行政院國家科學委員會補助專題研究計畫成果報告

多屬性多方法潛伏成長模型及其實證應用

Latent Growth Modeling with MTMM and Its Empirical Applications

計畫類別：個別型計畫

計畫編號：NSC 98-2410-H-009-010-MY2

執行期間：98 年 8 月 1 日至 100 年 10 月 31 日

執行機構及系所：國立交通大學經營管理研究所

計畫主持人：丁　承

計畫參與人員：鄭天德、王宏瑞、李孟哲、洪文琪

成果報告類型(依經費核定清單規定繳交)：完整報告

本計畫除繳交成果報告外，另已繳交出席國際學術會議心得報告。

處理方式：得立即公開查詢

中　華　民　國　100　年　10　月　28　日

目　　錄

中英文摘要及關鍵詞

摘　要

本研究計畫分兩部分，第一部分為潛伏成長模型(LGM)中level-1誤差共變異結構之鑑定，我們提出一有效的誤差共變異結構之鑑定方法，該法係基於卡方差異檢定以及自我相關與偏自我相關檢定，所鑑定結果合乎模型配適與精簡之訴求，並有效降低模型誤設的可能性，另亦具體說明並示範如何利用SAS PROC CALIS進行level-1誤差共變異結構之配適，在操作技術上具應用價值，有益於潛伏成長模型之實證研究；第二部分利用LGM從事投資人情緒實證研究，我們引進潛伏成長模型分析證券市場投資人情緒變化趨勢，並提出基於個股資料量測投資人情緒水平與情緒趨勢的方法，包括評估指標之建立以及信度與效度分析，投資人情緒之分析有利於掌握證券市場干預時機，在市場管理上具參考價值。

關鍵詞：潛伏成長模型，誤差共變異結構，卡方差異檢定，自我相關，偏自我相關，穩態，模型誤設，投資人情緒，周轉率，流動性，首次公開發行，驗證性因素分析，信度，收斂效度，區別效度，套利機制，市場干預。

Abstract

This study consists of two parts. The first part is to propose an effective approach to identify the level-1 error covariance structure of a latent growth model (LGM). The approach is based on the chi-square difference test and the tests for autocorrelations and partial autocorrelations. The error covariance structure identified is as simple as possible under the condition of achieving model fit. The possibility of model misspecification can be reduced. In addition, a tutorial on using SAS PROC CALIS to fit error covariance structures of latent growth models has been provided. It is useful for empirical studies using LGM. The second part of this study is to conduct an empirical research for investor sentiment by using LGM. Moreover, how to measure investment sentiment is specifically addressed, including the establishment of appropriate indicators and the assessment of reliability and validity. Analyses of investment sentiment are informative for policy makers to capture the timing to conduct intervention to stabilize securities markets.

Keywords**:** latent growth modeling (LGM), error covariance structure, chi-square difference test, autocorrelation, partial autocorrelation, stationarity, model misspecification, investor sentiment, turnover ratio, liquidity, initial public offering (IPO), confirmatory factor analysis (CFA), reliability, convergent validity, discriminant validity, arbitrage mechanism, market intervention.

1. Using SAS PROC CALIS to Fit Level-1 Error Covariance Structures of Latent Growth Models

1.1 Introduction

The latent growth model (LGM) plays an important role in repeated-measure analysis over a limited occasion in large sample data (e.g., Meredith & Tisak, 1990; Muthén & Khoo, 1998; Preacher, Wichman, MacCallum, & Briggs, 2008, p. 12; Singer & Willett, 2003, p. 9). The model can not only characterize intraindividual (within-subject) change over time but also examine interindividual (between-subject) difference by means of random growth coefficients, and is a typical application of hierarchical linear modeling (HLM). The within-subject errors over time and the between-subject errors are conventionally referred to as level-1 and level-2 errors, respectively. LGM can also be handled by using structural equation modeling (SEM) (e.g., Bauer, 2003; Boolen & Curran, 2006; Chan, 1998; Curran, 2003; Duncan, Duncan, & Hops, 1996; Mehta & Neal 2005; Meredith & Tisak, 1990; Willet & Sayer, 1994). SEM and HLM stem from different statistical theory, and each has developed its own terminology and standard ways of framing research questions. However, there exists much overlap between the two methodologies under some circumstances. Typically, when a two-level data structure arises from the repeated observations of a variable over time for a set of individuals (such that time is hierarchically nested within each individual), SEM is analytically equivalent to HLM (e.g., Bauer, 2003; Bovaird, 2007; Curran, 2003; MacCallum, Kim, Malarkey, & Kiecolt-Glaser, 1997; Raudenbush, 2001; Rovine & Molenaar, 2000; Willett & Sayer, 1994). The SEM approach provides advantages over the HLM approach in examining model fit, modeling the change over time for latent constructs, with the curve-of-factors model, embeding LGM into a larger latent variable model, with the factor-of-curves model, and incorporating measurement models for latent predictors (e.g., Bauer, 2003; Bollen & Curran, 2006, Chap. 7, 8; Bovaird, 2007; Chan, 1998; Curran, 2003; Duncan, Duncan, & Strycker, 2006, Chap. 4; MacCallum, et al., 1997; Raudenbush, 2001; Rovine & Molenaar, 2000). However, the SEM approach suffers from a tedious and error-prone data management task. Many steps are needed to properly structure the data and the SEM code quickly becomes unwieldy. In contrast, the HLM approach allows for simpler model specification, is computationally more efficient, and can easily be expanded to higher-level growth models for manifest variables (Curran, 2003; Wu, West, & Taylor, 2009). Detailed comparison between HLM and SEM can be seen in Bauer (2003) and Curran (2003).

Specialized software for SEM such as EQS (Bentler & Wu, 2005), LISREL (Jöreskog & Sörbom, 2001), M*plus* (Muthén & Muthén, 2007), Mx (Neale, Boker, Xie, & Maes, 2003), and SAS PROC CALIS (SAS Institute Inc., 2010) are readily available. HLM (Raudenbush, Bryk, & Congdon, 2005), MLwiN (Rasbash et al., 2000), and SAS PROC MIXED (SAS Institute Inc., 2010) are typical software for HLM. Due to the isomorphism between SEM and HLM for the same growth model, parameter estimates with SEM and those with HLM should be equivalent. Any minor variations can be attributed to different computational methods used (standard maximum likelihood (ML) estimation or full information maximum likelihood (FIML) estimation for SEM and restricted maximum likelihood estimation for HLM). Relevant discussions have been given in Bauer (2003), Bovaird (2007), Curran (2003), and Mehta and Neale (2005).

Level-1 errors could be autocorrelated. Autocorrelations, considered to be nuisance parameters, might result from carryover effects, memory effects, practice effects, or other unmodeled associations, and might not be present when a more complex model or a more appropriate time structure is used (Grimm & Widaman, 2010; Sivo & Fan, 2008). For example, the growth curve ARMA(p, q) model has been proposed to absorb error autocorrelations (e.g., Sivo, Fan, & Witta, 2005; Sivo & Fan, 2008). When level-1 errors are autocorrelated, misspecification of their covariance structure has a substantial impact on the inference for model parameters (Ferron, Dailey, & Yi, 2002; Kwok, West, & Green, 2007; Murphy & Pituch, 2009). However, correct covariance structure is difficult to specify by theory (Kwok et al., 2007, p. 588). Therefore, a specification search becomes needed. Littell et al. (2006, Chap. 5) illustrated two types of tools with SAS PROC MIXED to help select a covariance structure. First are graphical tools to visualize correlation patterns among residuals. Second are information criteria measuring the relative fit of competing covariance structures. AIC (Akaike, 1974) and BIC (Schwarz, 1978) are commonly used descriptive measures. The model that minimizes AIC or BIC is preferred. Before using these methods, researchers should first rule out covariance structures that are obviously inconsistent with the characteristics of the data. On the other hand, although linear growth curve models are often fitted because of their ease in estimation, theory may suggest that more complex growth models be used, as they can better capture developmental patterns. Correctly specifying the growth model might lead to a simple covariance structure (Grimm & Widaman, 2010). Moreover, when the growth model is misspecified, statistical inference during the search process can be misleading (Yuan & Bentler, 2004). Therefore, the growth model should be well determined before searching for an “optimal” covariance structure for level-1 errors.

A variety of processes underlying level-1 errors may be specified (e.g., Newsom, 2002; Singer & Willett, 2003, Chap. 7; Wolfinger, 1996). SAS PROC MIXED contains more than 30 different types of level-1 preprogrammed error processes. However, some important processes are unavailable and any modification of existing processes is not allowed. In contrast, there exists much flexibility in PROC CALIS when specifying error covariance structures. For example, the second-order autoregressive process, not available in PROC MIXED, can be handled with PROC CALIS. The strength of PROC CALIS is always accompanied with technical coding work, which needs to be specifically addressed, and is the focus of this study. In addition to PROC CALIS, any comparable SEM software could be used.

There seems to be no commonly acceptable criteria for assessing model fit based on the indices such as AIC and BIC resulting from PROC MIXED. In contrast, there is some agreement on the cutoff criteria of conventional fit indices based on the likelihood ratio test in SEM such as RMSEA, CFI, and NNFI (TLI) (e.g., Hu and Bentler, 1999). However, since in SEM-based LGM, the factor loadings are usually fixed at time points rather than freely estimated, and the fit of the model to the mean structure should be reflected as well, assessment of model fit by using conventional SEM-based fit indices should be cautious (Mehta & Neale, 2005; Wu, West, & Taylor, 2009). When every individual is observed at the same fixed set of time points (called balanced) with no missing values (called complete), ML estimation is used; otherwise FIML estimation is used (Wu, West, & Taylor, 2009). With FIML estimation, the model-implied means and covariances are computed for each individual, and the maximum likelihood chi-square fit function is obtained by summing −2 log likelihood across all of the individual data vectors (Bovaird, 2007). For balanced and complete data, FIML simplifies to ML, and, in this case, RMSEA, CFI, and NNFI among the SEM-based fit indices have shown good potential performance in evaluating the fit of LGM (Wu & West, 2010; Wu, West, & Taylor, 2009). For unbalanced designs or missing data, conventional guidelines for adequate fit with these indices may be misleading (Wu, West, & Taylor, 2009).

During the search process, we need instruments for the implementation of fitting various types of error covariance structures. The primary motivation to use PROC CALIS is to take advantage of its flexibility in specifying level-1 error covariance structures and its capability to deal with growth modeling for both manifest variables and latent constructs. PROC CALIS performs better than PROC MIXED, but more sophisticated coding work is required. The purpose of this study is to address this issue by giving a tutorial on the syntax using PROC CALIS to fit many types of level-1 error covariance structures in LGM for a manifest variable as well as for a latent construct. Illustrations will be conducted with the data generated from two given latent growth models. SAS is a general-purpose and publicly available software. Its ability to do data management and analysis within a single package would make the instruments we provide attractive to many researchers.

1.2 Latent Growth Models

In this section, we briefly introduce the LGM with a variety of level-1 error covariance structures through a typical example depicted in Figure 1-1. In the figure, $y_1 - y_4$ denote the repeated measures of y on four occasions and X a level-2 predictor. η_{α_i} is the unobserved intercept representing the initial status for individual i, and η_{β_i} the unobserved slope showing the individual's linear rate of change per unit increase in time. η_{α_i} and η_{β_i} are both latent factors. The level-1 model can be written as

$$\mathbf{y} = \mathbf{\Lambda}_{\mathrm{y}}^{*}\boldsymbol{\eta} + \boldsymbol{\varepsilon}, \tag{1-1}$$

where $\boldsymbol{y} = [y_1\ y_2\ y_3\ y_4]'$, $\mathbf{\Lambda}_{\mathrm{y}}^{*\prime} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{bmatrix}$, $\boldsymbol{\eta} = [\eta_\alpha\ \eta_\beta]'$, and $\boldsymbol{\varepsilon} = [\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \varepsilon_4]'$. λ_t is the measurement time points (t = 1, 2, 3, 4) and $\boldsymbol{\varepsilon}$ denotes level-1 errors. The solid line with four arrowheads presented in Figure 1-1 indicate that ε_t are pairwise correlated. The factor loading associated with initial status are all fixed at 1, whereas those associated with the slope are set at the value λ_t to reflect the particular time point t for individual i . A common coding of λ_t for different time points is to set λ_1 = 0 for baseline and $\lambda_t = t-1$ for the follow-ups. For this model, subject i's growth trajectory is a straight line, $\eta_{\alpha_i} + \lambda_t\eta_{\beta_i}$, λ_t = 0, 1, 2, 3. (For simplicity, subscript i is omitted for the rest part of this section.) The loading matrix $\mathbf{\Lambda}_{\mathrm{y}}^{*}$ containing fixed values has a superscript * to distinguish from the traditional notation used for the unknown loadings in confirmatory factor analysis (CFA). The model is a restricted CFA model.

The level-2 model can be written as

$$\boldsymbol{\eta} = \mathbf{\Gamma}_0 + \mathbf{\Gamma}_{\mathrm{x}}\mathbf{x} + \boldsymbol{\zeta}_{\boldsymbol{\eta}}, \tag{1-2}$$

where $\mathbf{\Gamma}_0 = [\gamma_{00}\ \gamma_{01}]'$, $\mathbf{\Gamma}_{\mathrm{x}} = [\gamma_{10}\ \gamma_{11}]'$, $\mathbf{x} = [X]$, and $\boldsymbol{\zeta}_{\boldsymbol{\eta}} = [\zeta_{\eta_\alpha}\ \zeta_{\eta_\beta}]$. Growth factors η_α and η_β (a random intercept and a random slope) are both predicted by a time invariant subject-level covariate X. γ_{00} and γ_{10} denote, respectively, the intercept and slope of the regression of η_α on X, γ_{01} and γ_{11} are those of η_β on X, and ζ_{η_α} and ζ_{η_β} are level-2 errors. Two or more time invariant predictors of change may be included. Since it is not our focus, for simplicity, we consider only one predictor here. $\boldsymbol{\zeta}_{\boldsymbol{\eta}}$ and $\boldsymbol{\varepsilon}$ are assumed to be uncorrelated. The models can be rewritten in combined form as

$$\boldsymbol{y} = \mathbf{\Lambda}_{y}^{*}(\mathbf{\Gamma}_0 + \mathbf{\Gamma}_{\mathrm{x}}\mathbf{x}) + \mathbf{\Lambda}_{y}^{*}\boldsymbol{\zeta}_{\eta} + \boldsymbol{\varepsilon}, \tag{1-3}$$

based on which the model-implied mean vector $\boldsymbol{\mu}$ and the model-implied covariance matrix $\boldsymbol{\Sigma}$ of the manifest variables y_1–y_4 and X can be expressed as functions of the model parameters as follows (Bollen & Curran, 2006, p. 134-135):

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{y}} \\ \boldsymbol{\mu}_{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}_{\mathbf{y}}^{*}(\boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_{\mathbf{x}}\boldsymbol{\mu}_{\mathbf{x}}) \\ \boldsymbol{\mu}_{\mathbf{x}} \end{bmatrix}, \tag{1-4}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Lambda}_{\mathbf{y}}^{*}(\boldsymbol{\Gamma}_{\mathbf{x}}\boldsymbol{\Sigma}_{\mathbf{xx}}\boldsymbol{\Gamma}_{\mathbf{x}}' + \boldsymbol{\Psi}_{\zeta_\eta})\boldsymbol{\Lambda}_{\mathbf{y}}^{*\prime} + \boldsymbol{\Theta}_{\varepsilon} & \boldsymbol{\Lambda}_{\mathbf{y}}^{*}\boldsymbol{\Gamma}_{\mathbf{x}}\boldsymbol{\Sigma}_{\mathbf{xx}} \\ \boldsymbol{\Sigma}_{\mathbf{xx}}\boldsymbol{\Gamma}_{\mathbf{x}}'\boldsymbol{\Lambda}_{\mathbf{y}}^{*\prime} & \boldsymbol{\Sigma}_{\mathbf{xx}} \end{bmatrix}, \tag{1-5}$$

where $\boldsymbol{\Theta}_{\varepsilon}$ and $\boldsymbol{\Psi}_{\zeta_\eta}$ denote the variance-covariance matrices of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\zeta_\eta}$, respectively, and $\boldsymbol{\mu}_{\mathbf{x}}$ and $\boldsymbol{\Sigma}_{\mathbf{xx}}$ denote, respectively the mean vector and the variance-covariance matrix of predictors ($\boldsymbol{\mu}_{\mathbf{x}} = \mu_X$ and $\boldsymbol{\Sigma}_{\mathbf{xx}} = \sigma_X^2$ for this model since there is only one predictor).

The level-1 errors, ε_1, ε_2, ε_3, and ε_4, are assumed to be normally distributed with zero means. The general error covariance matrix (ECM) is unstructured, and is given by

$$\boldsymbol{\Theta}_{\varepsilon} = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ \sigma_{\varepsilon_2\varepsilon_1} & \sigma_{\varepsilon_2}^2 & & \\ \sigma_{\varepsilon_3\varepsilon_1} & \sigma_{\varepsilon_3\varepsilon_2} & \sigma_{\varepsilon_3}^2 & \\ \sigma_{\varepsilon_4\varepsilon_1} & \sigma_{\varepsilon_4\varepsilon_2} & \sigma_{\varepsilon_4\varepsilon_3} & \sigma_{\varepsilon_4}^2 \end{bmatrix}. \tag{1-6}$$

The corresponding option given in SAS PROC MIXED is TYPE=UN. Other types of ECM, with fewer parameters may be desirable. The level-2 errors ζ_{η_α} and ζ_{η_β} are assumed to be normally distributed with zero means. Their covariance matrix is usually specified as unstructured (Murphy & Pituch, 2009):

$$\boldsymbol{\Psi}_{\zeta_\eta} = \begin{bmatrix} \sigma_{\zeta_{\eta_\alpha}}^2 & \sigma_{\zeta_{\eta_\alpha}\zeta_{\eta_\beta}} \\ \sigma_{\zeta_{\eta_\alpha}\zeta_{\eta_\beta}} & \sigma_{\zeta_{\eta_\beta}}^2 \end{bmatrix}. \tag{1-7}$$

1.2.1 Types of the level-1 error covariance structure and SAS statements

Any type of the level-1 ECM ($\boldsymbol{\Theta}_{\varepsilon}$) can be expressed as a set of linear and/or nonlinear constraints on the parameters involving the covariance structure. SAS PROC MIXED provides a REPEATED statement, in which many types of the level-1 error covariance structure can be specified through the TYPE= option (e.g., Singer, 1998). However, some important processes such as higher-order autoregressive and moving average ones are not included. Moreover, PROC MIXED cannot handle LGM for constructs.[1] To improve, use PROC CALIS. The STD, COV, and PARAMETERS statements in PROC CALIS can be used together to specify any type of ECM. The STD statement defines variances to estimate for

[1] Although PROC NLMIXED could be used to fit linear or nonlinear LGM for constructs (e.g., Blozis, 2006), no option is available in the procedure for specifying types of ECM. Relevant coding is laborious.

exogenous and error variables. The COV statement defines covariances to estimate for exogenous and error variables. The PARAMETERS statement defines additional parameters that are not specified in the models, and uses both the original and additional parameters for modeling ECM. In other words, each specific type of ECM is composed of functions of the original and additional parameters. The SAS statements in PROC CALIS for fitting different types of the level-1 error covariance structures, including AR(1) (the first-order autoregressive), MA(1) (the first-order moving average), ARMA(1,1) (the first-order autoregressive moving average), AR(2) (the second-order autoregressive), MA(2) (the second-order moving average), ARH(1) (heterogeneous AR(1)), TOEPH (heterogeneous Toeplitz), and UN (unstructured), with four equally spaced occasions are summarized in Table 1-1. AR(1), MA(1), ARMA(1,1), AR(2), and MA(2) are members of the ARMA family. Documentation for LGM with ARMA(1,1), TOEPH, and AR(2) for level-1 errors is given as follows:

*Example 1: ARMA(1,1)***.** The ARMA(1,1) process is defined as $\varepsilon_t = \phi_1\varepsilon_{t-1} + v_t - \theta_1 v_{t-1}$, where ϕ_1 denotes the autoregressive parameter, θ_1 the moving average parameter, and v_t an i.i.d. disturbance process (Box, Jenkins, & Reinsel, 1994, p. 77). Its interpretation is that the level-1 error at time *t* can be predicted by the level-1 error at time *t*–1 and the independent disturbance at time *t*–1. The resulting ECM is given by

$$\mathbf{\Theta}_{\varepsilon} = \sigma_{\varepsilon}^2 \begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ \rho_2 & \rho_1 & 1 & \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}, \tag{1-8}$$

where σ_{ε}^2 denotes the common variance of ε_t, t = 1, 2, 3, 4, and ρ_k denotes their autocorrelation coefficient at lag *k*, given by $\rho_1 = \dfrac{(\phi_1 - \theta_1)(1-\phi_1\theta_1)}{(1-2\phi_1\theta_1 + \theta_1^2)}$, $\rho_k = \phi_1\rho_{k-1}$, k = 2, 3, with the constraints of $|\phi_1| < 1$ and $|\theta_1| < 1$. Program 1 in Appendix 1-A demonstrates how to use PROC CALIS for modeling LGM with the ARMA(1,1) covariance structure for level-1 errors and the unstructured covariance for level-2 errors for four equally spaced time points. The UCOV and AUG options are specified to analyze the mean structures in an uncorrected covariance matrix. The dataset to be analyzed is augmented by an intercept variable INTERCEPT that has constant values equal to 1. The LINEQS statement given below is used to specify the level-1 model (the restricted CFA model) shown in Equation 1-1 and the level-2 model shown in Equation 1-2.

```
LINEQS
    Y1 = 1 F_Alpha + 0 F_Beta + E1,
    Y2 = 1 F_Alpha + 1 F_Beta + E2,
    Y3 = 1 F_Alpha + 2 F_Beta + E3,
    Y4 = 1 F_Alpha + 3 F_Beta + E4,
    F_Alpha = GA00 INTERCEPT + GA01 X + D0,
    F_Beta = GA10 INTERCEPT + GA11 X + D1;
```

where F_ALPHA and F_BETA represent latent factors η_{α_i} and η_{β_i}. Factor loadings are fixed values (in $\mathbf{\Lambda}_y^*$). Level-1 errors $\varepsilon_1 - \varepsilon_4$ are named E1–E4, and level-2 errors ζ_{η_α} and ζ_{η_β} are named D0 and D1. GA00, GA01, GA10, and GA11 represent estimates of growth parameters γ_{00}, γ_{01}, γ_{10}, and γ_{11}.

