfolios together with low volatility (especially for large caps), capturing a ‘boom’ market. For both large-cap and small-cap portfolios regime 2 has mean return close to zero and relatively low volatility. An interesting result is that the volatility in the boom market is lower than the moderate regime for the large-cap portfolios. Regime 3 can be characterized as a ‘bear’ market, as it shows a substantially negative mean return accompanied by high volatility, indicating the large uncertainty during these time periods. Summarizing, the first and third regime capture the ‘turbulent’ times with abrupt fluctuations in stock

⁵Note that the model with three regimes not only achieves higher marginal likelihood and predictive likelihood values than the model with two regimes, but also than the model with four regimes. For the PS and SPS specifications this is even more pronounced, with the marginal likelihood value declining monotonically as the number of regimes increases. While this may seem surprising at first, it can be understood by recalling that marginal likelihood (or Bayes factors) tends to select the more parsimonious model when improper priors are used for the regime dependent parameters, as discussed in Section 3.4.

Table 2: Posterior results for three-regime MS-VAR model with asymmetric synchronization and regime-dependent heteroskedasticity

		Regime 1	Regime 2	Regime 3
Large caps	Mean $\mu_{1,j}$	3.772 (0.626)	0.114 (0.224)	-1.337 (0.964)
	Volatility $\sigma_{1,j}$	2.901 (0.395)	3.353 (0.181)	7.804 (0.795)
Small caps	Mean $\mu_{2,j}$	6.310 (0.767)	-0.519 (0.368)	-2.077 (1.480)
	Volatility $\sigma_{2,j}$	4.159 (0.650)	3.909 (0.241)	11.192 (1.228)
	Correlation ρ_j	0.305 (0.198)	0.724 (0.032)	0.709 (0.065)
		<u>Transition probabilities p_{ij}</u>		
	Regime 1	0.516 (0.010)	0.440 (0.098)	0.043 (0.041)
	Regime 2	0.064 (0.024)	0.875 (0.028)	0.061 (0.020)
	Regime 3	0.215 (0.081)	0.066 (0.057)	0.718 (0.091)
		<u>Phase shifts κ_j</u>		
		0.000 (0.000)	-1.000 (0.000)	-0.049 (0.217)

Note: The table presents posterior means and standard deviations (in parentheses) of parameters in the MS-VAR model with three regimes, imperfect synchronization with asymmetric phase shifts, and regime-dependent heteroskedasticity, estimated for monthly excess returns of large cap and small-cap portfolios over the period July 1963 - December 2007. The volatilities $\sigma_{l,j}$ have the same regime dynamics as the corresponding expected excess returns $\mu_{l,j}$, $l = 1, 2$. The correlation ρ_j has the same regime dynamics as the mean and volatility of the large-cap portfolio. Posterior results are based on 50 000 simulations of which the first 20 000 are discarded for burn-in.

