Mixture Hidden Markov Models in Finance Research

Chapter · January 2010

DOI: 10.1007/978-3-642-01044-6_41 · Source: DBLP

José G. Dias
ISCTE-Instituto Universitário de Lisboa

Jeroen K. Vermunt
Tilburg University

Sofia B. Ramos
ESSEC - Ecole Supérieur des Sciences Economiques et Commerciales

Mixture Hidden Markov Models in Finance Research

José G. Dias, Jeroen K. Vermunt, and Sofia Ramos

Abstract Finite mixture models have proven to be a powerful framework whenever unobserved heterogeneity cannot be ignored. We introduce in finance research the Mixture Hidden Markov Model (MHMM) that takes into account time and space heterogeneity simultaneously. This approach is flexible in the sense that it can deal with the specific features of financial time series data, such as asymmetry, kurtosis, and unobserved heterogeneity. This methodology is applied to model simultaneously 12 time series of Asian stock markets indexes. Because we selected a heterogeneous sample of countries including both developed and emerging countries, we expect that heterogeneity in market returns due to country idiosyncrasies will show up in the results. The best fitting model was the one with two clusters at country level with different dynamics between the two regimes.

Keywords Finite mixture model · Hidden Markov model · Market volatility · Model-based clustering · Stock indexes.

1 Introduction

Finite mixture modeling has been a powerful tool for capturing unobserved heterogeneity in a wide range of social and behavioral science data (see, for example, McLachlan & Peel, 2000 or Dias & Vermunt, 2007). Modeling the dynamics of stock market returns has been an important challenge in modern financial econometrics. The statistics and dynamics of correctly specified distributions provide more accurate and detailed input for financial asset pricing and risk management. For example, investors buy or sell securities according to their expectation of the market state. In addition, portfolio risk reduction might be achieved by procedures that take into account the synchronization of market regimes. We introduce a specific

José G. Dias (✉)
Department of Quantitative Methods, ISCTE - Lisbon University Institute, Edifício ISCTE, Av. das Forças Armadas, 1649–026 Lisbon, Portugal,
e-mail: jose.dias@iscte.pt

A. Fink et al., (eds.), *Advances in Data Analysis, Data Handling and Business Intelligence*, Studies in Classification, Data Analysis, and Knowledge Organization,
DOI 10.1007/978-3-642-01044-6_41,

finite model for financial time series analysis that takes into account unobserved heterogeneity across space and time. Here, this methodology is used to model the dynamics of the returns of 12 stock market indexes.

As illustrated below, the proposed approach is flexible in the sense that it can deal with the specific features of financial time series data, such as asymmetry, kurtosis and unobserved heterogeneity. Having selected a heterogeneous sample of countries including both developed and emerging countries from Asia, we expect that heterogeneity in market returns due to country specificities will show up in the results. For instance, emerging market return distributions show larger deviations from normality; i.e., are more skewed and have fatter tails (Harvey, 1995).

The paper is organized as follows: Sect. 2 presents the full mixture hidden Markov model; Sect. 3 describes the 12 stock market time series that are used throughout this paper. Section 4 reports MHMM estimates. The paper concludes with a summary of the main findings.

2 The Mixture Hidden Markov Model

We model simultaneously the time series of n stock markets returns. Let y_{it} represent the response of observation (stock market) i at time point t, where $i \in \{1, \ldots, n\}$, $t \in \{1, \ldots, T\}$. In addition to the observed "response" variable y_{it}, the MHMM contains two different latent variables: a time-constant discrete latent variable and a time-varying discrete latent variable. The former, which is denoted by $w \in \{1, \ldots, S\}$ captures the unobserved heterogeneity across stock markets; that is, stock markets are clustered based on differences in their dynamics. We will refer to a model with S clusters as MHMM-S. The two-state time-varying latent variable is denoted by $z_t \in \{1, 2\}$.