By Equation 1-8, level-1 error variances are equal, their autocovariances at lag 1 are equal, and their autocovariances at lag 2 are equal as well. Level-2 error variances/covariances are unstructured, as shown in Equation 1-7. Therefore, the STD and COV statements are given as follows:

```
STD
    E1=VARE, E2=VARE, E3=VARE, E4=VARE, D0=VARD0, D1=VARD1;
COV
    E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1,
    E1 E3=COV_lag2, E2 E4=COV_lag2,
    E1 E4=COV_lag3,
    D0 D1=COVD0D1;
```

in which VARE represents the estimate of the common variance σ_ε^2 of the four level-1 errors, and VARD0 and VARD1 the estimates of the variances, $\sigma_{\zeta_{\eta_\alpha}}^2$ and $\sigma_{\zeta_{\eta_\beta}}^2$, of the two level-2 errors. COV_lag1 and COV_lag2 represent, respectively, the common level-1 error autocovariance estimates at lag 1 and lag 2. COV_lag3 is the estimate of the error autocovariance at lag 3. CD0D1 is the estimate of $\sigma_{\zeta_{\eta_\alpha}\zeta_{\eta_\beta}}$, the covariance of ζ_{η_α} and ζ_{η_β}.

Since there exist extra parameters in ECM, they need to be defined and the work can be achieved by using the PARAMETERS statement given by

```
PARAMETERS
    PHI1 RHO1;
    COV_lag1=RHO1*VARE;
    COV_lag2=PHI1* COV_lag1;    /* i.e., COV_lag2=PHI1*RHO1* VARE; */
    COV_lag3=PHI1* COV_lag2;    /* i.e., COV_lag3=(PHI1**2)*RHO1*VARE; */
```

in which PHI1 and RHO1 represent the estimates of ρ_1 and ϕ_1, defined through their relationships with the autocovariances shown in Equation 1-8. 'COV_lag1=RHO1*VARE' corresponds to the requirement that the common autocovariance at lag 1 be equal to $\sigma_\varepsilon^2 \rho_1$. The syntax corresponding to the requirements for the autocovariances at lag 2 ($=\sigma_\varepsilon^2 \phi_1 \rho_1$) and lag 3 ($=\sigma_\varepsilon^2 \phi_1 \rho_2 = \sigma_\varepsilon^2 \phi_1^2 \rho_1$) is given in a similar way.

The constraint of $|\phi_1| < 1$ is specified by the following BOUNDS statement:

```
BOUNDS
    –1. < PHI1 < 1.;
```

Example 2: TOEPH. The ECM resulting from heterogeneous Toeplitz is given by

$$\mathbf{\Theta}_{\varepsilon}=\begin{bmatrix}\sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho_1 & \sigma^2_{\varepsilon_2} & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho_2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho_1 & \sigma^2_{\varepsilon_3} & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho_3 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho_2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho_1 & \sigma^2_{\varepsilon_4}\end{bmatrix}, \tag{1-9}$$

where σ_{ε_t} denotes the standard deviation for ε_t, t = 1, 2, 3, 4, and ρ_k the autocorrelation at lag k, k = 1, 2, 3. The level-1 error variances are unequal but the autocorrelations at the same lag are equal. The STD and COV statements are given as follows:

```
STD
    E1=VARE1, E2=VARE2, E3=VARE3, E4=VARE4,
    D0=VARD0, D1=VARD1;
COV
    E1 E2=COVE1E2, E1 E3=COVE1E3, E1 E4=COVE1E4,
    E2 E3=COVE2E3, E2 E4=COVE2E4, E3 E4=COVE3E4,
    D0 D1=COVD0D1;
```

in which VARE1–VARE4 represent the estimates of the four level-1 error variances, and VARD0 and VARD1 those of the two level-2 error variances. COVE1E2–COVE3E4 represent the corresponding level-1 error autocovariance estimates, and COVD0D1 the level-2 error autocovariance estimate. Since the error covariances $\sigma_{\varepsilon_t\varepsilon_{t'}}$ of ε_t and $\varepsilon_{t'}$ is given by $\sigma_{\varepsilon_t\varepsilon_{t'}}=\sigma_{\varepsilon_t}\sigma_{\varepsilon_{t'}}\rho_{\varepsilon_t\varepsilon_{t'}}$ and the autocorrelations at the same lag are constrained to be equal, the following PARAMETERS statement needs to be added:

```
PARAMETERS
    RHO1 RHO2 RHO3;
    COVE1E2=SQRT(VARE1)*SQRT(VARE2)*RHO1;
    COVE2E3=SQRT(VARE2)*SQRT(VARE3)*RHO1;
    COVE3E4=SQRT(VARE3)*SQRT(VARE4)*RHO1;
    COVE1E3=SQRT(VARE1)*SQRT(VARE3)*RHO2;
    COVE2E4=SQRT(VARE2)*SQRT(VARE4)*RHO2;
    COVE1E4=SQRT(VARE1)*SQRT(VARE4)*RHO3;
```

where RHO1, RHO2, and RHO3 are estimates of ρ_1, ρ_2, and ρ_3. The LINEQS statement used for this example is the same as that given in Example 1.

Example 3: AR(2). It is not possible to model AR(2) for level-1 errors by using PROC MIXED, but the task can be done by using PROC CALIS, with the statements shown in Table 1-1. The AR(2) process, given by $\varepsilon_t=\phi_1\varepsilon_{t-1}+\phi_2\varepsilon_{t-2}+v_t$, where ϕ_1 and ϕ_2 are autoregressive parameters and v_t an i.i.d. process (Box, Jenkins, & Reinsel, 1994, p. 54), leads to the following level-1 ECM:

$$\sigma_\varepsilon^2 \begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ \rho_2 & \rho_1 & 1 & \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}, \tag{1-10}$$

where σ_ε^2 denotes the common variance of ε_t, t = 1, 2, 3, 4, and ρ_k denotes their autocorrelation at lag k, given by $\rho_0 = 1$, $\rho_1 = \phi_1 / (1 - \phi_2)$, and $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$, $k = 2,\ 3$, with the constraints of $|\phi_2| < 1$, $\phi_2 + \phi_1 < 1$, and $\phi_2 - \phi_1 < 1$. It follows that the autocovariances at lag 1, 2, and 3, denoted respectively by σ_1, σ_2, and σ_3, are given by $\sigma_1 = \rho_1\sigma_\varepsilon^2$, $\sigma_2 = \rho_2\sigma_\varepsilon^2 = \phi_1\rho_1\sigma_\varepsilon^2 + \phi_2\sigma_\varepsilon^2 = \phi_1\sigma_1 + \phi_2\sigma_\varepsilon^2$, and $\sigma_3 = \rho_3\sigma_\varepsilon^2 = \phi_1\rho_2\sigma_\varepsilon^2 + \phi_2\rho_1\sigma_\varepsilon^2 = \phi_1\sigma_2$ $\phi_2\sigma_1$. Note that the last two constraints are specified by using the LINCON statement. Relevant SAS statements are given as follows:

```
STD
    E1-E4=4*VARE, /* i.e., E1=VARE, E2=VARE, E3=VARE, E4=VARE */
    D0=VARD0, D1=VARD1;
COV
    E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1,
    E1 E3=COV_lag2, E2 E4=COV_lag2, E1 E4=COV_lag3,
    D0 D1=CD0D1;
PARAMETERS PHI1 PHI2;
    RHO1= PHI1/(1–PHI2);
    COV_lag1=RHO1*VARE;
    COV_lag2=PHI1*COV_lag1+ PHI2 *VARE;
    COV_lag3=PHI1*COV_lag2+PHI2*COV_lag1;
LINCON
    PHI2 + PHI1 < 1., PHI2 –PHI1 < 1.;
BOUNDS
    –1. < PHI2 < 1.;
```

In addition to those presented in Table 1-1, more level-1 error covariance structures for equally spaced data, including ARMA(p, q) (autoregressive moving average of order (p, q)), CS (compound symmetry), TOEP(q) (Toeplitz with q bands, q = 1,…, 4, in which the first q bands of the matrix are to be estimated, setting all higher bands equal to zero), CSH (heterogeneous CS), TOEPH(q) (heterogeneous Toeplitz with q bands, q = 1, …, 4), and UN(q) (UN with q bands, q = 1, …, 4), are summarized in Appendix 1-B. In particular, TOEP(1) indicates i.i.d. level-1 errors. SAS statements in PROC CALIS for each of them can be obtained in a similar way as shown in Table 1-1.

The level-1 error covariance structures displayed in Table 1-1 and Appendix 1-B are frequently seen in the LGM literature (e.g., Beck & Katz, 1995; Blozis, Harring, & Mels, 2008; Dawson, Gennings, & Carter, 1997; Eyduran & Akbas, 2010; Ferron et al., 2002; Goldstein, Healy, & Rasbash, 1994; Heitjan & Sharma, 1997; Keselman, Algina, Kowalchuk, & Wolfinger, 1998; Kowalchuk & Keselman, 2001; Kwok et al., 2007; Littell, Henry, & Ammerman ,1998; Littell, Rendergast, & Natarajan, 2000; Mansour, Nordheim, & Rutledge, 1985; Murphy & Pituch, 2009; Orhan, Eyduran, & Akbas, 2010; Rovine & Molennaar,1998, 2000; Singer & Willett, 2003, Chap. 7; Velicer & Fava, 2003; Verbeke & Molenberghs, 1997; West & Hepworth, 1991; Willett & Sayer, 1994; Wolfinger, 1993, 1996; Wulff & Robinson, 2009). The SAS statements provided can facilitate the implementation of their specification.

1.2.2 Illustration

An illustration is given based on the dataset generated from the linear growth model shown in Figure 1-1 with the ARH(1) level-1 error covariance structure and the UN level-2 error covariance structure. Population parameters are given in Table 1-2. The sample size of 300 was used (Muthén & Muthén, 2002). The RANDNORMAL function in SAS PROC IML was used to generate multivariate normal data based on the population model-implied mean vector $\boldsymbol{\mu}$, shown in Equation 1-4, and the population model-implied variance-covariance matrix $\boldsymbol{\Sigma}$, shown in Equation 1-5, of **y** and **x.** The population mean vector and covariance matrix as well as sample mean and covariance are reported in Table 1-2.

The parameter estimates resulting from fitting ARH(1) with PROC CALIS (the SEM approach) and PROC MIXED (the HLM approach), given in Table 1-3, are very close and verify each other. Furthermore, the fit results from PROC CALIS (chi-square = 11.076 with df = 6, p = .086; CFI = .998; NNFI = .996; RMSEA = .05) indicate good model fit.

1.3 Second-Order Latent Growth Models

A second-order latent growth model can be a curve-of-factors model or a factor-of-curves model (e.g., Duncan, Duncan, & Strycker, 2006, Chap. 4; Hancock, Kuo, & Lawrence, 2001). The curve-of-factors model is used to investigate the growth trajectory of a construct over time. It incorporates the multiple indicators (items) representing the latent construct observed at different time points into the model. Repeated latent constructs are termed the first-order factors and growth factors (i.e., random intercept and slope) are termed the second-order factors. The factor-of-curves model includes higher order common factors for random intercepts and random slopes associated with manifest variables used in LGM. In this model, growth factors are the first-order factors and the underlying common intercept and common slope are the second-order factors, accounting for common developmental patterns. Both the curve-of-factors model and the factor-of-curves model can be well handled by using PROC CALIS.

In this section, the second-order demonstration is given for the curve-of-factors model. The model has several advantages (Blozis, 2006; Preacher et al., 2008; Sayer & Cumsille, 2001). First, the model explicitly recognizes the presence of measurement errors in repeated measures and captures the growth of repeated constructs adjusted for the presence of these errors. Second, the model allows the separation of variation due to departure from the trajectory (temporal instability) and unique variation due to measurement error (unreliability). Third, the model permits the test of longitudinal factorial invariance.

For example, let latent construct F be measured by three indicators, observed at four occasions, denoted by $y_{1t} - y_{3t}$, t = 1, 2, 3, 4. The latent constructs $F_1 - F_4$ at the four occasions are the first-order factors, and the growth factors, denoted by η_α and η_β, are the second-order factors. Let ξ, measured by indicators $x_1 - x_3$, be a time-invariant latent predictor for the growth factors. The second-order curve-of-factors LGM is pictorially presented in Figure 1-2, and can be expressed in matrix form as

$$
\begin{aligned}
\mathbf{y} &= \mathbf{\Lambda}_{\mathrm{y}}\boldsymbol{F} + \boldsymbol{\varepsilon}, \\
\mathbf{x} &= \mathbf{\Lambda}_{\mathrm{x}}\xi + \boldsymbol{\delta}, \\
\boldsymbol{F} &= \mathbf{\Lambda}_{\mathrm{y}}^{*}\boldsymbol{\eta} + \boldsymbol{\zeta}_{\boldsymbol{F}}, \\
\boldsymbol{\eta} &= \mathbf{\Gamma}_0 + \mathbf{\Gamma}_{\xi}\xi + \boldsymbol{\zeta}_{\boldsymbol{\eta}},
\end{aligned}
\tag{1-11}
$$

where $\boldsymbol{y} = [y_{11}\ y_{21}\ y_{31}\ y_{12}\ y_{22}\ y_{32}\ y_{13}\ y_{23}\ y_{33}\ y_{14}\ y_{24}\ y_{34}]'$, $\mathbf{x} = [x_1\ \ x_2\ \ x_3]'$, $\boldsymbol{F} = [F_1\ \ F_2\ \ F_3\ \ F_4]'$, $\boldsymbol{\eta} = [\eta_{\alpha}\ \ \eta_{\beta}]'$, $\boldsymbol{\varepsilon} = [\varepsilon_{11}\ \ \varepsilon_{21}\ \ \varepsilon_{31}\ \ \varepsilon_{12}\ \ \varepsilon_{22}\ \ \varepsilon_{32}\ \ \varepsilon_{13}\ \ \varepsilon_{23}\ \ \varepsilon_{33}\ \ \varepsilon_{14}\ \ \varepsilon_{24}\ \ \varepsilon_{34}]'$, $\boldsymbol{\delta} = [\delta_1\ \delta_2\ \delta_3]'$, $\boldsymbol{\zeta}_{\boldsymbol{F}} = [\zeta_{F_1}\ \ \zeta_{F_2}\ \ \zeta_{F_3}\ \ \zeta_{F_4}]'$, and $\boldsymbol{\zeta}_{\boldsymbol{\eta}} = [\zeta_{\eta_{\alpha}}\ \zeta_{\eta_{\beta}}]'$. $\mathbf{\Lambda}_{\mathrm{y}}$ and $\mathbf{\Lambda}_{\mathrm{x}}$ in the measurement model denote the loading matrices showing the relations of indicators to their underlying constructs. One of the indicators for each construct is selected as the reference indicator and its loading is fixed to 1 at each time point for scaling purpose (Blozis, 2006; Sayer & Cumsille, 2001; Chan, 1998). $\mathbf{\Lambda}_{\mathrm{y}}^{*}$ denotes the loading matrix (with fixed values) of $\boldsymbol{F}$ on $\boldsymbol{\eta}$. $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_{\xi}$ denote, respectively, the vector of intercepts and slopes of the regressions of the growth factors $\boldsymbol{\eta}$ on the latent predictor ξ. $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ denote, respectively, the measurement errors for $\boldsymbol{F}$ and ξ. $\boldsymbol{\zeta}_{\boldsymbol{F}}$ and $\boldsymbol{\zeta}_{\boldsymbol{\eta}}$ denote, respectively, the errors reflecting the departure of the repeated latent constructs from the trajectory and the errors associated with the random intercept and slope. $\boldsymbol{\varepsilon}$ and $\boldsymbol{\zeta}_{\boldsymbol{F}}$ are level-1 errors, and $\boldsymbol{\delta}$ and $\boldsymbol{\zeta}_{\boldsymbol{\eta}}$ are level-2 errors. The assumptions include (a) $\boldsymbol{\varepsilon}$, $\boldsymbol{\zeta}_{\boldsymbol{F}}$, $\boldsymbol{\delta}$, and $\boldsymbol{\zeta}_{\boldsymbol{\eta}}$ are uncorrelated; (b) ζ_{F_1}, ζ_{F_2}, ζ_{F_3}, and ζ_{F_4} are uncorrelated; (c) The measurement errors associated with different indicators are uncorrelated. However, those associated with the same indicator at different points in time are allowed to covary; (d) $\zeta_{\eta_{\alpha}}$ and $\zeta_{\eta_{\beta}}$ are correlated (see, e.g., Blozis, 2006; Bollen & Curran, 2006, p. 249; Preacher et al., 2008, p. 63; Sayer & Cumsille, 2001). The correlated measurement errors are depicted in Figure 1-2 by the linkage of three solid lines with four arrowheads, one line for each indicator. Based on the above assumptions, the structures of $\mathbf{\Psi}_{\zeta_F}$ and $\mathbf{\Theta}_{\boldsymbol{\delta}}$ are both TOEPH(1), the structure of $\mathbf{\Psi}_{\zeta_{\eta}}$ is UN, and the covariance structure of the correlated measurement errors needs to be identified.

Weak factorial invariance is usually assumed in the second-order LGM to allow meaningful interpretations of growth trajectories. Weak factorial invariance requires the equality of the loadings in the measurement model for the same indicator across time (Blozis, 2006; Bollen and Curran, 2006, p. 255; Chan, 1998; Hancock et al., 2001; Preacher et al., 2008, p. 63; Sayer & Cumsille, 2001).