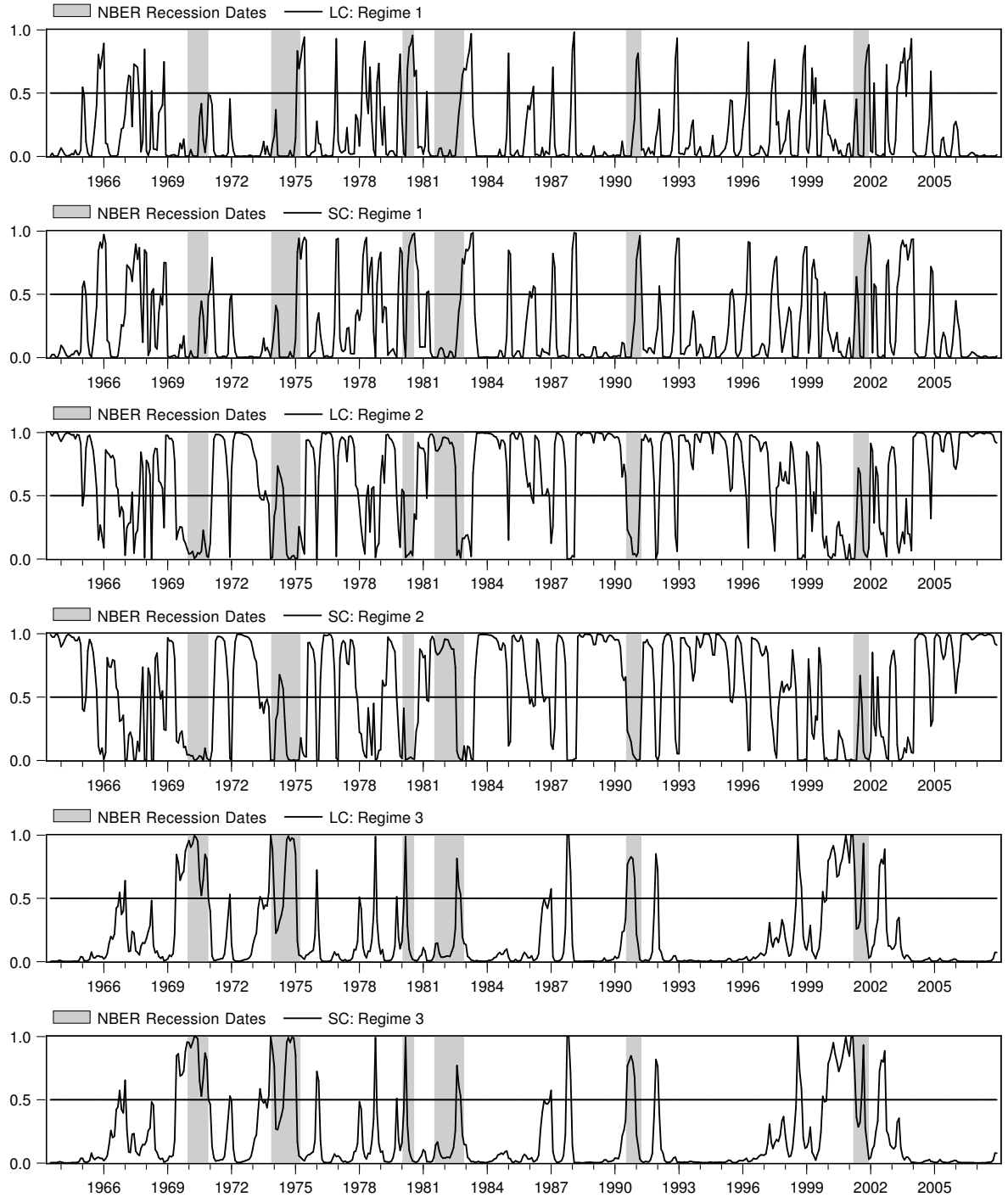
market and the second regime captures ‘normal’ times with either small gains or losses together with limited volatility.

The transition probability estimates also reveal an interesting pattern in the relative duration of the different regimes and the cyclical dynamics in the portfolio returns. The probability of staying in the boom regime is 0.52 and hence the duration of this regime is short, on average only two months. As the transition probability of 0.88 indicates, the average duration of the moderate regime 2 is longer than the other regimes with almost eight months. The probability of staying in the bear regime is 0.73 implying an average duration of about four months. Transition from the boom regime is ten times more likely to occur to the moderate regime than to the bear regime, with transition probability of 0.44 and 0.043, respectively. Hence, the boom regime typically does not end with a crash, but is followed by a period with moderate returns. By contrast, when the bear regime is left, this is much more likely to happen to the boom market regime than to the moderate regime, with transition probabilities of 0.215 and 0.066, respectively. This suggests that bear markets are most often followed by a (short) period with high returns, providing evidence of a ‘recovery’ or ‘bounce-back effect’ in these stock returns, similar to findings for the US business cycle in, for example, Kim *et al.* (2005).

We display the posterior regime probabilities in Figure 1. We also include recessions as dated by the NBER in shaded areas to elaborate the links between the regime timings with the business cycle.

First, from the graphs we can infer that the regimes are estimated quite precisely in the sense that generally the regime probabilities are close to zero or one, although there is some uncertainty about the regime assignment in a few periods. The conclusion drawn above about the regimes capturing turbulent and normal times is evident in Figure 1. The boom regime, for example, mostly occurs at the end of recessions (among others). In addition, except for the recession in 1981-2, the moderate regime

Figure 1: Posterior probabilities of regimes of LC and SC



appears considerably less frequently during recessions than during expansions. The bear market regime mostly occurs before and during the start of the recession. In addition, this regime captures all stock market crashes that occurred during the sample period, such as ‘Black Monday’ (October 1987), the Russia crises (August-September 1998), the collapse of the dot-com bubble (2000-1), and 9/11 (September 2001).

As an alternative way of assessing the regime dynamics, Figure 2 shows the implied model volatilities, obtained as the weighted average of the regime-specific volatilities with the posterior regime probabilities as weights, together with the monthly realized volatility constructed from the daily portfolio returns.⁶ It is seen that the patterns in realized volatility are captured fairly well by the model, in the sense that periods with a higher level of realized volatility generally are matched by increased ‘model-implied’ volatility.

Finally, we turn to the estimates of the phase shift parameters to assess the implications of the model concerning the lead-lag effects of the large-cap and small-cap portfolios. The posterior results of the phase shift parameters κ_j are shown in the bottom panel of Table 2 and in Figure 3.