Let $f(\mathbf{y}_i; \varphi)$ be the (probability) density function associated with the index return rates of stock market i, where φ is the vector of parameters in the model. The MHMM-S defines the following parametric model for this density:

$$f(\mathbf{y}_i; \varphi) = \sum_{w=1}^{S} \sum_{z_1=1}^{2} \cdots \sum_{z_T=1}^{2} f(w) f(z_1|w) \prod_{t=2}^{T} f(z_t|z_{t-1}, w) \prod_{t=1}^{T} f(y_{it}|z_t). \qquad (1)$$

As in any mixture model, the observed data density $f(\mathbf{y}_i; \varphi)$ is obtained by marginalizing over the latent variables. Because in our model these are discrete variables, this simply involves the computation of a weighted average of class-specific probability densities where the (prior) class membership probabilities or mixture proportions serve as weights (McLachlan & Peel, 2000). We assume that within cluster w the sequence $\{z_1, \ldots, z_T\}$ is in agreement with a first-order Markov chain. Moreover, we assume that the observed return at a particular time point depends only on the regime at this time point; i.e., conditionally on the latent state z_t, the response y_{it} is independent of returns at other time points, which is often referred to as the local independence assumption. As far as the first-order Markov assumption for the latent regime switching conditional on cluster membership w is concerned, it is important to note that this assumption is not as restrictive as one may initially think. It does clearly not imply a first-order Markov structure for the responses y_{it}. The standard hidden Markov model (HMM) (Baum, Petrie, Soules, & Weiss, 1970) is a special case of the MHMM-S that is obtained by eliminating the time-constant latent variable w from the model, that is, by assuming that there is no unobserved heterogeneity across countries.

The characterization of the MHMM is provided by:

- $f(w)$ is the prior probability of belonging to a particular cluster w with multinomial parameter $\pi_w = P(W = w)$.
- $f(z_1|w)$ is the initial-regime probability; that is, the probability of having a particular initial regime conditional on belonging to cluster w with Bernoulli parameter $\lambda_{kw} = P(Z_1 = k|W = w)$.
- $f(z_t|z_{t-1}, w)$ is a latent transition probability; that is, the probability of being in a particular regime at time point t conditional on the regime at time point $t-1$ and cluster membership; assuming a time-homogeneous transition process, we have $p_{jkw} = P(Z_t = k|Z_{t-1} = j, W = w)$ as the relevant Bernoulli parameter. In other words, within cluster w one has the transition probability matrix

$$\mathbf{P}_w = \begin{pmatrix} p_{11w} & p_{12w} \\ p_{21w} & p_{22w} \end{pmatrix},$$

  with $p_{12w} = 1 - p_{11w}$ and $p_{22w} = 1 - p_{21w}$. Note that the MHMM-S allows that each cluster has its specific transition or regime-switching dynamics, whereas in a standard HMM it is assumed that all cases have the same transition probabilities.
- $f(y_{it}|z_t)$, the probability density of having a particular observed stock return in index i at time point t, conditional on the regime occupied at time point t, is assumed to have the form of a univariate normal (or Gaussian) density function. This distribution is characterized by the parameter vector $\theta_k = (\mu_k, \sigma_k^2)$ containing the mean (μ_k) and variance (σ_k^2) for regime k. Note that these parameters are assumed invariant across clusters, an assumption that may, however, be relaxed.

Since $f(\mathbf{y}_i; \varphi)$, defined by (1), is a mixture of densities across clusters w and regimes, it defines a flexible Gaussian mixture model that can accommodate deviations from normality in terms of skewness and kurtosis. The two-state MHMM-S has $4S + 3$ free parameters to be estimated, including $S - 1$ class sizes, S initial-regime probabilities, $2S$ transition probabilities, two conditional means, and two conditional variances.

Maximum likelihood (ML) estimation of the parameters of the MHMM-S involves maximizing the log-likelihood function: $\ell(\varphi; \mathbf{y}) = \sum_{i=1}^{n} \log f(\mathbf{y}_i; \varphi)$, a problem that can be solved by means of the Expectation-Maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977). The E step computes the joint conditional distribution of the $T + 1$ latent variables given the data and the current provisional estimates of the model parameters. In the M step, standard complete data ML methods are used to update the unknown model parameters using an expanded data matrix with the estimated densities of the latent variables as weights. Since the EM algorithm requires us to compute and store the $S \times 2^T$ entries in the E step this makes this algorithm impractical or even impossible to apply with more than a few time points. However, for hidden Markov models, a special variant of the EM algorithm has been proposed that is usually referred to as the forward-backward or Baum–Welch algorithm (Baum et al., 1970). The Baum–Welch algorithm circumvents the computation of this joint posterior distribution making use of the conditional independencies implied by the model. As shown by Vermunt, Tran, and Magidson (2008), the Baum–Welch algorithm for HMMs can easily be generalized to the mixtures of HMMs.