Program 2 in Appendix 1-A demonstrates using PROC CALIS to fit a second-order linear trajectory model for four equally spaced time points, in which AR(1) is specified for three series, ε_{1t}, ε_{2t}, and ε_{3t}, t = 1, 2, 3, 4, TOEPH(1) is specified for $\boldsymbol{\zeta_F}$ and $\boldsymbol{\delta}$, and UN is specified for $\boldsymbol{\zeta}_{\boldsymbol{\eta}}$. The LINEQS statement, based on Equation 1-11, is given below. It is an extended version from that in Program 1 by incorporating the measurement models for $\boldsymbol{F}$ and the latent predictor ξ.

```
LINEQS
   Y11 = 1 F1 + EY11,    Y21 = LY21F1 F1 + EY21,    Y31 = LY31F1 F1 + EY31,
   Y12 = 1 F2 + EY12,    Y22 = LY22F2 F2 + EY22,    Y32 = LY32F2 F2 + EY32,
   Y13 = 1 F3 + EY13,    Y23 = LY23F3 F3 + EY23,    Y33 = LY33F3 F3 + EY33,
```

Y14 = 1 F4 + EY14, Y24 = LY24F4 F4 + EY24, Y34 = LY34F4 F4 + EY34,
X1 = 1 F7 + EX1, X2 = LX2F7 F7 + EX2, X3 = LX3F7 F7 + EX3,
F1 = 1 F_Alpha + 0 F_Beta + EZF1,
F2 = 1 F_Alpha + 1 F_Beta + EZF2,
F3 = 1 F_Alpha + 2 F_Beta + EZF3,
F4 = 1 F_Alpha + 3 F_Beta + EZF4,
F_Alpha = GA00 INTERCEPT + GA10 F7 + EZF5,
F_Beta = GA01 INTERCEPT + GA11 F7 + EZF6,
F7 = F7_int INTERCEPT + EZF7;

where F1–F4 are the first-order factors at the four occasions, F_ALPHA and F_BETA represent the second-order latent factors η_{α_i} and η_{β_i}. Y*jt* denotes the observed score on the *j*th indicator for *F* at occasion *t*, *j* = 1, 2, 3; *t* = 1, 2, 3, 4. X*j* (*j* = 1, 2, 3) denotes the observed score on the *j*th indicator for construct ξ, named F7. The loadings of Y1*t* on F*t* (*t* = 1, 2, 3, 4) are fixed to 1. LY*jt*F*t* represents the estimate of the first-order loading of Y*jt* on F*t*, *j* = 2, 3; *t* = 1, 2, 3, 4. EY*jt* denotes the corresponding measurement error. Similarly, the loadings of X1, X2, and X3 on F7 are 1, LX2F7, and LX3F7, respectively. EX1–EX3 are the corresponding measurement errors. Second-order factor loadings are fixed values (in $\mathbf{\Lambda}_y^*$). Level-1 errors ζ_{F_1}, ζ_{F_2}, ζ_{F_3}, and ζ_{F_4} are named EZF1–EZF4, and level-2 errors ζ_{η_α} and ζ_{η_β} are named EZF5 and EZF6. F7_int denotes the mean of F7, and EZF7 is the deviation of F7 from its mean.

The parameters in AR(1) for ε_{jt}, j = 1, 2, 3, include error variance $\sigma^2_{\varepsilon_j}$ and the autocorrelation at lag 1 $\phi_{1\varepsilon_j}$. The resulting ECM for $\boldsymbol{\varepsilon}$ is given by

$$\mathbf{\Theta}_\varepsilon = Cov[\varepsilon_{11} \quad \varepsilon_{21} \quad \varepsilon_{31} \quad \varepsilon_{12} \quad \varepsilon_{22} \quad \varepsilon_{32} \quad \varepsilon_{13} \quad \varepsilon_{23} \quad \varepsilon_{33} \quad \varepsilon_{14} \quad \varepsilon_{24} \quad \varepsilon_{34}]'$$

$$= \begin{bmatrix}
\sigma^2_{\varepsilon_1} \\
0 & \sigma^2_{\varepsilon_2} \\
0 & 0 & \sigma^2_{\varepsilon_3} \\
\phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & 0 & 0 & \sigma^2_{\varepsilon_1} \\
0 & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & 0 & 0 & \sigma^2_{\varepsilon_2} \\
0 & 0 & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & 0 & 0 & \sigma^2_{\varepsilon_3} \\
\phi^2_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & 0 & 0 & \phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & 0 & 0 & \sigma^2_{\varepsilon_1} \\
0 & \phi^2_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & 0 & 0 & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & 0 & 0 & \sigma^2_{\varepsilon_2} \\
0 & 0 & \phi^2_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & 0 & 0 & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & 0 & 0 & \sigma^2_{\varepsilon_3} \\
\phi^3_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & 0 & 0 & \phi^2_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & 0 & 0 & \phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & 0 & 0 & \sigma^2_{\varepsilon_1} \\
0 & \phi^3_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & 0 & 0 & \phi^2_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & 0 & 0 & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & 0 & 0 & \sigma^2_{\varepsilon_2} \\
0 & 0 & \phi^3_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & 0 & 0 & \phi^2_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & 0 & 0 & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & 0 & 0 & \sigma^2_{\varepsilon_3}
\end{bmatrix}, \quad (1\text{-}12)$$

which can be reexpressed as follows to facilitate readability:

$$\boldsymbol{\Theta}_{\varepsilon} = Cov[\varepsilon_{11} \quad \varepsilon_{12} \quad \varepsilon_{13} \quad \varepsilon_{14} \quad \varepsilon_{21} \quad \varepsilon_{22} \quad \varepsilon_{23} \quad \varepsilon_{24} \quad \varepsilon_{31} \quad \varepsilon_{32} \quad \varepsilon_{33} \quad \varepsilon_{34}]'$$

$$= \begin{bmatrix}
\sigma^2_{\varepsilon_1} \\
\phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \sigma^2_{\varepsilon_1} \\
\phi^2_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \sigma^2_{\varepsilon_1} \\
\phi^3_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \phi^2_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \sigma^2_{\varepsilon_1} \\
0 & 0 & 0 & 0 & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & \phi^2_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & \phi^3_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \phi^2_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{\varepsilon_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \sigma^2_{\varepsilon_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi^2_{1\varepsilon_3}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \sigma^2_{\varepsilon_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi^3_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \phi^2_{1\varepsilon_3}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \sigma^2_{\varepsilon_3}
\end{bmatrix}, \quad (1-13)$$

where $\phi^k_{1\varepsilon_j}$ is the autocorrelation at lag k for ε_{jt}, k = 1, 2, 3; t = 1, 2, 3, 4; j = 1, 2, 3, with the constraints of $|\phi_{1\varepsilon_j}| < 1$. For each indicator, their error variances at different time points are equal, their error autocovariances at lag 1 are equal, and their error autocovariances at lag 2 are equal as well. Therefore, the STD and COV statements are given as follows:

```
STD
   EY11−EY14=4*VARE1, EY21−EY24=4*VARE2, EY31−EY34=4*VARE3,
   EX1=VAREX1, EX2=VAREX2, EX3=VAREX3,
   EZF1=VARZF1, EZF2=VARZF2, EZF3=VARZF3, EZF4=VARZF4,
   EZF5=VARE_Intercept, EZF6=VARE_Slope, EZF7=VARZF7;
COV
      /* for the level-1 measurement errors associated with indicator 1 */
   EY11 EY12=COV1_lag1, EY12 EY13=COV1_lag1, EY13 EY14=COV1_lag1,
   EY11 EY13=COV1_lag2, EY12 EY14=COV1_lag2, EY11 EY14=COV1_lag3,
      /* for the level-1 measurement errors associated with indicator 2 */
   EY21 EY22=COV2_lag1, EY22 EY23=COV2_lag1, EY23 EY24=COV2_lag1,
   EY21 EY23=COV2_lag2, EY22 EY24=COV2_lag2, EY21 EY24=COV2_lag3,
      /* for the level-1 measurement errors associated with indicator 3 */
   EY31 EY32=COV3_lag1, EY32 EY33=COV3_lag1, EY33 EY34=COV3_lag1,
   EY31 EY33=COV3_lag2, EY32 EY34=COV3_lag2, EY31 EY34=COV3_lag3,
      /* for the level-2 errors associated with growth factors */
   EZF5 EZF6=CZF5ZF6;
```

in which VARE1, VARE2, and VARE3 represent, respectively, the estimates of the common variances $\sigma^2_{\varepsilon_1}$, $\sigma^2_{\varepsilon_2}$, and $\sigma^2_{\varepsilon_2}$. VAREX1−VAREX3 represent the estimates of variances of $\delta_1 - \delta_3$. VARZF1−VARZF4 represent the estimates of variances of $\zeta_{F_1} - \zeta_{F_4}$. VARE_Intercept, VARE_Slope, and CZF5ZF6 represent, respectively, the estimates of variances and covariance of the second-order factor errors ζ_{η_α} and ζ_{η_β}. VARZF7 represents the estimate of variance of the latent predictor ξ. COV1_lag1, COV1_lag2, and COV1_lag3 represent, respectively, the estimates of common autocovariance at lags 1, 2, 3 for ε_{1t}. Similarly, COV2_lag1, COV2_lag2, and COV2_lag3 represent those for ε_{2t}, and COV3_lag1, COV3_lag2, and COV3_lag3 represent those for ε_{3t}.

The following PARAMETERS statement is needed to bring three additional parameters, $\phi_{1\varepsilon_1}$, $\phi_{1\varepsilon_2}$, and $\phi_{1\varepsilon_3}$, based on Equation 1-13:

```
PARAMETERS    PHI1    PHI2    PHI3;
      /* for the level-1 measurement errors associated with indicator 1 */
   COV1_lag1=PHI1*VARE1;    COV1_lag2= (PHI1**2)*VARE1;
   COV1_lag3=(PHI1**3)*VARE1;
      /* for the level-1 measurement errors associated with indicator 2 */
   COV2_lag1=PHI2*VARE2;    COV2_lag2=(PHI2**2)*VARE2;
   COV2_lag3=(PHI2**3)*VARE2;
      /* for the level-1 measurement errors associated with indicator 3 */
   COV3_lag1=PHI3*VARE3;    COV3_lag2=(PHI3**2)*VARE3;
   COV3_lag3=(PHI3**3)*VARE3;
```

in which PHI1, PHI2, and PHI3 represent the estimates of $\phi_{1\varepsilon_1}$, $\phi_{1\varepsilon_2}$, and $\phi_{1\varepsilon_3}$. 'COV1_lag1= PHI1*VARE1' corresponds to the requirement that the common autocovariance at lag 1 for ε_{1t} be equal to $\phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1}$. 'COV1_lag2=(PHI1**2)*VARE1' corresponds to the requirement that the common autocovariance at lag 2 be equal to $\phi^2_{1\varepsilon_1}\sigma^2_{\varepsilon_1}$. 'COV1_lag3=(PHI1**3)*VARE1' corresponds to the requirement that the autocovariance at lag 3 be equal to $\phi^3_{1\varepsilon_1}\sigma^2_{\varepsilon_1}$. The relevant statements for ε_{2t} and ε_{3t} are given similarly.

The constraints of $|\phi_{1\varepsilon_1}|<1$, $|\phi_{1\varepsilon_2}|<1$, $|\phi_{1\varepsilon_3}|<1$ are specified by the following BOUNDS statement:

```
BOUNDS
   –1.< PHI1<1.,   –1.< PHI2<1.,   –1.< PHI3<1.;
```

Under the assumption of weak factorial invariance, the LINCON statement should be added to equalize the loadings for the same indicator across occasions as follows:

```
LINCON
   LY21F1=LY22F2, LY21F1=LY23F3, LY21F3=LY24F4,
   LY31F1=LY32F2, LY31F1=LY33F3, LY31F3=LY34F4;
```

1.3.1 Illustration

Another illustration is given with another dataset of size 300 generated from the second-order LGM in Figure 1-2. The population parameters with the AR(1) covariance structure for level-1 error processes ε_{1t}, ε_{2t}, and ε_{3t} and the sample covariance matrix of **y** and **x** resulting from the simulated dataset are presented in Table 1-4. The RANDNORMAL function in PROC IML was used again to generate multivariate normal data based on the population model-implied mean vector and variance-covariance matrix of **y** and **x** in Figure

1-2 (see Appendix 1-C for the derivation). The parameter estimates by fitting AR(1) for ε_{1t}, ε_{2t}, and ε_{3t} are summarized in Table 1-5. The resulting parameter estimates are all close to the corresponding population values specified in Table 1-4 and the model fit is excellent (chi-square = 90.49 with df = 109, p = .9009; CFI =1.0; NNFI =1.0; RMSEA <.0001).

1.4 Conclusion

We present a systematic coding for various level-1 error covaraince structures in LGM by using SAS PROC CALIS. The joint use of the STD, COV, PARAMETERS, LINCON, and BOUNDS statements in PROC CALIS can be extended for other types of ECM in a similar way to meet analysts' need. The advantages to use PROC CALIS include its flexibility in specifying ECM and its capabilities to better assess model fit for balanced complete data and to deal with latent constructs. A tutorial on the syntax has been provided for manifest variables and latent constructs. It is our hope that the coding provided will help applied researchers with LGM studies.

Although our demonstration is based on linear growth models, SAS statements in PROC CALIS for specifying level-1 ECM are applicable for quadratic and polynomial growth models. Theory may suggest appropriate growth models. As mentioned previously, misspecification of the growth model can lead to incorrect selection of the error covariance structure. The coding provided is useful when the growth model has been well determined and the level-1 error covariance structure is to be identified.

2. Identifying Level-1 Error Covariance Structures in Latent Growth Modeling

2.1 Introduction

The processes underlying level-1 errors can be categorized as stationary and nonstationary. An error process $\{\varepsilon_t\}$ is stationary if the mean of ε_t and the covariance of ε_t and ε_{t-k} are both time-invariant, where k is an arbitrary integer. More specifically, $\{\varepsilon_t\}$ is stationary if (a) $E(\varepsilon_t) = \mu_\varepsilon$, which is a constant, always assumed to be zero, and (b) $Var(\varepsilon_t) = \sigma_\varepsilon^2 = \sigma_0$ and $Cov(\varepsilon_t, \varepsilon_{t-k}) = \sigma_k$, $k > 0$. That is, the variances are equal and the covariances at lag k are equal. It follows that $\rho_k = \sigma_k / \sigma_0$, $k \geq 0$. σ_k and ρ_k are called, respectively, the autocovaraince and autocorrelation of ε_t at lag k (See, e.g., Box, Jenkins, & Reinsel, 1994, Chap. 3). Table 1-1 and Appendix 1-B summarize error covariance structures for equally spaced data, frequently seen in LGM studies. Stationary structures include those resulting from ARMA processes, Toeplitz with q bands, q = 1, …, T. denoted by TOEP(q), and compound symmetry, denoted by CS. TOEP(T), also simply denoted by TOEP, is the saturated stationary structure (the most general stationary structure). Nonstationary structures include TOEPH(q) (heterogeneous TOEP(q)), CSH (heterogeneous CS), ARH(1) (heterogeneous AR(1)), and UN(q) (unstructured with q bands). TOEPH(T) is simply denoted by TOEPH. UN(T), simply denoted by UN, is the saturated nonstationary structure.

2.2 An Effective Approach for Identifying an "Optimal" Level-1 Error Covariance Structure

As indicated in Grimm and Widaman (2010), although linear growth curve models are often fit because of their ease in estimation, theory may suggest that more complex models be used, as they have interpretable parameters and can better capture developmental patterns. Correctly specifying the growth model might lead to a simple covariance structure. In this case, using a simpler growth model and searching for an "optimal" autocorrelational structure is not worthwhile.

Assuming that the growth pattern has been well determined, researchers still need to deal with the problem of specification of the error covariance structure. Since little theoretical knowledge about the error covariance structure is available, a specification search needs to be conducted. Those that are clearly inconsistent with data characteristics should first be ruled out (Littell et al., 2006, p. 177). For example, TOEP(1) and TOEPH(1) are inappropriate if error covariances are significant. The ARMA family are inappropriate if the process is nonstationary.

An approach to search for an "optimal" level-1 structure from those shown in Table 1-1 and Appendix 1-B is proposed, assuming that the data are equally spaced in time. The approach attempts to identify a structure that is as simple as possible under the condition of achieving model fit. The procedure is implemented as follows:

Stage 1: Testing for stationarity of the error process.

The conditions of stationarity include the equality of error variances and the equality of error autocovariances at any lag. For example, for the model shown in Figure 1-1, the null hypothesis of stationarity is given by H0: $\sigma_{\varepsilon_1}^2 = \sigma_{\varepsilon_2}^2 = \sigma_{\varepsilon_3}^2 = \sigma_{\varepsilon_4}^2$, $\sigma_{\varepsilon_2\varepsilon_1} = \sigma_{\varepsilon_3\varepsilon_2} = \sigma_{\varepsilon_4\varepsilon_3}$,

$\sigma_{\varepsilon_3\varepsilon_1} = \sigma_{\varepsilon_4\varepsilon_2}$. Stationarity can be tested by using the chi-square difference test, where the difference of the chi-square fit statistic constrained by H_0 and the unconstrained chi-square fit statistic is distributed as the chi-square distribution with degrees of freedom being the number of constraints under H_0 ($df = 6$ for the above example). The former can be obtained by fitting TOEP and the latter by fitting UN.

Stage 2: Identifying an "optimal" level-1 error covariance structure.

The sequential chi-square difference test (SCDT, adapted from Anderson & Gerbing, 1988) is used. If stationarity, tested in Stage 1, is supported, then identify the structure from the stationary class; otherwise identify the structure from the nonstationary class. In addition to SCDT, the tests for autocorrelations and partial autocorrelations are used to help identify the order of an AR(p) or MA(q) process when stationarity is satisfied.

The saturated stationary structure is TOEP, and the saturated nonstationary structure is UN. The structure to be identified is as simple (parsimonious) as possible under the condition that it is nested within the saturated structure and it produces no significantly worse model fit than the saturated structure. Structure search is conducted sequentially, starting from the simplest (i.e., the most constrained) model and then a less constrained one. The simplest stationary structure is TOEP(1), also known as variance components (VC) (Murphy & Pituch, 2009), and the simplest nonstationary structure is TOEPH(1) (= UN(1)). At each step, the chi-square difference test is used to examine if the model fit with the current temporary structure, denoted by M_T, is significantly different from the model fit with the saturated structure, denoted by M_S. The null hypothesis showing no fit difference between the temporary structure and the saturated structure is denoted by H_0: $M_T = M_S$. If the test is significant, indicating significantly worse model fit resulting from the current M_T, then update M_T with a less constrained one and compare the model fit between the new M_T and M_S. The process is terminated by returning the temporary structure as the final structure when the test is nonsignificant.

A structure nested within M_S implies that the structure is just the M_S with some constraints. For example, AR(1) is equivalent to TOEP with the constraints of setting $\rho_k = \rho_1^{\,k}$ for $k > 1$. CS is equivalent to TOEP with the constraints of setting ρ_k to be equal for $k > 0$. TOEP(2) is equivalent to TOEP with the constraints of setting ρ_k to be zero for $k > 1$. The criterion to compare the degree of being constrained between two structures nested within M_S is the number of parameters. The structure with fewer parameters is the more constrained one. It is likely that there are two or more structures with the same number of parameters. For example, MA(1), TOEP(2), CS, and AR(1) are all stationary structures with two parameters. MA(1) is equivalent to TOEP with the constraints of setting $\rho_1 = -\theta_1 / (1+\theta_1^2)$ and $\rho_k = 0, k > 1$. MA(1) differs from TOEP(2) in that it hypothesizes a particular structure for ρ_1 rather than assuming just an association. Although MA(1) is structurally different from TOEP(2), their chi-square fit statistics are identical because their fit functions are the same and their parameters are one-to-one related. If MA(1) can be identified by using a different approach (to be addressed below), there is no need to consider TOEP(2) during the search process because it cannot improve fit at all. On the other hand, CS requires that the autocorrelations at lag k ($k > 0$) be all equal. Although autocorrelations ρ_k are allowed to exist for $k > 1$, they are constrained to be identical. If CS has been well identified, we need not consider AR(1) or MA(1). Typically, adjacent errors tend to be more correlated than errors farther apart in time (e.g., Diggle, Liang, & Zeger, 2002, p. 82; Guttman, 1954; Littell et al., 2006, p. 175). Thus, AR(1) should be more frequently used than MA(1), TOEP(2), and CS because its autocorrelation function decays exponentially.