The estimates provide convincing evidence for the presence of asymmetric phase shifts. Specifically, we find large and small stocks enter into the boom and bear regimes contemporaneously. For the boom regime we find that all posterior probability mass is located at $\kappa_1 = 0$. For regime 3, there is only a modest probability that $\kappa_3 = 1$. In contrast, for the moderate regime the large-cap portfolio leads the small-cap portfolio by one month. The posterior uncertainty for this parameter is very small.⁷

⁶Specifically, the monthly realized variance is computed as the sum of squared daily returns plus twice the sum of the cross-product of returns on consecutive trading days. The latter provides a correction for the presence of autocorrelation in the daily returns, see Zhou (1996).

⁷Clearly, the posterior standard deviations of the κ parameters are very small. Note that we impose strong priors to deal with the uncertainty embodied in the data but still this prior assigns equal probability to all types of synchronization. Hence, the result is not affected by this reasonably

Figure 2: Comparison of volatilities implied by the model and realized volatilities

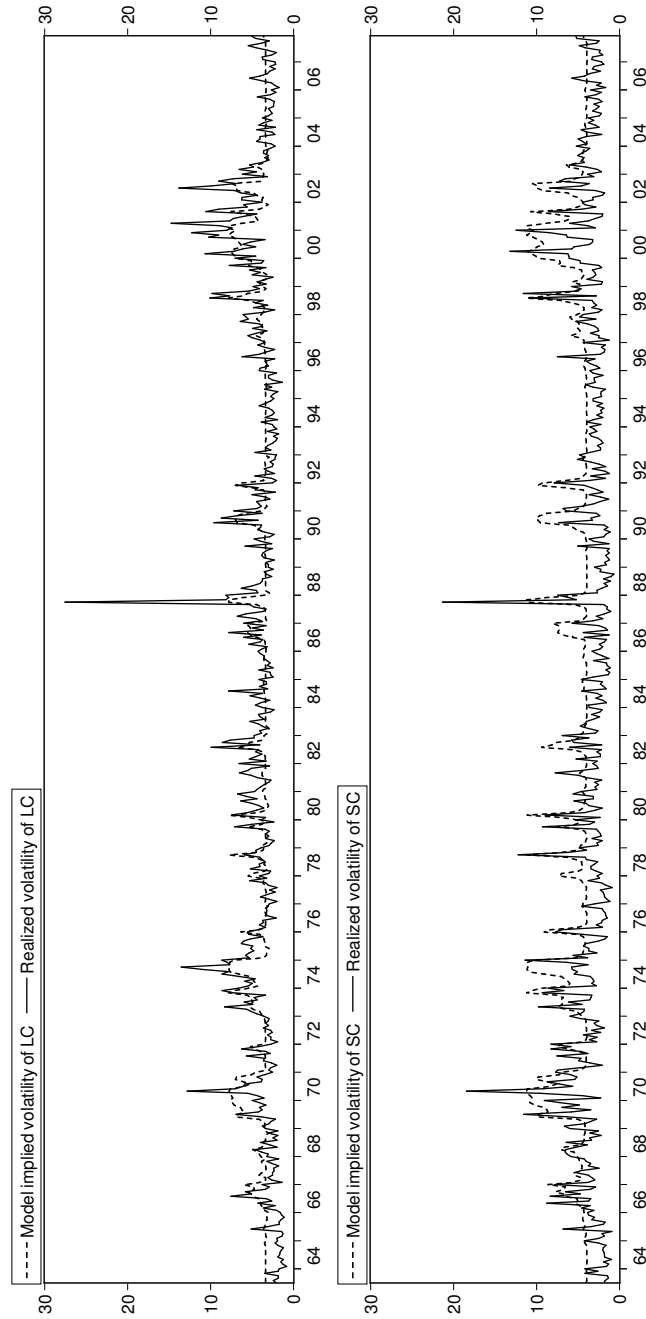
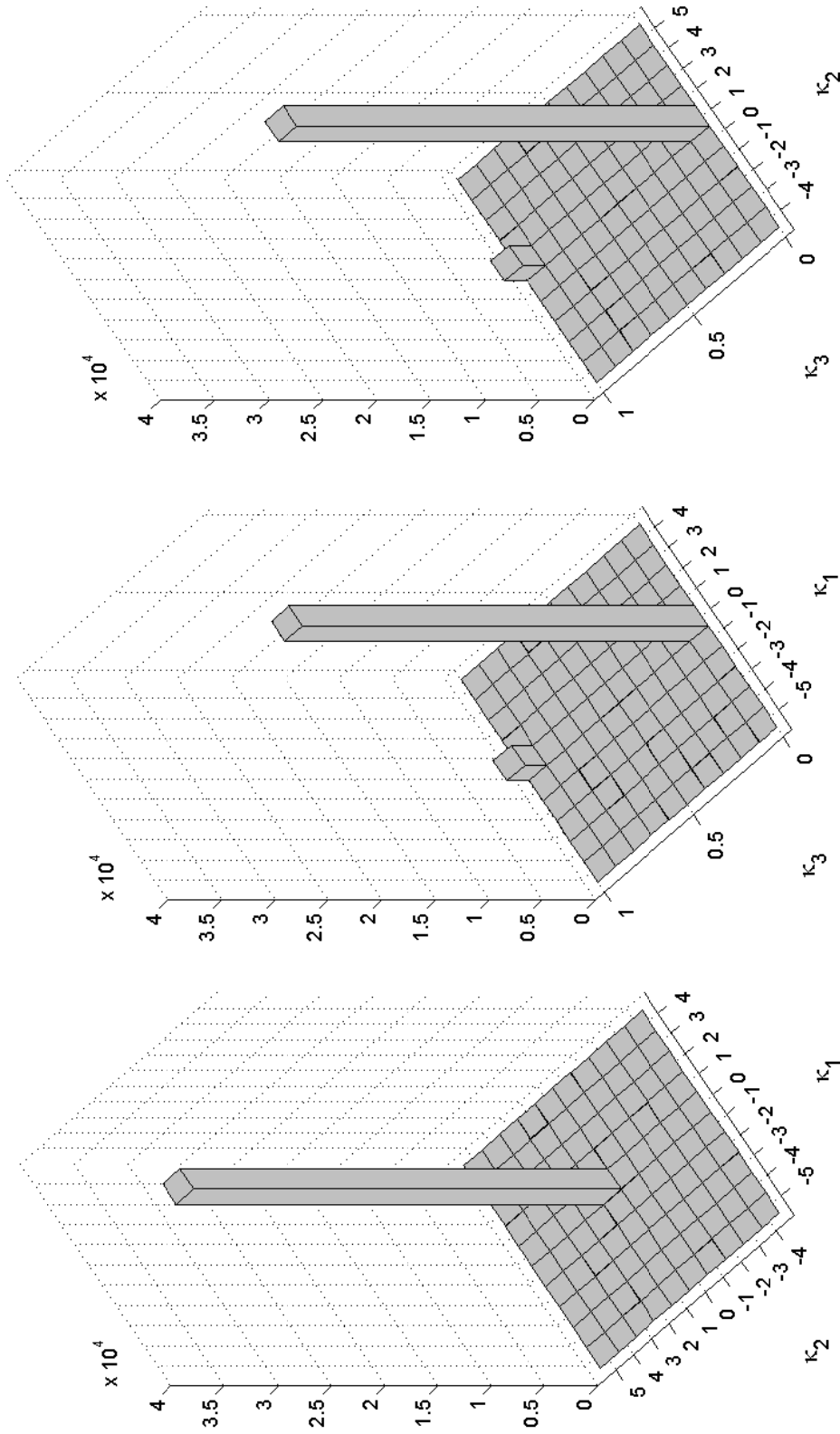