An important modeling issue is the setting of S, the number of clusters needed to capture the unobserved heterogeneity across stock markets. The selection of S is typically based on information statistics such as the Bayesian Information Criterion (BIC) (Schwarz, 1978). In our application we select S that minimizes the BIC value defined as:

$$BIC_S = -2\ell_S(\hat{\varphi}; \mathbf{y}) + N_S \log n, \tag{2}$$

where N_S is the number of free parameters of the model and n is the sample size.

3 Data Set

The data set used in this article are daily closing prices from 4 July 1994 to 27 September 2007 for 12 Asian stock market indexes drawn from Datastream database and listed in Table 1. The series are expressed in US dollars. In total, we have 3,454 end-of-the-day observations per country. Let P_{it} be the observed daily closing price of market i on day t, $i = 1, \ldots, n$ and $t = 0, \ldots, T$. The daily rates of return are defined as the percentage rate of return $y_{it} = 100 \times \log(P_{it}/P_{i,t-1})$, $t = 1, \ldots, T$, with $T = 3{,}454$.

The sample has some appealing features as it mixes developed and emerging markets of the Asian region. Major companies like S&P or MSCI develop regional indices because of the presumption that neighbor countries are economically interrelated. For instance, neighbor countries have more intense trade and, as a result, "cycles" related to one neighbor are likely to affect the other neighbor country. Therefore, one could expect some homogeneity on the behavior of such countries. One the other hand, international stock markets are divided in developed in emerging markets because of distinguished features of both markets. Therefore, the methodology will provide an opportunity to investigate how countries cluster in that region and whether it is indeed the case that neighbor countries have similar regime-switching propensities.

Table 1 provides descriptive statistics of the time series, while Fig. 1 depicts the full time series. The sample period includes periods of market instability as the Asian Flu Crises of 1997, the Russian Crises of 1998, and the global stock market downturn of the 2001 following the dot com bubble. It can be seen that both the

Table 1 Summary statistics

Stock market	Mean	Median	Std. deviation	Skewness	Kurtosis	Jarque–Bera test	
						Statistics	p-Value
Australia (AU)	0.043	0.047	1.066	−0.265	4.011	2,340	0.000
China (CH)	0.049	0.012	1.864	0.007	5.128	3,760	0.000
Hong Kong (HK)	0.033	0.009	1.506	0.001	11.402	18,620	0.000
India (IN)	0.038	0.041	1.556	−0.448	4.756	3,350	0.000
Japan (JP)	−0.003	0.000	1.362	0.109	3.114	1,390	0.000
Malaysia (MY)	0.004	0.000	1.822	−1.565	73.976	785,490	0.000
New Zealand (NZ)	0.027	0.042	1.057	−0.612	9.353	12,740	0.000
Pakistan (PK)	0.009	0.000	1.874	−0.377	6.491	6,110	0.000
Philippines (PH)	0.000	0.000	1.556	0.832	15.513	34,870	0.000
Singapore (SG)	0.020	0.047	1.263	−0.007	7.090	7,200	0.000
Taiwan (TA)	0.010	0.000	1.685	−0.145	3.176	1,450	0.000
Thailand (TH)	−0.011	0.000	2.098	0.332	8.409	10,190	0.000