AR(p) and MA(q) can be identified by using traditional time series methodology. Once they are identified, they are further compared with M_S (TOEP) to check model fit. The autocorrelation function (ACF) is used to identify MA(q) and the partial autocorrelation functions (PACF) to identify AR(p) (e.g., Box, Jenkins and Reinsel, 1994, Chap. 10). The ACF of an MA(q) process ($q \geq 1$) has a cutoff after lag q (that is, the autocorrelation at lag k is zero for $k > q$). However, its PACF has no cutoff. In contrast, the PACF of an AR(p) process ($p \geq 1$) has a cutoff after lag p. Its ACF does not exhibit any cutoff. The cutoff point for an MA (AR) process can be determined by examining the significance/nonsignificance of the sample autocorrelations (partial autocorrelations) at lags k, $k = 1, \ldots, T-1$. If ACF and PACF both have a cutoff after lag 0, then TOEP(1) is identified. If autocorrelations and partial autocorrelations are significant at all lags (i.e., with no cut-off), then neither MA(q) nor AR(p) is appropriate, and we need to proceed to identify an ARMA(p,q). Table 2-1 summarizes the decision rule.

If MA(q) is identified but AR(p) is not, then MA(q) is further compared with TOEP to check the model fit by using the chi-square difference test. MA(q) is selected as the final structure only when the test is non-significant. If AR(p) is identified but MA(q) is not, then it is further compared with TOEP. AR(p) is selected as the final structure only when its fit is not significantly different from the fit by TOEP. If MA(q) and AR(p) are both identified, they both need to be compared with TOEP. The one with acceptable model fit and fewer parameters is selected as the final structure. The identification of an MA(q) process must pass the examination of ACF and the chi-square difference test. Similarly, the identification of an AR(p) process must pass the examination of PACF and the chi-square difference test. When MA(q) and AR(p) are both identified and with the same order, that is, $p = q$, then the structure with the smaller AIC (Akaike, 1974) is selected. If neither MA(q), identified by using ACF, nor AR(p), identified by using PACF, achieves model fit, then go to the next step to examine ARMA(p,q) ($p \neq 0$, $q \neq 0$) processes.

According to the degree of parsimony, the ARMA(p,q) ($p \neq 0$, $q \neq 0$) processes to be examined are in the order of ARMA(1,1), (ARMA(2,1), ARMA(1,2)), (ARMA(3,1), ARMA(2,2), ARMA(1,3)), …, and (ARMA(p,q), $p+q = T-2$), all nested within TOEP. The processes within the same parenthesis are those with the same number of parameters. (ARMA(p,q), $p+q = T-1$) are not candidates because they have the same number of parameters as TOEP, leading to zero degree of freedom for the chi-square difference test. The structures without significantly different model fit from the saturated stationary structure TOEP may not be unique.

In sum, specification search is carried out sequentially, for stationary structures, in the order of TOEP(1), CS, AR(p)/MA(q), ARMA(1,1), (ARMA(2,1), ARMA(1,2)), (ARMA(3,1), ARMA(2,2), ARMA(1,3)), …, and (ARMA(p,q), $p+q = T-2$), all nested within TOEP, and, for nonstationary structures, in the order of TOEPH(1), (TOEPH(2), CSH, ARH(1)), TOEPH(3), …, (TOEPH(T), UN(2)), …, and UN($T-1$), all nested within UN. If the number of structures identified is more than one, then the determination is based on AIC. The structure that minimizes AIC is selected.

A flowchart for identifying an "optimal" level-1 error covariance structure is given in Figure 2-1. The procedure to determine the order of ARMA(p,q) is included within a dash-line box in the figure.

2.3 Illustrations

2.3.1 Illustration 1

The first illustration is based on a dataset generated from the linear growth model shown in Figure 1-1 with the ARH(1) level-1 error structure and the UN level-2 error structure. Population parameters and the sample covariance matrix of y_1–y_4 and X are given in Table 2-2. The sample size of 300 was used (Muthén & Muthén, 2002). The influential results by following the procedure given above are also summarized in Table 2-2.

We first tested for stationarity. Since stationarity was rejected by using the chi-square difference test (the chi-square difference $\Delta\chi^2_{\Delta df=6} = \chi^2_{df=7}$ (TOEP) $-\chi^2_{df=1}$ (UN) = 58.411 − 5.17 = 53.241, p < .0001), we proceeded with a sequential search within the nonstationary class. We started by fitting TOEPH(1). Since the model fit with TOEPH(1) was significantly worse than the model fit with UN, the saturated nonstationary structure ($\Delta\chi^2_{\Delta df=6}$ = 25.789, p < .0001), TOEPH(1) was inappropriate. M_T was updated with a less constrained one than TOEPH(1). TOEPH(2), CSH, and ARH(1), all having (T+1) = 5 parameters, are the second simplest structures. Since TOEPH(2) and CSH also resulted in significantly worse model fit than UN ($\Delta\chi^2_{\Delta df=5}$ = 14.771 (p = .011) and 13.206 (p = .022)), neither one was appropriate. Fitting ARH(1) led to nonsignificant results ($\Delta\chi^2_{\Delta df=5}$ = 5.906, p = .315). Therefore, the sequential search was terminated by choosing ARH(1) as the level-1 error covariance structure. The improvement of ARH(1) over TOEPH(1) could be verified by the significant chi-square difference ($\Delta\chi^2_{\Delta df=1}$ = 19.883, p < .0001), the adequate model fit with ARH(1) ($\chi^2_{df=6}$ = 11.076, p = .086) and the inadequate model fit with TOEPH(1) ($\chi^2_{df=7}$ = 30.96, p = .0002). The final structure identified, ARH(1), is just the one specified in the population model.

2.3.2 Illustration 2

In the second illustration, four datasets (each with N = 300) were generated from the linear growth model in Figure 1-1 with four cases of the level-1 covariance structures: AR(1), AR(2), MA(2), and ARMA(1,1). Population parameters are summarized in Table 2-3. Results of identifying an "optimal" level-1 error covariance structure based on each of the four datasets are reported in Table 2-4.

For the dataset generated from the AR(1) process (Case 1), we first test for stationarity. Since stationarity was supported ($\Delta\chi^2_{\Delta df=6} = \chi^2_{df=7}$ (TOEP) $-\chi^2_{df=1}$ (UN) = 6.321–3.292 = 3.029, p = .805), we proceeded with a specification search within the stationary class. We started by fitting TOEP(1). Since TOEP(1) produced significantly worse model fit than TOEP (the saturated stationary structure) ($\Delta\chi^2_{\Delta df=3}$ = 85.035, p < .0001), TOEP(1) was excluded. CS was then fitted. The resulting chi-square difference was still significant ($\Delta\chi^2_{\Delta df=2}$ = 72.067, p < .0001) and CS was inappropriate. Examining the sample ACF and PACF, we found that the autocorrelations were significant at lags 1 and 2 but not at lag 3. The ACF had a cutoff after lag 2. Moreover, the partial autocorrelations were significant at lag 1 only. The PACF showed a cutoff after lag 1. Thus, both AR(1) and MA(2) were identified. Since AR(1) is more parsimonious than MA(2), AR(1) was first compared with TOEP. The chi-square difference test indicated nonsignificance ($\Delta\chi^2_{\Delta df=2}$ = .145, p = .930). Hence, AR(1) was selected as the final process. It is unnecessary to compare MA(2) with TOEP because AR(1) has been well identified and is more parsimonious than MA(2). Fitting AR(1) also led to acceptable overall model fit ($\chi^2_{df=9} = 6.466, p = .693$).

For the dataset generated from the AR(2) process (Case 2), stationarity was supported ($\Delta\chi^2_{\Delta df=6} = 3.573$, $p = .734$). We then conducted a specification search for an "optimal" structure within the stationary class. We still started by fitting TOEP(1) and then CS. They were inappropriate because of the significantly worse model fit than TOEP ($\Delta\chi^2 = 283.894$ and 282.134, $p < .0001$). Examining the sample ACF and PACF, we found that the PACF had a cutoff after lag 2 and ACF had no cutoff (The autocorrelations were significant at all lags). Therefore an AR(2) process was identified. Since the model fit between AR(2) and TOEP was nonsignificant ($\Delta\chi^2_{\Delta df=1} = .088$, $p = .766$), AR(2) was identified as the final process. The overall model fit with AR(2) was satisfactory ($\chi^2_{df=8} = 4.853, p = .773$).

For the dataset generated from the MA(2) process (Case 3), stationarity was supported again ($\Delta\chi^2_{\Delta df=6} = 3.142$, $p = .791$). During the identification process, TOEP(1) and CS were excluded because they produced the worse model fit than TOEP ($\Delta\chi^2 = 20.645$, $p < .0001$ for both TOEP(1) and CS). Subsequently, the ACF showed a cutoff after lag 2 and the PACF showed no cutoff, suggesting the adoption of the MA(2) process. Since the model fit between MA(2) and TOEP was nonsignificant ($\Delta\chi^2_{\Delta df=1} = .018$, $p = .893$), MA(2) was selected as the final structure. The overall model fit with MA(2) was satisfactory ($\chi^2_{df=8} = 5.51, p = .702$).

For the dataset generated from the ARMA(1,1) process (Case 4), stationarity was supported ($\Delta\chi^2_{\Delta df=6} = 2.722$, $p = .842$). During the search process, TOEP(1) and CS were excluded because of significantly worse model fit ($\Delta\chi^2 = 213.409$ and 169.279, $p < .0001$). Moreover, neither the ACF nor the PACF had a cutoff (The autocorrelations and partial autocorrelations were significant at all lags), and therefore the SCDT was conducted to identify an ARMA process, starting from ARMA(1,1). Since the model fit between ARMA(1,1) and TOEP was nonsignificant ($\Delta\chi^2_{\Delta df=1} = .059$, $p = .808$), ARMA(1,1) was the final choice. Its overall model fit was satisfactory ($\chi^2_{df=8} = 3.6, p = .892$).

It appears that the final level-1 error covariance structure identified in each case was just that specified in the population model and the corresponding parameter estimates were close to their parameter values. The results have reflected the usefulness of the approach proposed.

2.4 Discussion

As mentioned previously, the impact of the misspecification of the level-1 error covariance structure is substantial. It thus becomes important to correctly identify an error covariance structure. However, relevant issues were not specifically addressed in the LGM-related literature. In this study, we have proposed an effective approach to deal with its identification. The test for stationarity, the sequential chi-square difference test, and the tests for autocorrelations and partial autocorrelations are used, based on the principle of improving parsimony after model fit has been achieved. The approach has been illustrated with simulated data. The satisfactory results reflect the usefulness of the approach. It is recommended that the approach be used in LGM empirical studies to reduce the possibility of model misspecification.

To implement the approach, we need SEM software that allows nonlinear constraints on parameters. SAS PROC CALIS (SAS Institute Inc., 2010), LISREL (Jöreskog & Sörbom, 2001), EQS (Bentler & Wu, 2005), and M*plus* (Muthén & Muthén, 2007) are readily available.

Our demonstrations were based on a simple linear growth model. If alternative growth forms (e.g., a quadratic growth model) can better capture developmental patterns, they should be used instead of the simple linear form (Grimm & Widaman, 2010). The approach proposed in this study applies as well for higher-order growth models.

A second-order factor structure is used to investigate the growth trajectory of a construct over time. It incorporates the multiple indicators (items) representing the latent construct observed at different time points into the model, which is known as the second-order latent growth model (e.g., Blozis, 2006; Bollen & Curran, 2006, Chap. 8; Chan, 1998; Hancock, Kuo, & Lawrence, 2001; Preacher et al., 2008, Chap. 3; Sayer & Cumsille, 2001). Repeated latent constructs are termed the first-order factors and growth factors (i.e., intercept and slope) are termed the second-order factors. Weak factorial invariance is assumed in the second-order LGM to allow meaningful interpretations of growth trajectories. Weak factorial invariance requires the equality of the loadings in the measurement model for the same indicator across time. The measurement errors associated with different indicators are assumed to be uncorrelated. However, those associated with the same indicator at different points in time are allowed to covary. The covariance structures of correlated measurement errors need to be identified, and can be done by using the approach proposed in this study.

There exist some limitations of this study. First, the data are assumed to be equally spaced in time Secondly, when two or more structures are identified, the structure that minimizes AIC is selected. The decision rule is not an objective one. How to determine the final selection under the situation with a suitable statistical test needs further investigation. Thirdly, an "optimal" level-1 structure is selected from those shown in Table 1-1 and Appendix 1-B. The reasons to choose those stationary and nonstationary structure candidates include that they are commonly seen in the LGM literature and that the stationary ones are nested within TOEP and the nonstationary ones are nested within UN, satisfying the nested relationship required by the chi-square difference test. How to deal with other types of structures needs to be further studied.

3. Assessing the Change in Investor Sentiment over Time

3.1 Introduction

The history of the financial market is peppered with many remarkable events, such as the Great Crash of 1929, the Internet bubble of the 1990s, and the U.S. housing bubble that burst in 2007. Just prior to these events, the markets were full of optimistic forecasting. In 1929 before the Great Crash, the world was experiencing high commercial growth. In the 1990s, the decade when Internet and e-commerce technologies emerged, technology stocks on Nasdaq rose to unprecedented levels during a two-year period. In the U.S. housing bubble, housing prices peaked in 2005-2006, started to drop substantially in 2007–2008, and led to global financial turmoil in 2008 (Saxton 2008). All of these events started from optimistic trends that were followed by severe crashes. These dramatic fluctuations had considerable difficulty matching the traditional efficient market theory and motivated the emergence of the new field of behavioral finance.

Behavioral finance theory claims that irrational sentiment exists in the market. There are “irrational investors” or “noise traders” (Black 1986, DeLong et al. 1990). During good times, irrational investors become more optimistic as they are reinforced by others jumping on the bandwagon (Brown and Cliff 2004). When stock prices have been driven up, noise traders might become even more bullish tomorrow, and arbitragers must take the risk of a further price rise when they have to buy back the stock. The risk will limit their willingness to bet against noise traders. These optimistic investors will continuously boost market sentiment, and the market will become more and more inefficient. When bad times come, noise traders become overly pessimistic, as they believe others fire-selling various assets will eventually ruin investor confidence in the market. The higher level of noise trading, the more inefficient the market becomes.

3.2 Investor Sentiment on Stocks

3.2.1 Change of sentiment over time

Noise trader theory has emerged as an alternative view to efficient market theory. In noise trader theory, the financial market is not expected to be efficient. Rather, systematic and significant deviations from efficiency persist for a long period (Shleifer 2000). The basis for efficient market theory rests on three assumptions. First, investors are assumed to be rational and therefore will value securities rationally. Second, to the extent that some investors are irrational, their trades are random and thus the effects are canceled out by each other. Third, even if investors are irrational in a similar way, the rational arbitrageurs will nullify the influence of irrational behavior.

However, efficient market theory is challenged both theoretically and empirically. A number of financial anomalies, including the excess volatility of asset prices, the mean reversion of stock returns, and the underpricing of closed-end mutual funds, demonstrate considerable evidence that many investors are not rational and do not follow theoretical efficient market hypotheses.

It is difficult to sustain the case that all investors are rational. Noise trader theory assumes that there are two types of investors in the markets: rational investors and irrational investors. Irrational investors react to irrelevant information and form their own demand on securities. Those investors, called “noise traders”, are influenced by a combination of cognitive biases and psychological habits such as overconfidence (Alpert and Raiffa 1982), overreaction, and trend tracing. They trade on noise more than relevant information (Black 1986). Noise traders follow the pseudo-signals (Black 1986) coming from diverse sources of information such as advice from stock-brokers and financial gurus, falsely believing that they are following reliable "insider information", which will enable them to forecast future returns of risky assets. Investor sentiment results from psychological cognitive biases (Kahneman and Tversky 1979), which can lead investors to buy more stocks even when the price has already risen or sell more stocks even when the price has already dropped. Beliefs based on psychological perception rather than normative economic models are called investor sentiment. Investor sentiment can be considered to be the expectations of irrational market participants.

Psychological evidence shows that people do not deviate from rationality randomly, but rather deviate in the same way (Kahneman and Tversky 1979). Moreover, Shiller (1984) claims that investing in speculative assets is a social activity, so it is plausible that investors' behavior would be influenced by social movements. The evidence for a social movement driving the bull market between the late 1940s and late 1960s would be the growing number of individuals who participated in the market. When noise traders behave socially and exhibit common reaction to rumors, the problem of deviating from rationality becomes even more severe. The trading strategies based on pseudo signals, popular models, cognitive biases, and psychological habits are correlated and lead to an aggregate demand shift.

Social movements are dynamic processes (Shiller 1984). Mutual reinforcement through the exchange of information will form a condition for the emergence of a uniform response to stimulate investor sentiment to move in the same way. Investor sentiment is formed through a process over time (Smidt 1968, Brown and Cliff 2005). Irrational investors will be greatly influenced by others joining in the noise trading. The uptrend or downtrend of investor sentiment could reflect the existence of social movements and thus investors will continuously deviate from rationality.

3.2.2 The influence of industry type on the change in sentiment over time

Noise traders consider investing to be a social activity (Shiller 1984). They are influenced by social trends, especially when dealing with popular topics. Investors are anxious to buy shares of any firms that are in new "glamour" industries and have enormous growth potential (Cooper et al. 2001). Investment in high-technology industries plays a major role in encouraging new technology and has a great chance of bringing about considerable economic profit in the future. Therefore, high-tech stocks should receive more noise trades than non-high-tech stocks. Ofek and Richardson (2003) indicate that Internet stocks are owned relatively more often by individual investors, who are regarded as noise traders (Black 1986). Individual investors might merely snap up Internet or other high-tech stocks (Dorn 2009). It can be expected that the average linear growth trend of sentiment for high-tech stocks is steeper than that for non-high-tech stocks.

3.2.3 The influence of stock size on the change in sentiment over time

The stocks with stronger retail concentrations may be traded more irrationally. Kumer and Lee (2006) find that smaller firms, lower priced firms, and firms with less institutional ownership are associated with stronger retail trading activities. Due to the sparse information available about smaller firms, the market may be less efficient (Hirshleifer 2001). Therefore, we predict that the average rate of change in sentiment for small-size stocks is greater than that for large-size stocks.

3.2.4 The influence of margin trading on the change in sentiment over time

A theoretical argument for efficient markets is based on arbitrage. Even if sentiment is correlated across irrational investors, arbitrageurs may take the other side of the demand and bring the market to equilibrium. However, according to behavioral finance, real world arbitrage is risky and limited. Figlewski (1979) indicates that it might take a very long time for noise traders to lose most of their money if arbitrageurs bear the fundamental risk in betting against them. Shiller (1984) and Campbell and Kyle (1987) focus on arbitrageurs' aversion to fundamental risk in discussing the effect of noise traders on stock market prices. They find that aversion to fundamental risk can by itself severely limit arbitrage, even when arbitrageurs have infinite horizons. Moreover, to arbitrageurs, sentiment itself is another source of risk (DeLong et al. 1990). Noise traders can create their own space. They may earn more expected returns from their own influence. Thaler (1999) claims that the precondition of an efficient market is that only rational investors will sell short when stock prices are higher than intrinsic values. In fact, both rational and irrational investors can trade by margin. If noise traders, instead of rational investors, dominate margin trading, they will have more ability to enhance their sentiment by margin buying, and thereby make change in the sentiment more pronounced. Thus, margin trading made by noise traders will make markets deviate even more from efficient market equilibrium. In this case, it is expected that the average growth trend of sentiment for stocks with ease of margin trading would be steeper than that for stocks without ease of margin trading. On the other hand, if margin trading is made mostly by rational investors, showing no lasting sentiment, the average growth trend of sentiment for stocks with ease of margin trading would not be salient.

Irrational investors tend to choose high-tech stocks as their main investing targets and they could further leverage their money through margin trading. Therefore, the difference of the average rate of change in sentiment between the groups with and without ease of margin trading for high-tech stocks is greater than that for non-high-tech stocks.

3.3 Methods

3.3.1 Measures

Investor sentiment may be investigated by using surveys. UBS/Callup conducts surveys to household investors. The University of Michigan consumer confidence index, also based on surveys, is used to reflect consumers' confidence, and is highly correlated with the UBS/Callup Index. The J. P. Morgan investor confidence index is another one to see investor sentiment. Besides, investor sentiment may be captured by observing patterns of market trades including mutual fund flows, closed-end fund discounts, volume of initial public offerings (IPOs), first-day returns on IPOs, and trading volume, etc.