Figure 3: Distribution of lead/lag times for LC and SC regimes



We interpret our findings concerning the phase shift parameters as follows. The contemporaneous switches of the two portfolios to the boom and bear regimes suggest that during turmoil periods, with either good or bad news, information diffusion occurs rapidly and lead-lag effects are not present. The positive phase shift for the second regime indicates that news ‘travels slowly’ during more moderate times, in the sense that it is incorporated only gradually into small stocks, such that their prices adjust with a delay compared to large stocks. We note that our findings of asymmetric phase shifts are quite different to the conclusions reached by McQueen *et al.* (1996) and Hou (2007) concerning the asymmetric nature of lead-lag effects. McQueen *et al.* (1996) find that large caps lead small caps only in response to good news, but not bad news. Hou (2007) documents the opposite effect, namely that small stocks adjust slowly to negative news but not to positive news. Incorporating lead-lag effects in a regime-switching framework with multiple regimes, we obtain intuitively plausible results that possibly reconcile the findings of McQueen *et al.* (1996) and Hou (2007). Our posterior results suggest that information diffusion in fact occurs rapidly for both good and bad news items that have considerable impact on stock prices, in the sense that large-caps and small-caps switch to the extreme boom and bear regimes at the same time. Small-caps respond with a delay during moderate times, possibly due to the fact that less ‘exciting’ news during those periods is incorporated in their prices less rapidly.