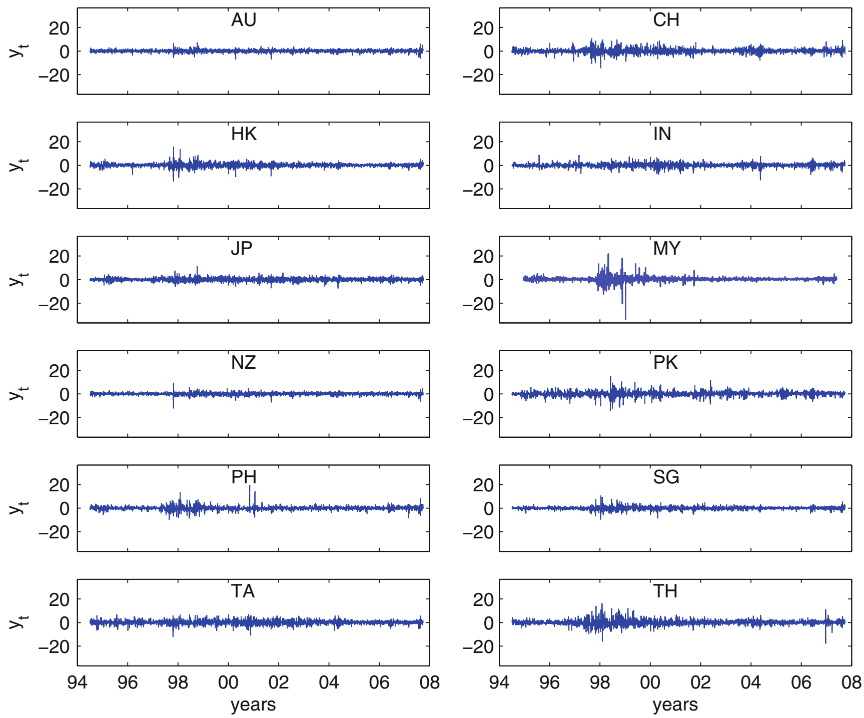

Fig. 1 Time series of index rates for 12 Asian region stock markets

mean and the median return rates are positive and close to zero, except for Japan and Thailand. Stock markets show, instead, very diverse patterns of dispersion, where the largest standard deviations are found in Thailand, China and Malaysia and the smallest dispersion in New Zealand and Australia. Higher standard deviations are typical for emerging markets, known for their high risk. Return rate distributions are diverse in terms of skewness and the kurtosis (which equals 0 for normal distributions) shows high positive values, indicating heavier tails and more peakness than the normal distribution. The Jarque–Bera test rejects the null hypothesis of normality for all 12 stock markets. Overall, these stock market features seem well suited to be modeled using MHMMs.

4 Results

This section reports the results obtained when applying the MHMM-S described before to these 12 stock markets. We estimated models characterized by different number of clusters ($S = 1, \ldots, 8$), using for the estimation of each of them 300 different starting values for the parameters to avoid local maxima. The model with two clusters ($S = 2$) yielded the lowest BIC value ($\ell_2(\hat{\varphi}; \mathbf{y}) = -70,256.1081$, $N_2 = 11$ and $BIC_2 = 140,539.6$).

Table 2 summarizes the results related to the distribution of stock market across clusters which gives the size of each cluster. The prior class membership probability shows that both clusters have the same size. From the posterior class membership probabilities, the probability of belonging to each of the clusters conditional on the observed data (Table 2), we found six countries assigned to cluster 1 (China, India, Japan, Pakistan, Taiwan, and Thailand) and six countries as well assigned to cluster 2 (Australia, Hong Kong, Malaysia, New Zealand, Philippines, and Singapore). Notice that from the posterior probabilities the modal allocation into classes is precise (the probability of the most likely cluster is always one or very close to one). Notice also that cluster 1 has mostly emerging market countries with the exception of the Japan, while cluster 2 is composed mainly by developed countries with the exception of Malaysia and Philippines. By combining the classification

Table 2 Estimated prior and posterior probabilities, and modal clusters for the MHMM-2

Stock market	Cluster 1	Cluster 2	Modal cluster
Prior probabilities	0.501	0.499	
Posterior probabilities			
Australia (AU)	0.000	1.000	2
China (CH)	1.000	0.000	1
Hong Kong (HK)	0.000	1.000	2
India (IN)	1.000	0.000	1
Japan (JP)	0.992	0.008	1
Malaysia (MY)	0.000	1.000	2
New Zealand (NZ)	0.000	1.000	2
Pakistan (PK)	1.000	0.000	1
Philippines (PH)	0.019	0.981	2
Singapore (SG)	0.000	1.000	2
Taiwan (TA)	1.000	0.000	1
Thailand (TH)	1.000	0.000	1

Table 3 Estimated marginal probabilities of the regimes and within Gaussian parameters