The existing measures have some drawbacks. First, the newsletter survey should be considered only if the news could affect the beliefs of market traders (Brown and Cliff 2004). Second, the measures such as the average closed-end fund discount, volume of IPOs, the average first-day return on IPOs, and composite indices of the above (e.g., Lee et al.1991, Ritter 1991, Neal and Wheatley 1998, Brown and Cliff 2004, 2005, Baker and Wurgler 2006) target the entire market rather than individual stocks. Although they aggregate market data, most of them consist of only selected stocks, which restrict their representativeness of the entire market. Baker and Wurgler (2006) argue that different stocks are subject to different levels of sentiment because of either different shocks on sentiment-based demands or different arbitrage constraints. The intensity of investor sentiment may vary with different stocks. To conduct the analysis of sentiment based on individual stocks, it is necessary to find an appropriate stock-specific sentiment proxy.

Baker and Stein (2004) note that if short-selling is costly or forbidden, irrational investors will be active to trade and thus add liquidity only when they are optimistic and betting on rising stocks. On the other hand, when irrational investors are pessimistic, the short-selling constraint keeps them out of the market. High liquidity could be considered as a symptom that the market is dominated by irrational investors. Scheinkman and Xiong (2003) state that volume reveals underlying differences of opinions, which is in turn related to valuation levels when short selling is difficult. Therefore, market turnover is a simple and effective proxy for this concept. Although margin trading in the Taiwan security market has been going on for many years, the naked short selling is forbidden, and the requirements of the regular short selling are stricter than those of margin buying. This implies that short selling in Taiwan is costly. In this study, turnover ratio is used as a proxy for investor sentiment. Turnover ratio data were collected for individual stocks. However, levels of turnover ratio may partially reflect fundamentals of stocks. To remove such effects, at least partially, we use earnings per share (EPS) as a measure for fundamentals and incorporate EPS into the model as a control variable.

3.3.2 Latent growth modeling

Latent growth modeling (LGM) plays an important role in repeated-measure analysis over a limited occasions in large sample data (e.g., Singer and Willett 2003, p.9, Preacher et al. 2008, p.12). LGM requires that data on a focal variable be collected from individuals at multiple points in time (Chan 1998), and is a typical application of the structural equation modeling (SEM) (e.g., Chan 1998, Bauer 2003, Curran 2003, Mehta and Neal 2005). SEM is popular in psychology, management and marketing, but not in finance, although Titman and Wessels (1988), Maddala and Nimalendran (1995), and Chang et al. (2009) have notably used SEM in corporate finance. The sentiment involves both level and change (Brown and Cliff, 2004). It is formed through a process over time (Smidt 1968, Brown and Cliff 2005). As mentioned previously, analysis for the growth trajectory of investor sentiment can provide useful information for management, but it was not specifically addressed in the finance literature. Thus, the purpose of this study is to demonstrate the use of LGM for assessing the change in investor sentiment over time on individual stocks and predicting the patterns of longitudinal changes.

LGM can not only characterize intraindividual (within-stock) change over time but also examine interindividual (between-stock) difference by means of a random intercept and random slopes. The within-stock errors over time and the between-stock errors (representing random effects for the intercept and slopes) are conventionally referred to as the level-1 and level-2 errors, respectively.

The first-level submodel, describing individual change, controlling for EPS, is given by

$$Y_{it} = \beta_{0i} + \beta_{1i} TIME_t + \phi EPS_t + \varepsilon_{it},\ t = 1, 2, ..., T, \tag{3-1}$$

where T is the total number of time points, $TIME_t$ represents a particular time point t and serves as an explanatory variable, time points are usually equally spaced, set as 0, 1, …, T–1, Y_{it} is the level of turnover ratio for stock i at time t, and ε_{it} the corresponding error. β_{0i} and β_{1i} denote, respectively, the intercept (initial status) and slope (rate of change) of the linear growth trajectory of sentiment for stock i. They are random because stocks differ in their initial sentiment levels and linear trajectory. ϕ denotes the fixed regression parameter for the control variable EPS. The model in Equation (3-1) indicates Y_{it} can be depicted as a linear function of time, EPS, and ε_{it}. ε_{it} is called the first level error associated with stock i at time t (intrastock), reflects departure from the growth trajectory for stock i. ε_{it}'s are serially correlated for stock i. The structure of autocovariance of ε_{it}, assumed to be identical for all stocks, needs to be identified. Although AR(1) (the first-order autoregressive) may be the most commonly used one (Littell et al. 2006, p. 175), ARH(1) (heterogeneous AR(1)) is more appropriate since error variances and autocovariances may not be homogeneous. The error covariance matrix for $T = 4$ based on ARH(1) is given by

$$\mathbf{\Theta}_{\varepsilon} = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho & \sigma_{\varepsilon_2}^2 & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho^2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho & \sigma_{\varepsilon_3}^2 & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho^3 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho^2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho & \sigma_{\varepsilon_4}^2 \end{bmatrix}, \tag{3-2}$$

where ρ_k denotes the autocorrelation at lag k, k = 1, 2, 3, $\rho_k = \rho^k$, $k > 0$. When error variances are equal ($\sigma_{\varepsilon_1}^2 = \sigma_{\varepsilon_2}^2 = \sigma_{\varepsilon_3}^2 = \sigma_{\varepsilon_4}^2$), ARH(1) reduces to AR(1).

The second-level submodel, describing interstock differences in intrastock change, is given by

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \zeta_{0i}, \\ \beta_{1i} &= \gamma_{10} + \zeta_{1i}, \end{aligned} \tag{3-3}$$

where γ_{00} and γ_{10} denote, respectively, the means of β_{0i} and β_{1i}, and ζ_{0i} and ζ_{1i} the corresponding level-2 errors. It is assumed that ζ_{0i} and ζ_{1i} have a bivariate normal distribution with E(ζ_{0i}) = 0, E(ζ_{1i}) = 0, Var(ζ_{0i}) $=\sigma_{\zeta_0}^2$, Var(ζ_{1i}) = $\sigma_{\zeta_1}^2$, and Cov(ζ_{0i}, ζ_{1i}) $=\sigma_{\zeta_0\zeta_1}$, and ζ_{0i} and ζ_{1i} are uncorrelated with ε_{it}. $\sigma_{\zeta_0}^2$, $\sigma_{\zeta_1}^2$ and $\sigma_{\zeta_0\zeta_1}$ are the variances and covariance of the random effects reflecting interstock differences. A positive value of γ_{10} indicates that the average linear growth trend of sentiment across the entire market is increasing, reflecting an optimistic tendency. In contrast, a negative value of γ_{10} reflects that the market is pessimistic. If γ_{10} = 0, then the market sentiment is unchanged. The maximum likelihood method is usually used to make inference for the growth parameters γ_{00} and γ_{10} as well as the level-1 and level-2 error variances and covariances.

The submodel in Equation (3-3) is often called the unconditional model since explanatory characteristics of stocks that affect the pattern of change are not included. It is denoted by Model (A) in this study. Interstock differences in growth could be attributed to systematic characteristics if β_{0i} and β_{1i} are related to time-invariant predictors. For the case of one level-2 predictor ξ, we have the following conditional model:

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{01}\ \xi_i + \zeta_{0i}, \\ \beta_{1i} &= \gamma_{10} + \gamma_{11}\ \xi_i + \zeta_{1i}, \end{aligned} \tag{3-4}$$

where γ_{00} and γ_{01} represent, respectively, the level-2 intercept and slope of ξ with respect to the initial status, and γ_{10} and γ_{11} represent those with respect to the rate of sentiment change over time. ζ_{0i} and ζ_{1i} are level-2 errors, with the same assumptions as those seen in Equation (3-3). In particular, γ_{11} is the difference of the average rate of sentiment changes by a unit increase of ξ. The growth parameters in this case include γ_{00}, γ_{01}, γ_{10}, and γ_{11}.

In the Taiwan stock market, the permissible amount of margin trading for the component stocks of exchange-traded fund or MSCI (Morgan Stanley Capital International Inc.) Taiwan index is twice more than the permissible amount of margin trading for other stocks. Therefore, margin trading is easier to implement for the component stocks than others. The component stocks are classified as those with ease of margin trading (EMT), and the others are those without ease of margin trading (NEMT). In this study, three level-2 predictors, the industry type (with the levels of 'high-tech (HT)' and 'non-high-tech (NHT)'), stock size (with the unit of billion shares), and the ease of margin trading (with the levels of 'Yes (EMT)' and 'No (NEMT)'), are considered. Dummy variable D_1 is defined for the industry type by letting $D_1 = 1$ for HT and $D_1 = 0$ for NHT. Dummy variable D_2 is defined for the ease of margin trading by letting $D_2 = 1$ for EMT and $D_2 = 0$ for NEMT. Furthermore, the dummy representations for the groups of (1) high-tech stocks with ease of margin trading (HT-EMT), (2) high-tech stocks without ease of margin trading (HT-NEMT), (3) non-high-tech stocks with ease of margin trading (NHT-EMT), and (4) non-high-tech stocks without ease of margin trading (NHT-NEMT) are given by

Group	D_1	D_2	D_1D_2
1: HT-EMT	1	1	1
2: HT-NEMT	1	0	0
3: NHT-EMT	0	1	0
4: NHT-NEMT	0	0	0

Let models (B), (C), and (D) denote, respectively, those with one level-2 predictor by substituting D_1 (the industry type), stock size, and D_2 (the ease of margin trading) into ξ in Equation (3-4). Model (E) is the model with predictors D_1, D_2, and their cross-product, given by $\beta_{0i} = \gamma_{00} + \gamma_{01} D_{1i} + \gamma_{02} D_{2i} + \gamma_{03} D_{1i} D_{2i} + \zeta_{0i}$ and $\beta_{1i} = \gamma_{10} + \gamma_{11} D_{1i} + \gamma_{12} D_{2i} + \gamma_{13} D_{1i} D_{2i} + \zeta_{1i}$.

The first-level and the second-level submodels are simultaneously fitted and estimated by using SEM. Since the p-value of the chi-square test for assessing the model fit is always small for large samples, leading to the rejection of an adequate model, two criteria indicating acceptable model fit, 'CFI (the comparative fit index) ≥ 0.95' and 'SRMR (the standardized root mean square residual) ≤ 0.09', are recommended (e.g., Iacobucci 2010).

3.3.3 Data

Individual investors, who are regarded as noise traders, spend far less time on investment analysis and rely heavily on a different set of information. Individual investors tend to buy or sell stock in concert with each other (Black 1986). Higher individual trading is associated with higher liquidity commonality and faster movement of investor sentiment (Choel and Yang 2010). A market with a high proportion of individual investors should be the target for analyzing investor sentiment. The data used for demonstration come from the Taiwan stock market, which contains about 70% individual investors (See Appendix 3-A).

Appendix 3-B shows the J. P. Morgan investor confidence indices (denoted by J. P. ICI in this study) and the volume of IPOs, a commonly used sentiment indicator, from January, 2005 through December, 2008 in the Taiwan market. It appears that J. P. ICI and the volume of IPOs display inconsistent tendency for each year. In particular, there exists an obvious upward trend of the volume of IPOs from September, 2007 to December, 2007, and the peak occurs in December 2007. However, J. P. ICI is 109.9 in September 2007 and 104.5 in December 2007, showing a downward trend. There exists inconsistency between the former (being optimistic) and the latter (being pessimistic). LGM, a more sophisticated approach, will be used to further examine the growth trajectory of sentiment during this period. The monthly turnover data from September, 2007 to December, 2007 for 692 stocks in the Taiwan market, obtained from the Taiwan Economic Journal Data Bank (TEJDB), are used. The heterogeneity of sentiment trajectories discovered by using LGM can interpret the inconsistency, and the issue will be discussed in detail in Section 5.3.

According to *Taiwan Stock Exchange Corporation (TWSE) Key Points for Classifying and Adjusting Categories of Industries of Listed Companies*, announced by TWSE in 2007, 52 semiconductor stocks belong to the industries of wafer fabrication, IC design, packing and testing, materials, and equipment manufacturing. These industries constitute the high-tech industry in Taiwan (Hsinchu Science Park Administration 2008).

3.4 Results

The results by using Model (1-A) (the combination of Model (1) and Model (A)) through Model (1-E) (the combination of Model (1) and Model (E)) based on the sample of 692 stocks are summarized in Table 3-1. Fit indices indicate acceptable fit for all models. Moreover, the level-1 error variances and covariances based on the ARH(1) structure and the level-2 error variances and covariances are all significant. The significance of the level-2 error variances and covariances supports the variability of intercept and slope. The estimate of the average rate of change in sentiment resulting from Model (1-A) based on the entire sample (692 stocks) is significantly negative ($\hat{\gamma}_{10} = -1.7980$). Therefore, the average linear growth trend of sentiment for the entire market during the period from September 2007 to December 2007 is declining, reflecting the pessimistic mood of investors (see Figure 3-1 (1-A)).

With Model (1-B), the difference between the average rates of sentiment change for HT and NHT stocks ($\hat{\gamma}_{11} = 3.6353$) is highly significant. The average slopes for HT and NHT stocks are 1.5678 (= −2.0675 + 3.6353) and −2.0675, respectively, and are presented in Figure 3-1 (1-B). The former shows a significantly optimistic trend ($p < 0.01$), exhibiting the bandwagon effect during this period, and the latter a significantly pessimistic trend ($p < 0.01$). Figure 3-1 (1-B) also contrasts the sentiment levels of HT and NHT stocks. The average sentiment levels of HT stocks at individual time points are higher than the corresponding ones of NHT stocks. Their difference is an increasing function of time.

According to the results by Model (1-C), the increment of the average rate of change in sentiment is highly significant ($\hat{\gamma}_{11} = 0.2206$) for a unit increase of stock size, leading to a smoother slope. Figure 3-1 (1-C) shows that, although the average linear trends for the stocks whose sizes are below the first quartile (Q1 = 0.133 billion shares) and for those whose sizes are above the third quartile (Q3 = 0.5335 billion shares) are both negative, the former is steeper than the latter.

The results by Model (1-D) indicate that the difference of the average rates of change in sentiment between EMT and NEMT stocks is significant ($\hat{\gamma}_{11} = 2.1771$). The stocks with NEMT show a significantly downward trend ($\hat{\gamma}_{10} = -2.3632$) but those with EMT do not (Their average slope is −2.3632 + 2.1771 = −0.1861 ($p = 0.5870$)). Margin trading might be used more by rational investors than irrational investors, leading to non-significant sentiment change over time.

In Model (1-E), for high-tech stocks without ease of margin trading, the average linear sentiment trend is non-significant ($\hat{\gamma}_{10}+\hat{\gamma}_{11}$ = −2.4320 + 1.6625 = –0.7695, p = 0.3844); the average linear sentiment trend for high-tech stocks with ease of margin trading becomes significantly upward ($=\hat{\gamma}_{10}+\hat{\gamma}_{11}+\hat{\gamma}_{12}+\hat{\gamma}_{13}$ = −2.4320 + 1.6625 + 1.5469 + 2.9072 = 3.6846, $p <$ 0.01). Margin trading may be utilized mainly by noise traders for high-tech stocks, thereby causing a marked increase in the optimistic trend of sentiment (see also Figure 3-1 (1-E)). In addition, the average linear sentiment trend for NHT-EMT stocks is significantly different than that for NHT-NEMT stocks ($\hat{\gamma}_{12}$ = 1.5469). The former is –0.8851 (p = 0.0151), and the latter is –2.4320 ($p < 0.01$). Although the average sentiment trends are both negative, pessimism is alleviated by rational investors through margin trading. The pattern is similar to that of Figure 3-1 (1-D).

3.5 Discussion

The empirical findings support the predictions mentioned in Section 3.2. The average linear growth trend of investor sentiment for the entire Taiwan stock market during the study period is declining. However, growth trends of sentiment for individual stocks are diverse, depending on their characteristics. There exist greater decline in investor sentiment for stocks with smaller sizes. The stocks with NEMT show, on average, a downward trend. Moreover, high-tech stocks show optimistic atmosphere. The heterogeneity of the change in sentiment over time for different stocks leads to some important implications on arbitrage. Relevant issues for IPOs are also discussed in this section.

3.5.1 Arbitrage mechanisms and market intervention

The success of arbitrage is based on two conditions. First, rational investors are the main players in the market. Second, short selling is allowed (with a sufficient supply of shares to do so) so that rational investors can arbitrage away any deviation from fundamental values. Haruvy and Noussair (2006) conclude that the availability of short selling may eliminate the bubble-and-crash phenomenon and induce prices to track fundamentals. However, market traders have different degrees of rationality, and lack of short-selling for some stocks would constrain rational traders. The existence of noise trading and the restriction of arbitrage (e.g., prohibition of margin trading) lead to the occurrence of sentiment.

Investors usually get more transparent information about EMT stocks, which can help arbitrageurs judge their values with more confidence. Thus, the EMT stocks should be able to be traded more efficiently than NEMT stocks. However, if noise traders dominate margin trading for some stocks, they can magnify their sentiment by margin buying, and drive the change in sentiment over time even more severely. A possible explanation why the average sentiment trajectory for HT-EMT stocks is steeper than that for HT-NEMT stocks is that margin trading is utilized mainly by noise traders for high-tech stocks. Under this situation, it is risky for arbitrageurs to sell short, and arbitrage becomes limited. The margin-trading behaviors of noise traders will reduce market efficiency.

In summary, the function of arbitrage depends on the sentiment of noise traders and stock characteristics. Past studies argue that sentiment would make markets deviate from efficient equilibrium, and suggest that noise trades should be eliminated. However, since noise trading cannot be avoided, when sentiment is found to be over-optimistic or over-pessimistic (by examining its growth trajectory), policy makers should take necessary actions to make it “cool down”.

Regulations about margin trading, such as short sales restrictions, are hot topics discussed in financial markets. However, the role of these regulations on market valuation has long been debated. Academics studying efficient market theory usually state that short-selling is undertaken by rational arbitrageurs, and help markets to correct short-term deviations of stock prices from fundamentals (e.g. Diether et al., 2008). On the other hand, other people believe that trading on margin does not play a particularly helpful role in stabilizing the overall stock market (e.g. Lamont and Stein 2004). We have found some empirical evidence to support the view that in the case of the high-tech industry, ease of margin trading may enhance the accumulation of sentiment.

From the perspective of market monitoring, Stiglitz (1989) and Summers and Summers (1989) consider that a significant part of market volatility reflects "noise trading", instruments discouraging noise trading such as "transactions tax" should contribute to reductions in volatility and improve the functioning of speculative markets. Reductions in noise trading will increase market efficiency. On the other hand, Black (1986) claims that noise trading could contribute to the enhancement of liquidity, and many empirical works such as Glosten and Migrom (1985), Admati and Pleiderer (1988), Berkman and Eleswarapu (1998), Greene and Smart (1999), and Liang (1999) demonstrate this argument. To sum up, noise trading causes market volatility and liquidity simultaneously, and no matter what rules to prevent noise trading are enforced, the market cannot be both efficient and liquid.

This study finds that not all noise trading would cause high liquidity. Only when the sentiment trend is optimistic and increasing would the market liquidity be increased. In contrast, a pessimistic trend would diminish market liquidity. Both optimistic and pessimistic sentiment trends would be the target of policy instruments for regulating market trading. The continuities of sentiment trends influence the practice of efficient markets and the stability of financial markets regardless of whether the sentiment trend is optimistic or pessimistic. The former may cause bubble-and-crash and the latter would increase the liquidity risk. Therefore, in addition to facilitating arbitrage mechanisms, governments should take necessary actions to avoid over-optimistic or over-pessimistic sentiment spreading in the markets. In order to stimulate markets, government could encourage institutional investors to participate in trade and make markets more active; likewise governments could take steps to stave off liquidity risk caused by over-pessimistic sentiment. For example, monetary policy would be influential on investor sentiment in bear market period (Kurov 2010). The influence of margin purchases and short selling on sentiment trends would be instructive to governments. If margin trading magnifies the over-optimistic sentiment, governments should restrict short selling or margin purchases. On the other hand, if the sentiment trend is caused by the limits on arbitrage, governments should relax the constraints. The strategies may be applied to the entire market, specific industries, or individual stocks, depending on the situation.