5 Conclusion

In this paper we have developed a Markov-Switching vector autoregressive model that allows for imperfect synchronization of cyclical regimes in multiple variables, due to phase shifts of a single common cycle. The model has three key features. First,

strong prior, though the amount of variation may be affected by its strength. Nevertheless, the results with weaker priors are very similar to the results shown here. Detailed results are available upon request.

the amount of phase shift different across regimes (as well as across variables). This feature is empirically relevant given the evidence for asymmetries in lead-lag effects in different types of assets, as well as in leading and coincident economic indicators. Second, the cycle can consist of any number of regimes $J \geq 2$. Allowing for multiple regimes can be necessary for accurate modeling of cyclical dynamics in financial asset returns, as well as business cycle dynamics in macroeconomic variables. Third, the model allows for regime-dependent volatilities and correlations. This is particularly important for identification of regimes in asset prices, where differences in volatilities (and correlations) often are more pronounced than differences in mean returns.

In an empirical application to monthly returns on size-based stock portfolios, we found that a three-regime model with asymmetric phase shifts and regime-dependent heteroscedasticity characterizes the joint distribution of returns most adequately. While large- and small-cap portfolios switch contemporaneously into boom and crash regimes, the large-cap portfolio leads the small-cap portfolio for switches to a moderate regime by a month. This suggests that during turmoil periods, with either good or bad news, information diffusion occurs rapidly and lead-lag effects are not present. News ‘travels slowly’ during more moderate times, in the sense that it is incorporated only gradually into small stocks, such that their prices adjust with a delay compared to large stocks.

Appendix A Conditional Posterior Distributions

In this appendix, we derive the posterior distributions that we use in the sampling scheme described in Section 3.3. For notational convenience, let $\mathcal{S}_{1,t}$ and $\mathcal{S}_{2,t}$ be the $J \times 1$ vectors of indicator functions for the first and second variable, respectively. The elements in these vectors take the values 0 or 1, indicating which of the J regimes occurs at time t .

We write the model (4) in matrix notation

$$Y_t - \mathcal{S}'_t \mu = \Phi(Y_{t-1} - \mathcal{S}'_{t-1} \mu) + \mathcal{E}_t, \quad \text{with } \mathcal{E}_t \sim N(0, \Sigma_t), \quad (\text{A.1})$$

where Y_t and \mathcal{E}_t are defined below (9) and

$$\mathcal{S}_t = \begin{pmatrix} \mathcal{S}_{1,t} & \mathbf{0}_J \\ \mathbf{0}_J & \mathcal{S}_{2,t} \end{pmatrix}, \quad (\text{A.2})$$

with $\mathbf{0}_J$ a $J \times 1$ vector of zeros.

A.1 Sampling of μ

To sample μ we first rewrite (A.1) as

$$\Sigma_t^{-\frac{1}{2}} (Y_t - \Phi Y_{t-1}) = \Sigma_t^{-\frac{1}{2}} (\mathcal{S}'_t - \Phi \mathcal{S}'_{t-1}) \mu + \Sigma_t^{-\frac{1}{2}} \mathcal{E}_t. \quad (\text{A.3})$$

This is a regression in the form of

$$Z_t = X_t \mu + \epsilon_t \quad \text{with } \epsilon_t \sim N(0, I_2), \quad (\text{A.4})$$

and hence the conditional distribution of μ is multivariate normal with mean $(X'X)^{-1} X'Z$ and variance $(X'X)^{-1}$, where $X = (X'_2, \dots, X'_T)'$ and $Z = (Z'_2, \dots, Z'_T)'$, see

Zellner (1971). The restrictions on the elements of μ that are imposed for identification of the regimes, as discussed in Section 3.2, can be applied by sampling from the corresponding truncated distribution or by using acceptance rejection sampling.

A.2 Sampling of Φ

Conditional on the remaining model parameters, the model in (A.1) is a multivariate regression of the form

$$Z = \Phi X + U, \quad (\text{A.5})$$

where $Z = (Z_2, \dots, Z_T)$ with $Z_t = Y_t - \mathcal{S}'_t \mu$, and $X = (X_2, \dots, X_T)$ with $X_t = (Y_{t-1} - \mathcal{S}'_{t-1} \mu)$. Using the fact that $\text{vec}(\Phi X) = (X' \otimes I_2) \text{vec}(\Phi)$, we can write this as a univariate regression

$$z = (X' \otimes I_2) \text{vec}(\Phi) + u, \quad (\text{A.6})$$

where $z = \text{vec}(Z)$ and $u = \text{vec}(U)$. It then follows that the conditional posterior distribution of $\text{vec}(\Phi)$ is multivariate normal with mean $((X' \otimes I_2)' \Omega^{-1} (X' \otimes I_2))^{-1} ((X' \otimes I_2)' \Omega^{-1} z)$ and covariance matrix $((X' \otimes I_2)' \Omega^{-1} (X' \otimes I_2))^{-1}$, where

$$\Omega^{-1} = \text{Cov}(u)^{-1} = \text{diag}(\Sigma_2^{-1}, \dots, \Sigma_T^{-1}).$$