	$P(Z)$		Return (mean)		Risk (variance)	
	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
Estimate	0.2545	0.7455	−0.1025	0.0596	7.3521	0.8802
Std. error	(0.0280)	(0.0280)	(0.0276)	(0.0060)	(0.1439)	(0.0116)

Table 4 Characterization of the switching regimes

	Cluster 1		Cluster 2	
	Regime 1	Regime 2	Regime 1	Regime 2
$P(Z\|W)$	0.3487	0.6513	0.1601	0.8399
	(0.0141)	(0.0141)	(0.0135)	(0.0135)
Transitions				
Regime 1	0.9047	0.0953	0.9349	0.0651
	(0.0068)	(0.0068)	(0.0063)	(0.0063)
Regime 2	0.0512	0.9488	0.0124	0.9876
	(0.0035)	(0.0035)	(0.0012)	(0.0012)

information with the descriptive statistics in Table 1, Cluster 1 tends to contain countries with higher volatility (except Japan) and cluster 2 aggregates countries with lower volatility, except mainly Malaysia. As it will become clear the main discrimination between these two groups has to do with other important factors.

Table 3 provides information on the two regimes that were identified; that is, the average proportion of markets in regime k over time and the mean and variance of the returns in regime k. The result is in line with the common dichotomization of financial markets into "bull" and "bear" markets. Consistently, the reported means show that one of the regimes is associated with positive returns (bull market) and the other with negative returns (bear market). The probability of being in the bear and bull regimes is 0.25 and 0.75, respectively. We would also like to emphasize that these results are coherent with the common acknowledgment of volatility asymmetry of financial markets. Volatility is likely to be higher when markets fall than when markets rise.

Table 4 reports the estimated probabilities of being in one of the regimes within each cluster. There is a clear distinction between these clusters. Cluster 1 has the largest probability of being in bear regime (0.35). For cluster 2 this probability becomes 0.16. Moreover, Table 4 provides another key insight from our analysis. It gives the transition probabilities between the two regimes for both clusters. First, notice that both clusters show regime persistence. Once a stock market jumps to a regime, it is likely to remain within the same regime for a while, which is coherent with stylized facts in financial markets. Second, cluster 2 shows lower propensity to move from a bull regime to a bear regime (0.012) than cluster 1. Third, cluster 1 shows higher probability to jump from a bear to a bull regime than cluster 2. This is in line with the idea that cluster 1 has more emerging markets, which are known for having more and longer financial crises than developed markets.

Figure 2 shows the regime-switching dynamics of the countries within both clusters. It depicts the posterior probability of being in bull regime at period t, where