3.5.2 Relevant issues for IPOs

The impact of investor sentiment is acute in hot markets. Kaustia and Knüpfer (2008) indicate that investor sentiment drives IPO demand. When there exist overoptimistic investors and short-selling constraints, both the IPO price and the aftermarket price will be driven above the stock's fundamental value (Derrien 2005, Cornelli et al. 2006, Ljungqvist et al. 2006, Dorn 2009). Security issuers and security underwriters will choose the timing with optimistic sentiment to issue IPOs. On the contrary, pessimistic sentiment creates no stimuli for issuing IPOs. Since IPOs represent only the stocks with optimistic sentiment, the volume of IPOs should not be used as an indicator of the sentiment for the entire market.

As mentioned previously, J. P. ICI reflect a pessimistic trend, but the volume of IPOs gives an optimistic signal from September, 2007 to December, 2007. According to the results by Model (1-A), the sentiment trend for the entire market is decreasing. The result is consistent with that indicated by J. P. ICI, but not with that indicated by the volume of IPOs. Although the volume of IPOs for the entire market is increasing during the period, reflecting the optimistic trend of sentiment, they are still the "minority" of the market. Interestingly, the increasing volume of IPOs and the optimistic trend of sentiment for the high-tech industry obtained by using Model (1-B) agree with each other. Since optimistic sentiment may not exists for all industries, it is likely that using the volume of IPOs as a sentiment measure for the entire market will lead to a different conclusion about the growth trajectory of sentiment.

Regarding IPO practice, security underwriters in both Korea and Taiwan provide withdrawal options (put-back option) for individual investors as an incentive to invest. Investors have a withdrawal option to return IPO shares to the issuers if the IPO aftermarket price drops significantly. Investors' option to withdraw reduces the information asymmetry between informed investors and uninformed investors but increases the firm-commitment underwriting risk (Lin et al. 2010), especially when investor sentiment declines. The market response in Korea is more pronounced during high sentiment periods, particularly for small-size, young, highly volatile, and low-profit stocks (Kim and Byun 2010). Although the regulatory change has given an incentive to lower underwriting fees, underwriters still charge higher fees for small-size companies issuing IPOs (Ahn et al. 2007). LGM incorporating stock characteristics can help better assess growth trajectories of sentiment, based on which the timing, the prices, and underwriting fees of issuing IPOs can be better determined.

3.6 Conclusion

Assessing the sentiment change over time in stock markets is important. Although measures of investor sentiment have been proposed based on market aggregate indicators, they may not well reflect the sentiment for the entire market. Irrational investors have different expectations for different stocks. Their sentiment-based expectations may vary with stock characteristics. In this study, we have demonstrated using latent growth modeling to examine interstock differences in intrastock sentiment change over time. The approach, aiming at the entire market, is on the basis of individual stocks. Analysis is conducted simultaneously for all individual stocks by using the turnover data, thereby making the empirical work more sophisticated. The results for the market in Taiwan during the period from September 2007 to December 2007 indicate that the average linear growth trend of sentiment depends on stock characteristics, including the industry type (high-tech / non-high-tech) and stock size, and whether margin trading is easy to implement. Although the volume of IPOs can reflect investor sentiment, it may apply only for a specific group (such as the high-tech industry demonstrated in this study) and should not be used as an indicator of the sentiment for the entire market.

Based on the results by LGM, policy implications about arbitrage mechanisms and market intervention and some relevant issues regarding IPOs have been discussed. There exist some directions for future research. First, the level-1 linear growth model can be extended to a broader class of polynomial functions for a longer period. Besides, more industry types and more stock characteristics such as capitalization, profitability, and company growth can be introduced in the level-2 model. These characteristics can be combined with the ease of margin trading (EMT/NEMT) to see their interactive influence on the sentiment change over time. Second, measures for investor sentiment and its growth trajectory should be developed, with satisfactory reliability and validity. Finally, market inefficiency resulting from sentiment needs to be further empirically examined. Kim and Byun (2010) find that investor sentiment affects the valuation of stocks in the Korea market. It is interesting to see the effects of the change in sentiment over time on stock returns and how the effects are moderated by stock characteristics. SEM may be used again for this purpose.

4. On the Measurement of the Change in Investor Sentiment over Time for Individual Stocks

4.1 Introduction

How the stock market bubbles start is an interesting issue in behavior finance. Examples of bubble stemming from optimistic investor sentiment trend might include 1987 Black Monday, internet bubble and the ensuing Nasdaq and telecom crashes in the late 1990s, and the U.S. housing bubble, leading to the global financial turmoil in 2008 (Saxton, 2008). Shiller (1990) conducts a survey, and finds that the main reason of the market crash during the week of October 19, 1987 is that it is overpriced. Arbitrage did not work to drive price to the equilibrium. Galbraith (1954) and Kindleberger (1978) emphasizes the irrational element inducing the public to invest in the bull “over-heating” market. The eagerness to buy stocks is driven by irrational euphoria among individual investors. Those aforementioned well-known stock market events show that the continuously optimistic sentiments trend resulting from general rise in prosperity cause the rise in stock prices, followed by severe crashes.

Noise traders are the irrational investors in financial market. They are not based on fundamental information (Shiller, 1984). Moods and emotions interact with cognitive processes when people make decisions. Noise traders believe they can forecast future returns of risky assets. During good time, they may still be optimistic even though the stocks they hold are fundamentally poor because they are affected by “good atmosphere”. During bad time, on the other hand, they may become pessimistic even though the stocks they hold are fundamentally sound because their confidence is shaken by “bad atmosphere”. Their sentiment, called investor sentiment (IS), is a belief about future asset values that is not justified by the facts (Baker and Wurgler, 2007). The belief leads them to irrationally buy or sell more stocks, and their trading behavior are potential to affect stock price (Barber et al., 2009).

Investing could be considered as a social activity. Human interactions spread moods and emotions, which cause uniformity in financial decision-making (Prechter, 1999). The spread of moods is similar to the spread of diseases, thus, investors’ behavior would be influenced by social movements (Shiller, 1984). Noise traders behave socially and follow each others’ mistakes by interact with others. Mutual reinforcement through interaction with others will finally stimulate IS to move in the same way for a period of time. The presence of noise traders in financial markets then causes prices and risk levels to deviate from expected levels even though rational traders exist, and can limit the willingness of arbitragers to bet against noise traders. For example, increasing sentiment is associated with social mood like optimism and hope which lead to overconfidence and euphoria. Optimistic investors will boost market sentiment continuously for a period of time, thereby making the market more and more inefficient, and lead to an eventual market crash.

Dealing with the IS trend rather than simply a snapshot on the process is appealing. The dynamic process by which social movements work takes time. That is, IS trend is formed through a process over time (e.g., Smidt 1968, Brown and Cliff 2005). Observing IS levels individually cannot capture IS trend over a period of time. Moreover, the lasting optimistic or pessimistic IS trend implies that people do not deviate from rationality randomly, but in the same way, demonstrating the existence of social movements. Persistent optimism may push prices a long way away from fundamentals and produce anomaly, followed by bubbles and crashes. Thus, IS trend should receive more attention than IS level. Although a number of researchers such as Grossman and Stiglitz (1980), Black (1986), DeLong et al. (1990), Campbell and Kyle (1993), Barberis et al. (1998), Daniel et al. (1998), and Hong and Stein

(1999) have formally modeled the role of sentiment and conducted empirical studies for sentiment, the dynamic change over time has rarely been discussed in the literature.

Investor sentiment is a latent construct that cannot be measured directly. Market indicators used as sentiment proxies have been proposed in the financial literature. Some studies used the principle component analysis for the market indicators to obtain a composite index (e.g., Brown and Cliff 2004, Baker and Wurgler 2007, Chen et al. 2010). However, how IS is appropriately measured (with acceptable reliability and validity) is not well addressed.

Confirmatory factor analysis (CFA), the measurement model in structural equation modeling (SEM), is useful for assessing reliability and validity of measures of a latent construct. SEM has been widely applied in behavior studies, but is rarely seen in finance. Titman and Wessels (1988), Maddala and Nimalendran (1995), and Chang et al. (2009) used the approach similar to SEM in corporate finance.

An important issue in investor sentiment is the measurement (Baker and Wurgler, 2007), and is the focus of this research. It differs from previous research in that IS trend, rather than IS level, is measured and the measurement is based on individual stocks rather than the entire market. Construct validity and reliability based on multiple indicators are assessed for both of the IS trend and the IS levels. The proxies for IS are reviewed and summarized. The case by case regression is used to analyze the IS trend based on each of the IS indicators identified. The output resulting from the case by case regression is then input for CFA. The common factor of indicator-based IS trends is extracted and its model fit is assessed. The usefulness of the results obtained is illustrated with an application to investigate the relationship between IS trend and stock return. The situations when IS trend can influence return are examined. Relevant policy implications are discussed..

4.2 Literature Review

4.2.1 The sentiment indicators in empirical studies

In empirical studies, there are some indices used for measuring IS. IS may be reflected by those based on surveys such as the UBS/Callup index and The University of Michigan consumer confidence index. On the other hand, IS may be captured by price-based variables such as mutual fund flows (Brown et al. 2002), closed-end fund discounts (CEFD) (Zweig, 1973; Lee et al., 1991; Neal and Wheatley, 1998; Lowry, 2003), volume of initial public offerings (IPOs) and first-day returns on IPOs (Stigler, 1964; Ritter, 1991; Baker and Wurgler, 2000). Liquidity-based indicators like trading volume and market turnover (Campbell et al., 1994; Cooper, 1999; Gervais et al., 2001; Chordia and Swaminathan, 2000, Lee and Swaminathan, 2000, Baker and Stein, 2002) are also widely used. A few studies use the principle component analysis for the aforementioned aggregate market indicators to obtain a composite index for IS (e.g., Brown and Cliff, 2004; Baker and Wurgler, 2006; Baker and Wurgler, 2007; Chen et al. 2010).

Although aggregate market indicators integrate the data for individual stocks, most of them consist of only selected stocks and their representativeness of the entire market is questionable. According to previous studies, investor sentiment vary with stock characteristics. Baker and Wurgler (2007) argue that different stocks are subject to different levels of sentiment because of different shock characteristics or different arbitrage constraints. Ofek and Richardson (2003) report that internet stocks are owned relatively more often by individual investors, who are regarded as noise traders in the literature. Individual investors might merely snap up stocks in the internet or other high-tech industries (Dorn, 2009). Thus, sentiments stemming from high-tech stocks are relatively higher. Besides, the stocks with stronger individual concentrations may be traded more irrationally. Kumar and Lee (2005) find that smaller firms, lower priced firms, or firms with lower institutional ownership, are associated with stronger noise trading activities. The sparser information available about smaller firms will make the market less efficient (Hirshleifer, 2001). Stocks with smaller sizes or stocks in high-tech industry tend to be more sensitive to waves of investor sentiment. Therefore, the indicators reflecting stock characteristics should be able to capture the sentiment trend more precisely than aggregate market indicators and the analysis of sentiment will be conducted based on individual stocks.

4.2.2 Measurement in psychometrics

IS reflects the moods and emotions of noise traders. Measurement error of such a construct is a severe problem throughout the social sciences (Peter, 1981). A number of steps have been suggested in the measure development process (Churchill, 1979; Gerbing and Anderson, 1988). These steps, such as reliability and validity test, emphasize that traversing the distance from the conceptual to the operational requires a systematic process. The operational measure of a latent construct is indirect (Nunnally, 1978). As the abstractness of a construct increases, the distance between the conceptual and the operational definitions increases. The fundamental objective in measurement is to produce observed scores which approximate true level of investor sentiment as closely as possible. The quality of measures depends on the evidence supporting their goodness, which takes the form of reliability or validity index (Churchill, 1979). Reliability and validity denote the accuracy or precision of a measuring instrument, and how well it measures that it purports to measure (Kerlinger, 1986). Following the tradition of psychometrics for the measuring latent constructs (e.g., Guilford, 1954; Nunnally, 1978; Churchill 1979), it is much better to use multi-item measures instead of single-item measures since no single item could provide a perfect representation of a concept. Multiple-item measures are inherently more reliable because they enable computation of correlations between items. The positive correlation and high average correlation (i.e., a high coefficient alpha) indicate the internal consistency of all the items in representing the presumed underlying construct. Besides, multiple-item measure captures more information than that can be provided by a single-item measure. Multiple-item measure is more likely to tap all facets of the construct of interest (Baumgartner and Homburg, 1996, p. 143) to capture the domain of interest to achieve content validity (Hinkin, 1999), and provides more response categories than the single-item measure. The more abstract is a construct, the more indicators are needed.

There exist two main concerns for assessed IS in psychometrics. First, most studies used single indirect indicator rather than multiple indicators. Second, for those studies taking into account multiple indicators, principle component factor analysis is used to form a composite measure of IS. However, it is only a part of preliminary stages in measure development, and provides preliminary information about unidimensionality. It should be followed by confirmatory factor analysis (CFA) (Hinkin, 1998). CFA could test reliability, validity, specified models of the relationship between items and factors, as well as the overall index of fit between the proposed model and the data. In order to highlight the heterogeneity of stocks and follow the rules of psychometrics, sentiment indicators reflecting stock characteristics are employed in this study.

4.3 Methods

4.3.1 IS indicators

Surveys and market indicators (such as dividend premium, CEFD, the volume of IPOs, first-day returns on IPOs, the equity share in new issues, and market turnover) are frequently used for measuring IS, but are mostly based on the data for the entire market. Since, as we have mentioned, IS can vary with stock characteristics, analysis of IS based on individual stocks can lead to more sophisticated results. However, from the aspects of IS for individual stocks, it is costly to obtain by surveys. On the other hand, most market indicators are unavailable for individual stocks. Even when they are available, the data may not exist for the required time period. For example, the availability of IPO and the dividend premium for individual stocks depends on time. In addition, new equity issues, the number of IPOs and CEFD appear for only selected stocks. Since the availability of market turnover is easy to achieve for each stock, it can serve as an IS indicator.

Market liquidity could be an IS indicators if short-selling is costly or forbidden (Baker and Stein 2004). Irrational investors will be active to trade and thus add liquidity only when they are optimistic and betting on rising stocks. On the contrary, when the irrational investors are pessimistic, the short-selling constraint keeps them out of market. Hence, high liquidity could be considered as a symptom that the market is dominated by irrational investors. Scheinkman and Xiong (2003) mention that liquidity reveals the common belief, which are in turn related to valuation levels when short-selling is difficult. In Taiwan securities market, naked short selling is forbidden, and the initial short selling margin ratio is 190%. This implies that short selling in Taiwan is costly. Market liquidity is an appropriate IS indicator in normal circumstance. Therefore, liquidity is a simple and effective proxy for IS.

The turnover rate of trading volume and the turnover rate of trading value are two commonly used liquidity indicators. Because noise traders trade more actively (with greater number of transactions), the number of transactions is additionally taken into account to avoid the misinterpretation of high turnover rate caused by little block trades (i.e., huge volume of single order). More specifically, the number of transactions / outstanding shares, called the turnover rate of transaction frequency in this study, is the third IS indicator, introduced in this study. Note that turnover ratios may partially reflect fundamentals of stocks. To remove such influences, at least partially, we use earnings per share (EPS) as a measure for fundamentals and use the residuals from the regression of the three market turnover indices (turnover rates of trading volume, trading value, and transaction frequency) on EPS as our IS indicators.

Only a few potential noisy proxies, such as, market price to book value ratio (e.g., Baker and Wurgler, 2007; Brennan and Wang, 2010; Hirshleifer, 2001) and short-selling turnover ratio (Chen, 2010) have been mentioned in previous literatures. However, only if noise traders are the main players and there exists limit of arbitrage, could stock price be influenced by sentiment. In most cases, there are both noise traders and rational investors in the stock market. Thus market price to book value ratio (PBR) should not be used to measure sentiment in most of time. Similarly, unless the market is occupied by noise traders and the margin trading is taken by irrational investors, short-selling turnover ratio (SSTR) should not be used to measure sentiment. Their limitations will be further discussed later in Section 4.4.2.

4.3.2 Measurement of IS trend

IS trend for an individual stock means the change in IS level over time. For each stock during the time period, the case by case regression (OLS) (Bollen *and Curran, 2006, Sec. 2.4*) with time as the independent variable and one IS indicator as the dependent variable is used to assess its IS trend by the estimated regression coefficient (slope). A positive value of the regression coefficient indicates that the IS linear trend for an individual stock is increasing, reflecting an optimistic sentiment. In contrast, a negative value of the regression coefficient reflects that the sentiment is pessimistic. If the coefficient is equal to zero, then the sentiment is unchanged.

Alternatively, IS trend could be assessed by the relative increment between the initial IS level and the final IS level. However, the advantage of the case by case approach is that it takes into account all IS levels during the period and the estimated regression coefficient for each individual is unbiased.

The conventional principal component approach to obtain factor scores uses different weights for the original indicators. They are data-specific and cannot be replicated across studies (Hair et al., 2010, p. 126-128). To improve, the scores of IS trend are obtained by computing the summated scores or the mean scores associated with its three indicators (Hair et al., 2010, p. 126-128). Another reason to use the summated scores is that reliability for a construct is always defined on the sum of indicators used to measure the construct (Lord and Novick 1968, Chap. 9). The slopes resulting from the case-by-case approach (OLS) are first scaled to remove the effects of different units before summated or mean scores are calculated.

4.3.3 Confirmatory factor analysis

In social sciences, it is frequently of interest to examine relationships for latent constructs. However, a latent construct cannot be directly measured, but can be represented by one or more observable indicators. CFA is a tool for testing how well indicators represent their underlying constructs (Hair et al., 2010, p. 693). The use of CFA should be based on measurement theory. CFA provides information for assessing reliability and construct validity. Higher reliability indicates lower measurement error. To adequately capture the construct domain, multiple indicators should be used instead of only a single indicator. Moreover, the use of multiple indicators can help avoid identification problems (Hair et al., 2010, p. 698). CFA provides various indices for assessing model fit.

A model for measuring constructs, taking into account measurement errors, is called a measurement model, and is formulated by using CFA. Let x_1, x_2, … x_J denote the observed scores of J indicators for measuring a construct, denoted by F. Centered by its mean μ_{x_i} and based on classical test theory (CTT), x_i can be formulated as (e.g., Joreskog, 1974; McDonald, 1999, p. 78; Reuterberg and Gustafsson, 1992) :

$$x_i - \mu_{x_i} = \lambda_i F + \delta_i,\ i = 1,\ ...,\ J, \tag{4-1}$$

where λ_i denotes the loading of x_i on F and δ_i the measurement error associated with x_i. It is assumed that the underlying latent construct F is uncorrelated with δ_i ($\sigma_{F\delta_i} = 0$), δ_i are uncorrelated ($\sigma_{\delta_i\delta_j} = 0,\ i \neq j$), and $\sigma_F^2 = 1$. It follows that λ_i is the covariance σ_{x_iF} of x_i and F.

Reliability of the summated scores $X = \sum_{i=1}^{J} x_i$, called composite reliability (denoted by CR) is the proportion of the variance of the summated observed indicators explained by the latent construct. It is given, based on Model (1), by (e.g., Lord and Novick 1968, Chap. 9, Fornell and Larcker 1981)

$$CR = \frac{(\sum_{i=1}^{J} \lambda_i)^2}{(\sum_{i=1}^{J} \lambda_i)^2 + \sum_{i=1}^{J} \sigma_{\delta_i}^2}. \tag{4-2}$$

Coefficient alpha (Cronbach 1951) is another commonly used reliability index for a construct. It is defined as

$$\alpha = \frac{J}{J-1}\left(1 - \frac{\sum_{i=1}^{J} \sigma_{x_i}^2}{\sigma_X^2}\right), \tag{4-3}$$

where $\sigma_{x_i}^2$ and σ_X^2 denote, respectively, the variance of the *i*th indicator and the variance of their sum.