A.3 Sampling of Variances and Correlations

To sample the variances we decompose the multivariate normal distribution of \mathcal{E}_t into the conditional distribution of $\varepsilon_{2,t}$ given $\varepsilon_{1,t}$ and the marginal distribution of $\varepsilon_{1,t}$. This results in

$$\prod_{t=2}^T f(\mathcal{E}_t) = \prod_{t=2}^T \frac{1}{\sigma_{1,t}} \phi\left(\frac{\varepsilon_{1,t}}{\sigma_{1,t}}\right) \frac{1}{\sigma_{2,t} \sqrt{(1 - \rho_t^2)}} \phi\left(\frac{\varepsilon_{2,t} - \rho_t \varepsilon_{1,t}}{\sigma_{2,t} (1 - \rho_t^2)}\right), \quad (\text{A.7})$$

where $\sigma_{l,t}^2$ denotes the variance of $\varepsilon_{l,t}$, $l = 1, 2$, which follows from (7) or (8). Hence, using standard results from a linear regression model with a diffuse prior for the variance, it follows that the conditional posterior distribution of $\sigma_{l,j}^2$, with $l = 1, 2$ and $j = 1, \dots, J$, is an inverted χ^2 distribution with scale parameter $\sum_{t=2}^T I[S_{l,t} = j] \varepsilon_{l,t}^2$ and with $\sum_{t=2}^T I[S_{l,t} = j]$ degrees of freedom.

To sample ρ_j from its conditional posterior distribution we can again use (A.7).

It is easy to see that only the time periods for which $S_{1,t} = j$ are relevant

$$(1 - \rho_j^2)^{-\frac{3}{2}} \prod_{t=2}^T \left(\frac{1}{\sqrt{(1 - \rho_t^2)}} \phi \left(\frac{\varepsilon_{2,t} - \rho_t \varepsilon_{1,t}}{\sigma_{2,t} \sqrt{(1 - \rho_t^2)}} \right) \right)^{I[S_{1,t}=j]}. \quad (\text{A.8})$$

We can easily implement the gridgy Gibbs sampler approach of Ritter and Tanner (1992). Given that $\rho_j \in (-1, 1)$ for all $j = 1, \dots, J$, we can setup a grid in this interval based on the precision we desire about the value of ρ_j .

A.4 Sampling of P

From the conditional likelihood function (12), it follows that the transition probabilities from state i can be written as

$$f(p_{i,1}, \dots, p_{i,J}) \propto \prod_{j=1}^J p_{i,j}^{T_{i,j}} \quad \text{for } i = 1, \dots, J, \quad (\text{A.9})$$

where T_{ij} denotes the number of transitions from state i to state j . This corresponds to the kernel of a Dirichlet distribution and hence the transition probabilities can be sampled from a Dirichlet distribution with parameters T_{ij} for $j = 1, \dots, J - 1$. The transition probability to state J , p_{iJ} , follows from the restriction that $\sum_{j=1}^J p_{ij} = 1$.

A.5 Sampling of Lead/Lag Parameters κ

As the κ_j parameters can only take discrete values we can compute the posterior probabilities for all $\kappa \in \mathcal{C}$ and sample from a multinomial distribution. We can sample all κ_j parameters at once or conditional on each other, one at a time. When the number of admissible κ 's is large, the latter approach may be more attractive as the number of combinations increases only linearly in the number of states, whereas it increases exponentially in the first case.

A.6 Sampling of Regimes

The conditional posterior density of $S_{1,t}$ denoted by $f(S_{1,t}|S_1^{-t}, \theta, Y^T)$ for $t = 1, \dots, T$ and $S_1^{-t} = S_1^T \setminus \{S_{1,t}\}$, is proportional to the transition probabilities due to the Markov structure and to the density of Y conditional on the regimes. Hence, we can write the posterior density of $S_{1,t}$

$$f(S_{1,t}|S_1^{-t}, \theta, Y^T) \propto f(S_{1,t}|S_{1,t-1}, \theta) f(S_{1,t+1}|S_{1,t}, \theta) \prod_{i=t-\kappa_{\min}}^{t+1+\kappa_{\max}} f(Y_i|Y^{i-1}, S^i, \theta), \quad (\text{A.10})$$

where $f(Y_t|Y^{t-1}, S^t, \theta)$ is given in (10), $\kappa_{\max} = \max(\kappa_1, \dots, \kappa_J)$ and $\kappa_{\min} = \min(\kappa_1, \dots, \kappa_J)$.