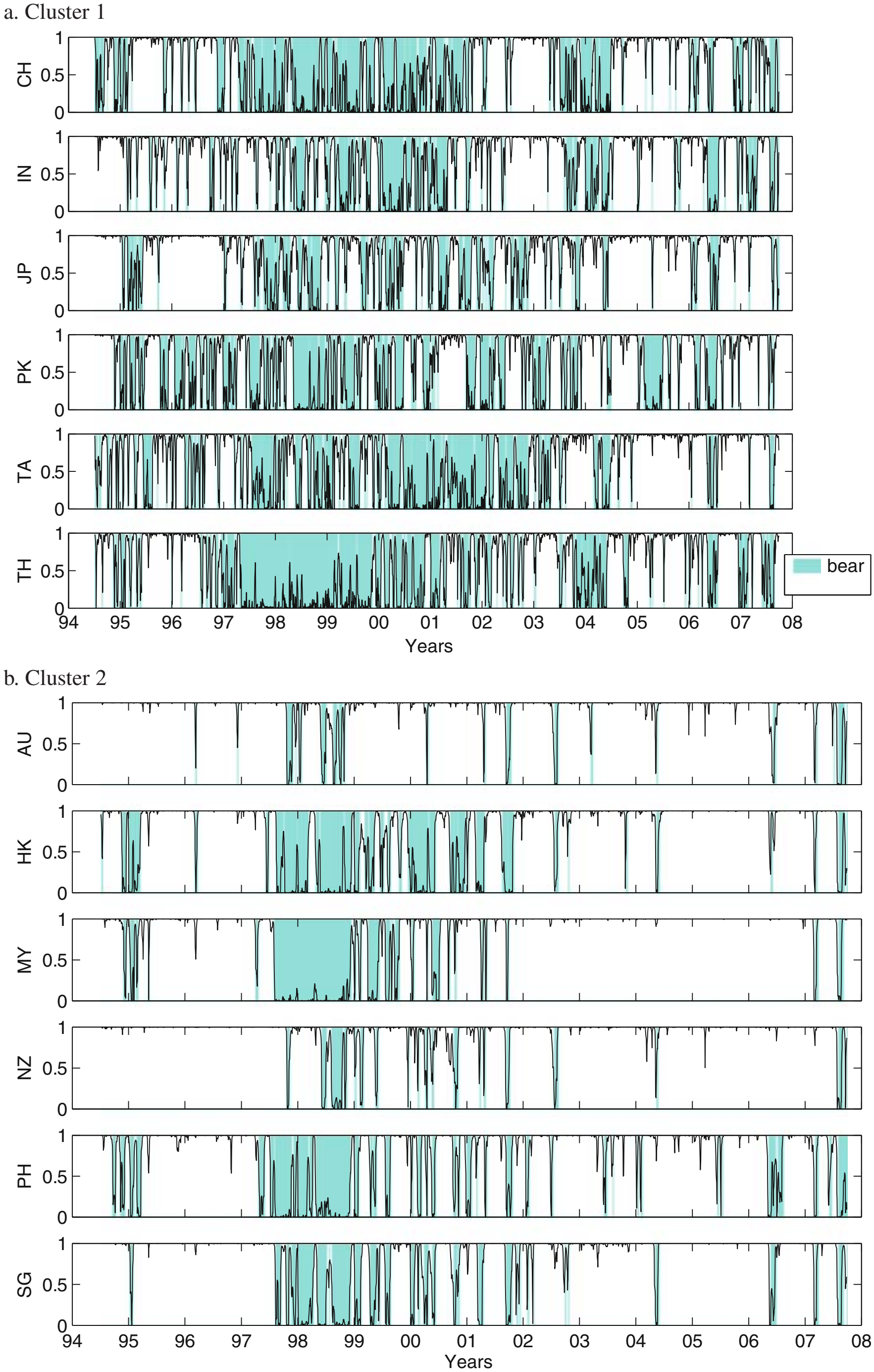

Fig. 2 Estimated posterior bull regime probability and modal regime

the grey color identifies periods in which this probability is below 0.5 which corresponds to a higher likelihood of being in the bear regime. It is visible a long period of "bear regimes" that starts at the end of 1997, with the Thailand's currency crisis and goes until 2002 that affects all the countries of the region. However, the behavior before 1997 and after 2002 is clearly different between countries from cluster 1 and 2. The two clusters of countries have rather different pattern of regime switching. Cluster 2 is more regime persistent with short duration bear regimes that did not turn out to be endemic during the period of analysis, despite critical periods around 1998. Cluster 1 is extremely dynamic and tends to move very fast between regimes, switching frequently between bear and bull states.

5 Conclusions

A mixture of hidden Markov models allows model-based clustering of financial time series. In the analysis of a sample of 12 stock markets providing observations for a period of 3,454 days the best fitting model was the one with two clusters. The two clusters clearly defined two distinct types of regime switching, which is coherent with many stylized facts in finance. Moreover, the simultaneous analysis of the 12 time series allows a better comparison of country dynamics in opposition to the application of Markov-switching approaches that estimate regimes for each country separatively (see, e.g., Wang & Theobald, 2008).

References

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

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

Dias, J. G., & Vermunt, J. K. (2007). Latent class modeling of website users' search patterns: Implications for online market segmentation. *Journal of Retailing and Consumer Services, 14*, 359–368.

Harvey, C. R. (1995). Predictable risk and returns in emerging markets. *Review of Financial Studies, 8*, 773–816.

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Vermunt, J. K., Tran, B., & Magidson, J. (2008). Latent class models in longitudinal research. In S. Menard (Ed.), *Handbook of longitudinal research: Design, measurement, and analysis* (pp. 373–385). Burlington, MA: Elsevier.

Wang, P., & Theobald, M. (2008). Regime-switching volatility of six East Asian emerging markets. *Research in International Business and Finance, 22*(3), 267–283.