CR can be estimated using CFA by

$$\hat{CR} = \frac{(\sum_{i=1}^{J} \hat{\lambda}_i)^2}{(\sum_{i=1}^{J} \hat{\lambda}_i)^2 + \sum_{i=1}^{J} \hat{\sigma}_{\delta_i}^2} \tag{4-4}$$

where $\hat{\lambda}_i$ are estimated loadings and $\hat{\sigma}_{\delta_i}^2$ are estimated error variances. Higher reliability is associated with lower measurement error. Unless the reliability is 100 percent (i.e., no measurement error), the true construct relationship will always be underestimated (Hair et al. 2010, p. 637). While a reliability of 0.7 should serve as an absolute minimum for newly developed measures, a reliability considerably higher than 0.7 is recommended in applied research (Hinkin, 1998).

Internal consistency means the degree of interrelatedness (intercorrelation) among indicators. Homogeneity means unidimensionality, implying that indicators all reflect the same unique construct. That is, they can be explained by only one underlying construct (Hair et al. 2010, p.696). When there are two or more constructs, no cross-loading is allowed. Each indicator loads only on its underlying construct. Coefficient α is a function of internal consistency and the number indicators (Cortina, 1993; Osburn, 2000). Since high α may be obtained when the number of indicators is large for those with low internal consistency or even with multi-dimensions (Cortina, 1993), reliability should be assessed after homogeneity has been established (Gerbing and Anderson, 1988).

When there is only one construct, unidimensionality can be assessed by using the principal component factor analysis, followed by CFA (Hair et al. 2010, p. 125, 696). The former is to see if there exists only one common factor, usually based on the 'eigenvalue -greater-than-one' criterion, and the latter is to assess convergent validity. Convergent validity reflects the extent to which indicators actually represent the underlying construct they are designed to measure (Hair et al., 2010, p. 689). A simple way to assess convergent validity is to test for significance of estimated loadings (Anderson and Gerbing, 1988) using the *t*-statistic (= the estimated loading / standard error of the estimated loading), which is asymptotically normal. If the estimated loadings are all significantly non-zero ($|t| > 1.96$), then convergent validity is supported. Unidimensionality is achieved if only one factor is identified by the principal component factor analysis and convergent validity is supported by CFA. Moreover, CFA provides various indices for assessing model fit. Two criteria indicating acceptable model fit, 'CFI (the comparative fit index) $\geq$ 0.95' and 'SRMR (the standardized root mean square residual) $\leq$ 0.09', are recommended (e.g., Iacobucci, 2010).

Once reliability and validities are all supported, subsequent analysis involving the construct can be conducted, using the summated scores as the data for the construct. Standardized data should be used for those with different measurement units.

4.3.4 Data

The data used for illustration are collected from the Taiwan stock market from September 2007 to December 2007, composed of 692 stocks. Monthly transaction data for individual stocks are obtained from the TEJDB (Taiwan Economic Journal Data Bank). As with several emerging stock markets, the Taiwanese stock market had historically set several limitations on foreign investment. For instance, in 1990, the government required documents should be attached, before investment, for the securities Competent Authority and Central Bank of Taiwan to review and approve. The level of foreigner juridical ownership is limited, therefore in terms of trading share or holding share, the main players in Taiwan stock market were individuals. Individual investors, who are regarded as noise traders (Black, 1986), spend far less time on investment analysis and rely heavily on a different set of information. Such investors tend to buy or sell stock in concert with each other. Barber et. al, (2006) analyzed trading records in Taiwan stock market during 1995 to 1999, individual investors in Taiwan may trade more actively.

However, since September 30, 2003, the investment rules for foreign investors have been changed gradually. First, for facilitate internationalization, the aforementioned review and approve system was changed to registration system. After that, foreign investors can invest in the stock market after simply registering with the Taiwan Stock Exchange Corporation (TWSE) instead of be reviewed and approved by securities competent authority and central bank of Taiwan. In addition, in 2004, the US SEC approved TWSE as a “designated offshore securities market”, which assisted listed companies in raising funds from overseas. Therefore, the level of juridical ownership in Taiwan has accelerated to more than 50% and the percentage of trading value has increased from 20% to 30%, indicating the structure of investors in Taiwan stock market has altered.

There are two reasons to choose the period from September 2007 to December 2007. First, in the aspect of arbitrage cost, TWSE launched a centralized securities borrowing and lending (SBL) system in June 2003 to meet the needs of institutional investors. Starting July 2007, securities firms and securities finance companies are allowed to conduct SBL business acting as principal. Investors thus have alternative channels, not only borrowing from the TWSE SBL system but also from the aforementioned institutions. In addition, the tax rate of SBL trading was decreased on August 20, 2007. Thus, since September 2007, institutional investors could make margin trading with lower cost and various channels to borrow securities. Second, from the perspective of arbitrage efficiency, if the level of ownership of rational investors is higher than noise trader, then the arbitrage can work easily, through margin trading by rational investors. The level of judicial ownership reached to 59.74% in 2007, which is close to the highest record (60.16%) happening in financial crisis in 2008. In addition, if sentiment trading could be found in the period with low ratio of individual investors (67.06%), only next to the lowest record - 64% in financial crisis in 2008. We can infer that sentiment trading exists in other periods of time with higher ratio of individual investors.

4.4 Results

4.4.1 Trend reliability and validity

The correlations among the IS indicators are highly correlated, providing an overall clue of internal consistency between indicators (Table 4-1). Table 4-2 presents the results of principal component factor analysis. The percentage of variation explained in first factor are between 94.08% to 97.45% for each time period, implying unidimensionalilty. Table 4-3 indicates that coefficient α and CR both exceed 0.7 for each time period, indicating satisfactory reliability of the IS levels. All indicators load significantly on their underlying construct–investor sentiment, demonstrating convergent validity of the IS levels.

The correlations among the trend of the three IS indicators are highly correlated, providing an overall clue of internal consistency between indicators (Table 4-4). Results of the principal component factor analysis for IS trend, which assessed by case by case regression, are presented in Table 4-5, the percentage of variation explained in first factor is 95.27% in IS trend, implying unidimensionalilty. Results of the reliability and convergent validity assessment for the IS trends based on the three indicators are reported in Table 4-6. Coefficient α and CR of trends are all greater than 0.7, indicating that the reliabilities are acceptable. Moreover, the factor loadings of the three IS trajectories on their common factor are all significant. Therefore, the convergent validity of the IS trend is achieved.

According to the output from the case by case regression approach, the mean trend of the three IS indicators across all individual stocks is −4.4358, a value showing pessimistic trend. It is consistent with the results of the survey for investor sentiment in Taiwan conducted by the Shih Hsin university Department of Finance (2011). The sentiment index drops form −2.13 in October 2007 to −104.72 in December 2007.

4.4.2 Issues of PBR and SSTR

Table 4-7 shows that, based on CFA, the trend of PBR and SSTR do not exhibit satisfactory convergent validity, indicating that these two indicators do not represent sentiment. Thus these two indicators are not suitable for observing behavior of noise traders. Table 4-8 shows that, also based on CFA, the correlation coefficient of PBR trend, SSTR trend, and IS trend are not significiant, showing that these 3 factors that uncorrelated, demonstrating discriminant validity.

Noise trader theory indicates that, when arbitrage is limited, the IS trend can affect stock return. If the market is dominated by noise traders, leading to arbitrage limit, PBR and SSTR could be used as proper sentiment proxies. The ratio of judicial ownership (59.74%) in the Taiwan market may not meet the requirement of the dominance by noise trades, and therefore PBR and SSTR are not internally consistent with other IS indicators. Previous studies show stocks in high-tech industries or stocks with high individual holdings tend to have bigger waves of investor sentiment. In Taiwan, the optoelectronic industry is deemed as a high-tech industry (Hsinchu Science Park Administration, 2009), and is mainly owned by individuals (with 64.66% individual holdings). Thus, the stocks in this industry are used in this study. As shown in Table 4-9, the model fit (CFI = 0.96, SRMR = 0.0295) and reliability indices (coefficient alpha = 0.69, CR = 0.85) are acceptable based on the five indicators. However, the factor loading associated with the IS trend based on SSTR is not significant ($t = -0.6489 < 1.96$), not supporting convergent validity, and therefore it should be removed. The reliability indices based on the rest of the four indicators are higher (coefficient alpha = 0.72, CR = 0.92), and convergent validity is achieved ($t > 1.96$ for each of the indicators). The model with the four indicators for IS, coupled with SSTR as a control, also show an adequate fit (CFI = 0.96, SRMR = 0.0295). Accordingly, for the situation when stocks belong to the high-tech industry and are mainly owned by individuals, PBR (but not SSTR) could be used as an indicator for measuring the IS trend. It is noteworthy that neither PBR nor SSTR should be used for the entire market.

4.5 Applications

This section discusses the influence of the heterogeneity of stocks on arbitrage, mispricings and market bubbles, and the timing of market intervention. Efficient market theory indicates that arbitrage can work all the time. On the contrary, Shleifer and Vishny (1997) mention that both the absolute and the relative values of stocks are much harder to calculate than those of fixed income securieties. Therefore, arbitrage opportunities are harder to identify in stock markets than in bond or foreign exchange markets. Limit of arbitrage is often seen in stock market. In fact, the previous studies show that investor sentiment varies with stock characteristics. Thus, the effectiveness of arbitrage depends on stock characteristics. Illustration, related issues, and policy implications are addressed in this section.

4.5.1 IS trend and effectiveness of arbitrage

IS Trend could be considered as the evidence demonstrating the existence of social dynamics. When the noise traders are the main players, arbitrage would be limited. Thus, for those stocks with mainly noise traders, IS trend could influence return. The role of margin trading varies with stock characteristics. While margin trading is made by noise traders, arbitrage mechanism will strengthen the influence of sentiment trend on return. On the contrary, if margin trading is made by rational investors, who own more stocks than individuals, arbitrage is easier to work. In the Taiwan stock market, judical investors prefer the component stocks in the exchange-traded fund or the MSCI (Morgan Stanley Capital International Inc.) Taiwan index. The permissible amount of margin trading for those stocks are twice more than the permissible amount of margin trading for other stocks. Thus the component stocks are considered as those with ease of margin trading (EMT), and the others are considered as those without ease of margin trading (NEMT).

In order to illustrate the function of arbitrage mechanisms in different industries, Optoelectronics stocks are considered as stocks mainly owned by noise traders (NT), and financial and insurance stocks, with 64% judicial holdings, are considered as stocks mainly owned by rational traders (RT). The regression is used to test the influence of sentiment trend on return for different types of stocks. Dummy variable E_1 is defined by letting $E_1 = 1$ for NT and $E_1 = 0$ for RT. Dummy variable E_2 is defined by letting $E_2 = 1$ for the category with the ease of margin trading (EMT) and $E_2 = 0$ for the category without the ease of margin trading (NEMT). The dummy representations for the four types of stocks are given as follows:

Type	E_1	E_2	E_1E_2
NT-NEMT	1	0	0
NT-EMT	1	1	1
RT-NEMT	0	0	0
RT-EMT	0	1	0

Table 4-10 presents the results of the regression analysis. From the overall perspective, the results indicate that IS trend can positively account for stocks return (regression coefficient =0.0206 with $p < 0.01$). It means that limit of arbitrage exist in entire market. From the perspective of high noise holdings, the result shows that, for the type of NT, sentiment trend positively influence return (regression coefficient for NT = 0.0367 with $p < .0001$), and, for the type of RT, the influence of sentiment trend on return is not significant (regression coefficient for RT = –0.0032 with p =0.8757). It appears that arbitrage mechanism works (i.e., can eliminate the influence of IS trend on return) for RT stocks only. In other words, arbitrage limit exists for NT stocks. Moreover, although the IS trend affects return for both types of

NT-EMT (regression coefficient = 0.1211 with $p < .01$) and NT-NEMT (regression coefficient = 0.0405 with $p < .05$), the influence of the former is greater (the difference of the regression coefficients is 0.0806, $p < .10$). It is noteworthy that arbitrage mechanism, ineffective for stocks with mainly noise traders, can even reinforce the influence of IS trend on return. As a result, whether arbitrage mechanism can effectively function depends on the stock characteristics.

4.5.2 Discussion

Noise trader theory points out that change of IS and limit to arbitrage will cause mispricing. Baker and Wurgler (2007) show that, IS could affect current market return and predict subsequent market returns. Thus, mispricing factors must be included in asset pricing (Brennan and Wang, 2010). However, in asset pricing practice, there exist limitations for current IS proxies. For example, it is costly to conduct IS survey for individual stocks. The appropriateness of existing market indicators need to be examined by using psychometric approaches. Our research results facilitate asset pricing for practitioners..

From the perspective of individual stock pricing, the investor sentiment, successfully measured by the indicators developed in this study, can be used as a mispricing factor. Limited arbitrage is required when using this mispricing factor. The relationship between IS trend and return can be used to help identify the stocks with limited arbitrage potential. From the viewpoint of market bubble formation, an empirical study conducted by Baker and Wurgler (2007) indicates that, if the current IS is high, then the current returns are high but subsequent returns are low. It seems that even the current market price deviates from fundamental value due to IS, market mechanism will draw it back at the subsequent period. If market mechanism can function, market bubble and crash would not happen. However, famous market bubbles always start from continuously optimistic sentiment, leading to irrational rise of market price. Thus, the positive relationship between IS trend and stock return discovered in this study would facilitate analyzing the formulation process of market bubble.

Efficient Market theory dictates that trade price is determined by free commercial activity by the two side parties, thus reflecting fundamental value. So, on the premise of the above-mentioned price mechanism, there is no need for government's interference with the market. However, the results of the survey about market collapse in October 1987 released by the U.S. government in 1988 revealed that, for the most part, it is resulted from investor psychology; thus, the circuit breaker mechanisms were established by the U.S. government to eliminate the psychological scare of the investors, that becomes the source of the market intrvention of governments. For example, a market-wide trading halt of New York Stock Exchange, in the event of a 10% decline in the Dow Jones Industrial Average (DJIA), would be one hour. The London Stock Exchange's circuit breaker system automatically halts any stock that is trading unusually lower or higher than 3% for five minutes.

According to the report by International Organization of Securities Commissions in 2010, many emerging markets and developed markets adopt circuit breaker mechanisms to deal with investors' panic problems. The supporters (e.g. Greenwald and Stein, 1991) claimed that circuit breakers provide investors with a cooling off period to calm fear and panic. On the contrary, the opponents (Lee et al, 1994) argued that halts are unhelpful for price discovery and do not actually reduce volatility in trading following the lifting of the halt. For instance, if fundamental information arrives at the time of the circuit breaker, the price adjustment process is delayed. Because no information is transmitted through trading if there is a halt, this may in turn increase price uncertainty.

Those mechanisms work in such a way that significant fluctuations in a securities' or index price would trigger an automatic halt in the trading of the security or a suspension of the entire market. However, market collapses is resulted from constantly price deviation caused by optimistic IS trend. Therefore, in addition to remedy mechanisms such as circuit breaker, prevention mechanisms should be taken to ease the influence of IS trend on stock price. This study presents a way to find out the stocks of which prices are effected by optimistic IS trends easily. Management tools such as taking advance collection of funds, decreasing margin purchase ratio could be taken to keep off continuous price rise, and cause market collapses.

Besides, in terms of circuit breaker mechanism, the market price indicator show that the price fluctuation is not necessary resulted from investors' panic. Though the market price may be influenced by sentiment, the influence will not be found on every stock. Therefore, this mechanism is started based on price fluctuations without determining whether the causes of the fluctuations are resulted from IS. The results of this study will be helpful in the improvement of the shortcomings in this mechanism.

First, for the starting base of this mechanism, in addition to considering the price fluctuations, IS should be also included to clarify if the cause of the price fluctuations is resulted from emotions and decide if the stop mechanism should be started. Secondly, as far as "the practical range" of this mechanism concerned, the method of this paper can be adopted to find out the stock characteristic group affected by IS susceptibly for practice. Furthermore, in terms of circuit breakers for individual stocks, the practical operation for circuit breaker is based on the timely dynamic information to determine if the trading should be stopped. The IS indicator with reliability and validity provided in this paper can meet this demand.

In summary, the purpose of the market intervention is to eliminate the impact of the psychological factors of the investors against the price mechanism. If IS can be included in the base range of the market intervention, it will be helpful to improve the effectiveness of interventions. Above all, the results and the analysis methods of this study will help the improvement of the performance of the mechanism.

4.6 Conclusion

As suggested by Baker and Wurgler (2007), measuring investor sentiment and understanding the variation in investor sentiment over time are important work to help interpret limited arbitrage. In this study, psychometric methods are used for assessing reliability and validity of levels and trends of investor sentiment for individual stocks. Turnover rate of trading volume, turnover rate of trading value, and turnover rate of transactions, three liquidity indicators, have been found to be appropriate proxies to measure IS for individual stocks. In addition, return can be affected by IS trends, depending on stock characteristics. The results imply that there exists heterogeneity in arbitrage among different stocks. Arbitrage is limited for those stocks in high-tech industry with high individual holdings. According to noise trader theory, irrational investors trade in the same way, causing a limit to arbitrage. Another implication of this study is that investor sentiment can serve as a preventive criterion for market supervisors to implement circuit breaker mechanisms. Instead of observing stock price, market supervisors should capture investor sentiment so that market efficiency can be improved. If an optimistic IS trend persists, the market price will deviate from fundamental values, and the deviation may finally lead to market collapse. Therefore, market intervention is needed when stock price is affected by IS trend. In the future, more sophisticated analytical techniques such as latent growth modeling, a typical application of SEM, can be used to analyze the influence of stock characteristics on IS trend. The information obtained can help build market intervention rules to make IS trend smoother.

參考文獻

Admati, A. R. & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies*, 1, 3-40.

Ahn, O.H., Kim, J. & Son, P. (2007). The pricing of underwriting services in the Korean IPO market: Characteristics and determinants. *Asia-Pacific Journal of Financial Studies,* 36, 731-764.

Alpert, M. & Raiffa, H. (1982). A progress report on the training of probability assessors. In *Judgment under Uncertainty: Heuristics and Biases*, edited by D. Kahneman, P. Slovic, and A. Tversky, pp. 249-305, Cambridge University Press: Cambridge.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411-423.

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transaction on Automatic Control*, AC-19, 716-723.

Baker, M. & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7, 271-299.

Baker, M., & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *Journal of Finance*, 55, 2219-2257.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61, 1645-1680.

Baker, M. & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21, 129-151.

Barber, B. M., Odean, T., & Zhu, N. (2009). Systematic noise. *Journal of Financial Markets*, 12, 547-569.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49, 307-343.

Bauer, D. J. (2003). Estimating multilevel linear models as structural models. *Journal of Educational and Behavioral Statistics*, 28, 135-167.

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: a review. *International Journal of Research in Marketing*, 13**,** 139-161.

Beck, N., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review*, 89, 634-647.

Bentler, P. M., & Wu, E. J. C. (2005). *EQS 6.1 - Structural equation modeling software for windows*. Encino, CA: Multivariate Software, Inc.

Berkman, H. & Eleswarapu, V. R. (1998). Short term traders and liquidity: A test using bombay stock exchange data. *Journal of Financial Economics*, 47, 339-355.

Black, F. (1986). Noise. *Journal of Finance*, 41, 529-543.

Blozis, S. A. (2006). A second-order structured latent curve model for longitudinal data. In K. van Montfort, H. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 189-214). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Blozis, S. A., Harring, J. R., & Mels, G. (2008). Using LISREL to fit nonlinear latent curve models. *Structural Equation Modeling*, 15, 346-369.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.