At time $t = T$ the term $f(S_{1,t+1}|S_{1,t}, \theta)$ drops out. The regime at $t = 1$ can be sampled from the conditional distribution

$$f(S_{1,1}|S_1^{-1}, \theta, Y^T) \propto f(S_1|\theta) f(S_{1,2}|S_{1,1}, \theta) \prod_{i=2}^{t+1+\kappa_{\max}} f(Y_i|Y^{i-1}, S^i, \theta), \quad (\text{A.11})$$

where the unconditional density $f(S_{1,1}|\theta)$ follows a multinomial distribution with ergodic probabilities of the Markov chain.

Sampling of the state variables can be implemented by starting from the most recent value of S_1^T and sampling the states backward in time, one after another. After each step, the t^{th} element of S_1^T is replaced by its most recent draw.

References

- Ang, A. and G. Bekaert (2002), International Asset Allocation With Regime Shifts, *Review of Financial Studies*, 15, 1137–1187.
- Artis, M., H.-M. Krolzig, and J. Toro (2004), The European Business Cycle, *Oxford Economic Papers*, 56, 1–44.
- Artis, M. J., Z. G. Kontolemis, and D. R. Osborn (1997), Business Cycles for G7 and European Countries, *Journal of Business*, 70, 249–279.
- Badrinath, S., J. Kale, and T. Noe (1995), Of Shepherds, Sheep, and the Cross-Autocorrelations in Equity Returns, *Review of Financial Studies*, 8, 401–430.
- Barnard, J., R. McCulloch, and X. Meng (2000), Modeling Covariance Matrices In Terms Of Standard Deviations And Correlations, With Application To Shrinkage, *Statistica Sinica*, 10, 1281–1311.
- Bekaert, G., C. R. Harvey, and A. Ng (2005), Market Integration and Contagion, *Journal of Business*, 78, 39–70.
- Board, C. (2001), *Business Cycle Indicators Handbook*, available at <http://www.conference-board.org/publications/describebook.cfm?id=852>.
- Boldin, M. (1996), A Check on the Robustness of Hamilton’s Markov Switching Model Approach to the Economic Analysis of the Business Cycle, *Studies in Non-linear Dynamics & Econometrics*, 1, 35–46.
- Bordo, M. D. and T. Helbling (2003), Have National Business Cycles Become More Synchronized?, nBER Working Paper No. 10130.
- Boudoukh, J., M. P. Richardson, and R. F. Whitelaw (1994), A Tale of Three Schools: Insights on Autocorrelations of Short-Horizon Stock Returns, *Review of Financial Studies*, 7, 539–573.
- Brennan, M., N. Jegadeesh, and B. Swaminathan (1993), Investment Analysis and the Adjustment of Stock Prices to Common Information, *Review of Financial Studies*, 6, 799–824.
- Candelon, B., J. Piplack, and S. Straetmans (2008), On Measuring Synchronization of Bulls and Bears: The Case of East Asia, *Journal of Banking and Finance*, 32, 1022–1035.
- Candelon, B., J. Piplack, and S. Straetmans (2009), Multivariate Business Cycle Synchronization in Small Samples, *Oxford Bulletin of Economics and Statistics*, 71, 715–737.
- Canova, F., M. Ciccarelli, and E. Ortega (2007), Similarities and Convergence in G-7 Cycles, *Journal of Monetary Economics*, 54, 850–878.

- Chauvet, M. and S. Potter (2000), Coincident and Leading Indicators of the Stock Market, *Journal of Empirical Finance*, 7, 87–111.
- Chordia, T. and B. Swaminathan (2000), Trading Volume and Cross-Autocorrelations in Stock Returns, *Journal of Finance*, 55, 913–935.
- Clements, M. P. and H.-M. Krolzig (2003), Business Cycle Asymmetries: Characterization and Testing Based on Markov-Switching Autoregressions, *Journal of Business & Economic Statistics*, 21, 196–211.
- Del Negro, M. and C. Otrok (2008), Dynamic Factor Models with Time-Varying Parameters: Measuring Changes in International Business Cycles, Staff Reports 326, Federal Reserve Bank of New York.
- Edwards, S., J. G. Biscarri, and F. Perez de Gracia (2003), Stock Market Cycles, Financial Liberalization and Volatility, *Journal of International Money and Finance*, 22, 925–955.
- Frühwirth-Schnatter, S. (2001), Markov Chain Monte Carlo Estimation of Classical and Dynamic Switching and Mixture Models, *Journal of the American Statistical Association*, 96, 194–209.
- Frühwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*, New York: Springer.
- Gelfand, A. E. and D. K. Dey (1994), Bayesian Model Choice: Asymptotics and Exact Calculations, *Journal of the Royal Statistical Society, Series B*, 56, 501–514.
- Geman, S. and D. Geman (1984), Stochastic Relaxations, Gibbs Distributions, and the Bayesian Restoration of Images, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 6, 721–741.
- Geweke, J. (2007), Interpretation and Inference in Mixture Models: Simple MCMC works, *Computational Statistics & Data Analysis*, 51, 3529–3550.
- Guidolin, M. and A. Timmermann (2006), An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns, *Journal of Applied Econometrics*, 21, 1–22.
- Hamilton, J. D. and G. Perez-Quiros (1996), What Do the Leading Indicators Lead?, *Journal of Business*, 69, 27–49.
- Harding, D. and A. Pagan (2006), Synchronization of Cycles, *Journal of Econometrics*, 132, 59–79.
- Hou, K. (2007), Industry Information Diffusion and the Lead-Lag Effect in Stock Returns, *Review of Financial Studies*, 20, 1113–1138.

- Kim, C.-J., J. Morley, and J. Piger (2005), Nonlinearity and the Permanent Effects of Recessions, *Journal of Applied Econometrics*, 20, 291–309.
- Kole, E., K. Koedijk, and M. Verbeek (2006), Portfolio Implications of Systemic Crises, *Journal of Banking & Finance*, 30, 2347–2369.
- Koopman, S. J. and J. V. E. Azevedo (2008), Measuring Synchronization and Convergence of Business Cycles for the Euro Area, UK and US, *Oxford Bulletin of Economics and Statistics*, 70, 23–51.
- Koopman, S. J. and A. Harvey (1997), Multivariate Structural Time Series models, in B. H. C. Heij, H. Schumacher and C. Praagman (eds.), *Systematic Dynamics in Economic and Financial Models*, New York: John Wiley and Sons, pp. 269–298.
- Kose, M. A., C. Otrok, and C. H. Whiteman (2003), International Business Cycles: World, Region, and Country-Specific Factors, *American Economic Review*, 93, 1216–1239.
- Kose, M. A., C. Otrok, and C. H. Whiteman (2008), Understanding the Evolution of World Business Cycles, *Journal of International Economics*, 75, 110–130.
- Krolzig, H.-M. (1997), *Markov-Switching Vector Autoregressions*, New York: Springer.
- Lo, A. W. and A. C. MacKinlay (1990), Contrarian Profits and Stock Market Overreaction, *Review of Financial Studies*, 3, 175–205.
- Maheu, J. M. and T. H. McCurdy (2000), Identifying Bull and Bear Markets in Stock Returns, *Journal of Business & Economic Statistics*, 18, 100–112.
- Maheu, J. M., T. H. McCurdy, and Y. Song (2000), Extracting Bull and Bear Markets from Stock Returns, working paper, University of Toronto.
- McQueen, G., M. Pinegar, and S. Thorley (1996), Delayed Reaction to Good News and the Cross-Autocorrelation of Portfolio Returns, *Journal of Finance*, 51, 889–919.
- Paap, R., R. Segers, and D. v. Dijk (2009), Do Leading Indicators Lead Peaks More Than Troughs?, *Journal of Business & Economic Statistics*, 27, 528–543.
- Pelletier, D. (2006), Regime Switching for Dynamic Correlations, *Journal of Econometrics*, 131, 445–473.
- Perez-Quiros, G. and A. Timmermann (2000), Firm Size and Cyclical Variations in Stock Returns, *Journal of Finance*, 55, 1229–1262.
- Perez-Quiros, G. and A. Timmermann (2001), Business Cycle Asymmetries in Stock Returns: Evidence from Higher Order Moments and Conditional Densities, *Journal of Econometrics*, 103, 259–306.

- Pesaran, M. H. and A. Timmermann (2009), Testing Dependence Among Serially Correlated Multicategory Variables, *Journal of the American Statistical Association*, 104, 325–337.
- Richardson, T. and D. R. Peterson (1999), The Cross-Autocorrelation of Size-Based Portfolios is not an Artifact of Portfolio Autocorrelation, *Journal of Financial Research*, 22, 1–13.
- Ritter, C. and M. A. Tanner (1992), Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy-Gibbs Sampler, *Journal of the American Statistical Association*, 87, 861–868.
- Tanner, M. A. and W. H. Wong (1987), The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association*, 82, 528–550.
- Timmermann, A. (2000), Moments of Markov Switching Models, *Journal of Econometrics*, 96, 75–111.
- Turner, C. M., R. Startz, and C. R. Nelson (1989), A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market, *Journal of Financial Economics*, 25, 3–22.
- Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.
- Zhou, B. (1996), High-Frequency Data and Volatility in Foreign-Exchange Rates, *Journal of Business & Economic Statistics*, 14, 45–52.