Bovaird, J. A. (2007). Multilevel structural equation models for contextual factors. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 149–182). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis*: *Forecasting and control* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Brennan, M. J. & Wang, A. W. (2010). The Mispricing Return Premium, Review of Financial Studies 23, 3437-3468.
Brown, G. W. & Cliff, M. T.(2004).Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, 11, 1-27.
Brown, G. W. & Cliff, M. T. (2005). Investor sentiment and asset valuation. *Journal of Business*, 78, 405-440.
Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2005). *HLM for Windows* (*Version 6.03*) [Computer Software]. Lincolnwood, IL: Scientific Software International.
Browne, M. W., & du Toit, S. H. C. (1991). Models for learning data. In L. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 47-68). Washington, DC: APA.
Campbell, J. Y., Grossman, S. J., & Wang, J. (1994). Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics*, 108, 905-939.
Campbell, J. Y. & Kyle, A. (1987). *Smart Money, Noise Trading, and Stock Price Behavior*, Princeton University Press: Princeton, NJ.
Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods*, 1, 421-483.
Chang, C., Lee, A. C. & Lee, C. F. (2009). Determinants of capital structure choice: A structural equation modeling approach. *Quarterly Review of Economics and Finance*, 49, 197-213.
Chen, H., Chong, T. T. L., & Duan, X. (2010). A principal-component approach to measuring investor sentiment. *Quantitative Finance*, 10, 339-347.
Choel, H. & Yang, C. W. (2010). Liquidity commonality and its causes: Evidence from the Korean stock market. *Asia-Pacific Journal of Financial Studies*, 39, 626-658.
Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16, 64-73.
Cooper, M. (1999). Filter rules based on price and volume in individual security overreaction. *The Review of Financial Studies*, 12, 901-935.
Cooper, M. J., Dimitrov, O. & Rau, P. R. (2001). A rose.com by any other name. *Journal of Finance*, 56, 2371-2388.
Cornelli, F., Goldreich, D. & Ljungqvist, A. (2006). Investor sentiment and pre-IPO markets. *Journal of Finance*, 61, 1187-1216.
Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 96–104.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-234.
Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, 38, 529-569.
Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and over-reactions. *Journal of Finance*, 53, 1839-1886.
Dawson, K. S., Gennings, C., & Carter, W. H. (1997). Two graphical techniques useful in detecting correlation structure in repeated measures data. *The American Statistician*, 51, 275-283.
DeLong, J. B., Shleifer, A., Summers, L. H. & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98, 703-738.
Derrien, F. (2005). IPO pricing in 'hot' market conditions: Who leaves money on the table? *Journal of Finance*, 60, 487-521.
Diggle, P. J., Liang, K., & Zeger, S. L. (2002). *Analysis of longitudinal data*. New York: Oxford University Press.

Dorn, D. (2009). Does sentiment drive the retail demand for IPOs? *Journal of Financial and Quantitative Analysis*, 44, 85-108.

Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal design. *Psychological Methods*, 1, 236-248.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Eyduran, E., & Akbas, Y. (2010). Comparison of different covariance structure used for experimental design with repeated measurement. *The Journal of Animal & Plant Sciences*, 20, 44-51.

Ferron, J., Dailey, R. F., & Yi, Q. (2002). Effects of misspecifying the first-level error structure in two-level models of change. *Multivariate Behavioral Research*, 37, 379-403.

Figlewski, S., Subjective information and market efficiency in a betting market. *Journal of Political Economy*, 1979, 87, 75-88.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-50.

Gerbing, D. W., & Anderson, J. C. (1998). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25, 186-192.

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The high-volume return premium. *Journal of Finance*, 56, 877-919.

Glosten, L. R. & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14, 71-100.

Goldstein, H., Healy, M. J. R., & Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, 13, 1643-1655.

Greene, J. T. & Smart, S. (1999). Liquidity provision and noise trading: Evidence from the investment dart board column. *Journal of Finance*, 54, 1885-1899.

Greenwald, B., & Stein, J. (1991). Transactional risk, market crashes, and the role of circuit breakers. *Journal of Business*, 64, 443-462.

Grimm, K. J., & Widaman, K. F. (2010). Residual structures in latent growth curve modeling. S*tructural Equation Modeling*, 17, 424-442.

Grossman, S., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70, 393- 408.

Guilford, J. P. (1954). *Psychometric methods*. McGraw-Hill, New York.

Guttman, L. (1954). A new approach to factor analysis: The radix. In P. F. Lazarfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258–348). New York: Columbia University Press.

Hair, Jr., J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis: A Global Perspective*,(7th ed.), Pearson Education International.

Hancock, G. R., Kuo, W. L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling*, 8, 470-489.

Haruvy, E. & Noussair, C. N. (2006). The effect of short selling on bubbles and crashes in experimental spot asset markets. *Journal of Finance*, 61, 1119-1157.

Heitjan, D. F., & Sharma, D. (1997). Modeling repeated series longitudinal data. *Statistics in Medicine*, 16, 347-355.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance,* 56, 1533-1597.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, 54, 2143-2184.

Hsinchu Science Park Administration. (2008). *Annual report: Hsinchu Science Park*, Hsinchu

Science Park Administration: Hsinchu, Taiwan.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, **6**, 1-55.
Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20, 90-98.
International Organization of Securities Commissions (IOSCO). (2010). Effectiveness of market interventions in emerging markets, IOSCO: Madrid, Spain.
Jöreskog, K. G. (1974). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 109-133.
Jöreskog, K., & Sörbom, D. (2001). *LISREL 8: User's reference guide,* Chicago: Scientific Software International.
Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-292.
Kaustia, M. & Knüpfer, S. (2008). Do investors overweight personal experience? Evidence from IPO subscriptions. *Journal of Finance*, 63, 2679-2702.
Kerlinger, F. (1986). *Foundations of Behavior Research*, (3rd ed.). Holt, Rinehart & Winston, New York.
Keselman, H. J., Algina, J., Kowalchuk, R. K., & Wolfinger, R. D. (1998), A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements, *Communications in Statistics: Simulation and Computation*, 27, 591-604.
Kim, K. & Byun, J. (2010). Effect of investor sentiment on market response to stock split announcement. *Asia-Pacific Journal of Financial Studies*, 39, 687-719.
Kurov, A. (2010). Investor sentiment and the stock market's reaction to monetary policy. *Journal of Banking and Finance*, 34, 139–149.
Kumar, A. & Lee, C. M. (2006). Retail investor sentiment and return comovements. *Journal of Finance,* 61, 2451-2486.
Kowalchuk, R. K., & Keselman, H. J. (2001). Mixed-model pairwise multiple comparisons of repeated measures means. *Psychological Methods*, 6, 282-296.
Kwok, O. M., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research*, 42, 557-592.
Lee, C. M., Ready, M., & Seguin P. J. (1994). Volume, volatility and New York stock exchange trading halts. *Journal of Finance*, 49, 183-214.
Lee, C. M., Shleifer, A. & Thaler, R. H. (1991).Investor sentiment and the closed-end fund puzzle. *Journal of Finance*, 46, 75-100.
Lee, C. M., & Swaminathan, B. (2000). Price momentum and trading volume. *Journal of Finance*, 55, 2017-2069.
Liang, B. (1999). Price pressure: Evidence from the "Dartbord" column. *Journal of Business*, 1999, 72, 119–134.
Lin, D. K. J., Kao, L. & Chen, A. (2010). Winner's curse in initial public offering subscriptions with investors' withdrawal options. *Asia-Pacific Journal of Financial Studies*, 39, 3-27.
Littell R. C., Henry, P. R., & Ammerman, C. B. (1998). Statistical analysis of repeated measures data using SAS procedures. *American Society of Animal Science*, 76, 1216-1231.
Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models* (2nd ed.), Cary, NC: SAS Institute Inc.
Littell R. C., Rendergast J., & Natarajan, R. (2000). Modeling covariance structure in the analysis of repeated measures data. *Statistics in Medicine*, 19, 1793-1819.
Ljungqvist, A. P., Nanda, V. & Singh, R. (2006). Hot markets, investor sentiment, and IPO pricing. *Journal of Business*, 79, 1667-1702.

MacCallum, R. C., Kim, C, Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, 32, 215-253.

Maddala, G. S. & Nimalendran, M. (1995). An unobserved component panel data model to study the effect of earnings surprise on stock prices, trading volumes, and spreads. *Journal of Econometrics*, 68, 229-242.

Mansour, H., Nordheim, E. V., & Rutledge, J. J. (1985). Maximum likelihood estimation of variance components in repeated measures designs assuming autoregressive errors. *Biometrics*, 41, 287-294.

McDonald, R. P. (1999). *Test theory: A Unified Treatment*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, 10, 259-284.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55, 107-122.

Murphy, D. L., & Pituch, K. A. (2009). The Performance of multilevel growth curve models under an autoregressive moving average process. *Journal of Experimental Education*, *77*, 255-282.

Muthén, B. O., & Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, 10, 73-101.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9, 599-620.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th ed.), Los Angeles, CA: Author.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.), Richmond: Virginia Commonwealth University.

Neal, R. & Wheatley, S. M. (1998). Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis*, 33, 523-548.

Newsom, J. T. (2002). A multilevel structural equation model for dyadic data. *Structural Equation Modeling,* 9, 431-447.

Nunnally, J. C. (1978). Psychometric theory, (2nd ed.). McGraw-Hill, New York.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *Journal of Finance*, 53, 1887-1934.

Ofek, E. and Richardson, M. (2003). DotCom mania: The rise and fall of internet stock prices. *Journal of Finance*, 58, 1113-1138.

Orhan, H., Eyduran, E., & Akbas, Y. (2010). Defining the best covariance structure for sequential variation on live weights of anatolian merinos male lambs. *The Journal of Animal & Plant Sciences*, 20, 158-163.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, 5, 343-355.

Prechter, R. R. (1999). *The wave principle of human social behavior and the new science of socionomics*. New Classics Library, Gainesville, GA.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks, CA: Sage.

Rasbash, J., Browne, W. J., Goldstein, H., Yang, M., Plewis, I., Healy, M., et al. (2000). *A user's guide to MLwiN* (Version 2.1). London: Institute of Education, University of London.

Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology, 52,* 501-525.

Ritter, J., The long-run performance of initial public offerings. *Journal of Finance*, 1991, 46, 3-27.

Rovine, M. J., & Molenaar, P. C. M. (1998). A LISREL model for the analysis of repeated measure with a patterned covariance matrix. *Structural Equation Modeling*, 5, 318-343.
Rovine, M. J., & Molenaar, P. C. M. (2000). A structural modeling approach to a multilevel random coefficients model. *Multivariate Behavioral Research*, 35, 51-88.
Saxton, J. (2008). *The U.S. Housing Bubble and the Global Financial Crisis: Vulnerabilities of the Alternative Financial System* (*Research Report #110-23*), Joint Economic Committee, United States Congress.
SAS Institute Inc. (2010). *SAS/STAT user's guide* (*SAS* 9.2). Cary, NC: Author.
Sayer, A. G., & Cumsille, P. E. (2001). Second-order latent growth models. In L. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 179-200). Washington DC: American Psychological Association.
Scheinkman, J. & Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, 111, 1183-1219.
Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
Shih Hsin University Department of Finance. (2011). Survey on Taiwan investor sentiment ,http://contract.shu.edu.tw/~emotion, accessed June 14.
Shiller, R. J. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity*, 2, 457-498.
Shiller, R. J. (1990). Speculative prices and popular models. *Journal of Economic Perspectives*, 4, 55-65.
Shleifer, A. (2000). *Inefficient Market: An Introduction to Behavioral Finance*, Oxford University Press: Oxford.
Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 24, 323-355.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
Sivo, S., Fan, X., & Witta, L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. *Structural Equation Modeling*, 12, 215-231.
Sivo, S., & Fan, X. (2008). The latent curve ARMA (p, q) panel model: Longitudinal data analysis in educational research and evaluation. *Educational Research and Evaluation*, 14, 363-376.
Smidt, S. (1968). A new look at the random walk hypothesis. *Journal of Financial and Quantitative Analysis*, 3, 235-261.
Stigler, G. J. (1964). Public regulation of the securities markets. *Journal of Business*, 7, 117-142.
Stiglitz, J. E. (1989). Using tax policy to curb speculative short term trading. *Journal of Financial Services Research*, 3, 101-115.
Summers, L. H. & Summers, V. P. (1989).When financial markets work too well: A cautious case for a securities transaction tax. *Journal of Financial Services Research*, 3, 261-286.
Thaler, R. H. (1999). The end of behavioral finance. *Financial Analysts Journal*, 55, 12-17.
Titman, S. & Wessels, R. (1988). The determinants of capital structure Choice. *Journal of Finance*, 43, 1-19.
Velicer, W. F., & Fava, J. L. (2003). Time series analysis. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology*. *Vol. 2*: *Research methods in psychology*, (pp. 581-606). Hoboken, NJ: Wiley.
Verbeke G., & Molenberghs G. (1997). *Linear mixed models in practice: A SAS-oriented approach*. Lecture Notes in Statistics 126, New York: Springer-Verlag.
West, S. G., & Hepworth, J. T. (1991). Data analytic strategies for temporal data and daily

events. *Journal of Personality*, 59, 609-662.
Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin,* 116, 363-381.
Wolfinger, R. (1993). Covariance structure selection in general mixed models. *Communications in Statistics—Simulation and Computatio*n, 22, 1079-1106.
Wolfinger, R. (1996). Covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics,* 1*,* 205-230.
Wu, W., & West, S. G. (2010). Sensitivity of fit indices to misspecification in growth curve models. *Multivariate Behavioral Research*, 45, 420-452.
Wu, W., West, S. G., & Taylor, A. B. (2009). Evaluating model fit for growth curve models: Integration of fit indices from SEM and MLM frameworks. *Psychological Methods*, 14, 183-201.
Wulff, S. S., & Robinson , T. J. (2009). Assessing the uncertainty of regression estimates in a response surface model for repeated measures. *Quality Technology & Quantitative Management*, 6, 309-324.
Yuan K., & Bentler, P. M. (2004). One chi-square difference and z tests in mean and covariance structure analysis when the base model is misspecified. *Educational and Psychological Measurement*, 64, 737-757.
Zweig, M. E. (1973). An investor expectations stock price predictive model using closed-end fund premiums. *Journal of Finance*, 28, 67-87.

附註：與執行本計畫相關之著作及學生畢業論文：

Ding, C. G., & Jane, T. D. (accepted). Using SAS PROC CALIS to Fit Level-1 Error Covariance Structures of Latent Growth Models. *Behavior Research Methods*.

Ding, C. G., & Jane, T. D. (2011). Identifying the Order of the ARMA Process for Level-1 Errors in Latent Growth Modeling. Paper presented at the 2011 Academy of Management Annual Meeting – Research Methods Division, San Antonio, Texas, USA, August 12-16, 2011.

鄭天德，「成長模型第一層誤差共變異結構之鑑定準則」，國立交通大學經營管理研究所博士論文，民國100年1月。

Table 1-1
SAS Statements in PROC CALIS for Specifying Different Types of the Level-1 Error Covariance Structure with Four Occasions

Structure ($\boldsymbol{\Theta}_\varepsilon$) and ECM	Statements in PROC CALIS
AR(1): $\varepsilon_t = \phi_1 \varepsilon_{t-1} + \nu_t,\ \lvert\phi_1\rvert < 1;$ $\sigma_\varepsilon^2 \begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ \rho_2 & \rho_1 & 1 & \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix},$ $\rho_k = \phi_1^k,\ k > 0.$	**STD** E1=VARE, E2=VARE, E3=VARE, E4=VARE, D0 =VARD0, D1 =VARD1; **COV** E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1, E1 E3=COV_lag2, E2 E4=COV_lag2, E1 E4=COV_lag3, D0 D1=CD0D1; **PARAMETERS** PHI1; COV_lag1= PHI1*VARE; COV_lag2=(PHI1**2)*VARE; COV_lag3= (PHI1**3) *VARE; **BOUNDS** –1. < PHI1 < 1. ;
MA(1): $\varepsilon_t = \nu_t - \theta_1 \nu_{t-1},\ \lvert\theta_1\rvert < 1;$ $\sigma_\varepsilon^2 \begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ 0 & \rho_1 & 1 & \\ 0 & 0 & \rho_1 & 1 \end{bmatrix},$ $\rho_1 = \frac{-\theta_1}{(1+\theta_1^2)},\quad \rho_k = 0, k > 1.$	**STD** E1=VARE, E2=VARE, E3=VARE, E4=VARE, D0 =VARD0, D1 =VARD1; **COV** E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1, D0 D1=CD0D1; **PARAMETERS** THE1; COV_lag1= (–THE1/(1+ THE1**2))*VARE; **BOUNDS** –1. < THE1 < 1. ;
ARMA(1,1): $\varepsilon_t = \phi_1 \varepsilon_{t-1} + \nu_t - \theta_1 \nu_{t-1},$ $\lvert\phi_1\rvert < 1,\ \lvert\theta_1\rvert < 1;$ $\sigma_\varepsilon^2 \begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ \rho_2 & \rho_1 & 1 & \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix},$ $\rho_1 = \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{(1 - 2\phi_1\theta_1 + \theta_1^2)},$ $\rho_k = \phi_1 \rho_{k-1},\ k > 1.$	**STD** E1=VARE, E2=VARE, E3=VARE, E4=VARE, D0=VARD0, D1=VARD1; **COV** E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1, E1 E3=COV_lag2, E2 E4=COV_lag2, E1 E4=COV_lag3, D0 D1=CD0D1; **PARAMETERS** PHI1 RHO1; COV_lag1=RHO1*VARE; COV_lag2=PHI1* COV_lag1; COV_lag3=PHI1* COV_lag2; **BOUNDS** –1. < PHI1 < 1.;

Table 1-1 (Continued)

Structure ($\boldsymbol{\Theta}_{\varepsilon}$) and ECM	Statements in PROC CALIS
AR(2): $\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \nu_t$, $\lvert \phi_2 \rvert < 1, \phi_2 + \phi_1 < 1, \phi_2 - \phi_1 < 1$; $\sigma_{\varepsilon}^2 \begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ \rho_2 & \rho_1 & 1 & \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}$, $\rho_0 = 1$, $\rho_1 = \phi_1 / (1 - \phi_2)$, $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2},\ k > 1$,	**STD** E1-E4=4*VARE, D0=VARD0, D1=VARD1; **COV** E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1, E1 E3=COV_lag2, E2 E4=COV_lag2, E1 E4=COV_lag3, D0 D1=CD0D1; **PARAMETERS** PHI1 PHI2; RHO1= PHI1/(1–PHI2); COV_lag1=RHO1*VARE; COV_lag2=PHI1*COV_lag1+ PHI2 *VARE; COV_lag3=PHI1*COV_lag2+PHI2*COV_lag1; **LINCON** PHI2 + PHI1 < 1., PHI2 –PHI1 < 1.; **BOUNDS** –1. < PHI2 < 1.;
MA(2): $\varepsilon_t = \nu_t - \theta_1\nu_{t-1} - \theta_2\nu_{t-2}$, $\lvert \theta_2 \rvert < 1, \theta_2 + \theta_1 < 1, \theta_2 - \theta_1 < 1$; $\sigma^2 \begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ \rho_2 & \rho_1 & 1 & \\ 0 & \rho_2 & \rho_1 & 1 \end{bmatrix}$, $\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{(1+\theta_1^2+\theta_2^2)},\ \rho_2 = \frac{-\theta_2}{(1+\theta_1^2+\theta_2^2)}$, $\rho_k = 0, k > 2$,	**STD** E1=VARE, E2=VARE, E3=VARE, E4=VARE, D0=VARD0, D1=VARD1; **COV** E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1, E1 E3=COV_lag2, E2 E4=COV_lag2, D0 D1=CD0D1; **PARAMETERS** THE1 THE2; COV_lag1=((–THE1+THE1*THE2)/ (1+THE1**2+THE**2))*VARE; COV_lag2=(–THE2/(1+THE1**2+THE2**2))*VARE; **LINCON** THE2+ THE1 < 1., THE2 –THE1 < 1.; **BOUNDS** –1. < THE2 < 1. ;
ARH(1) (heterogeneous AR(1)): $\begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho & \sigma_{\varepsilon_2}^2 & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho^2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho & \sigma_{\varepsilon_3}^2 & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho^3 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho^2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho & \sigma_{\varepsilon_4}^2 \end{bmatrix}$	**STD** E1=VARE1, E2=VARE2, E3=VARE3, E4=VARE4, D0=VARD0, D1=VARD1; **COV** E1 E2=COVE1E2, E1 E3=COVE1E3, E1 E4=COVE1E4, E2 E3=COVE2E3, E2 E4=COVE2E4, E3 E4=COVE3E4, D0 D1=CD0D1; **PARAMETERS** RHO; COVE1E2=SQRT(VARE1)*SQRT(VARE2)*RHO; COVE1E3=SQRT(VARE1)*SQRT(VARE3)*RHO**2; COVE1E4=SQRT(VARE1)*SQRT(VARE4)*RHO**3; COVE2E3=SQRT(VARE2)*SQRT(VARE3)*RHO; COVE2E4=SQRT(VARE2)*SQRT(VARE4)*RHO**2; COVE3E4=SQRT(VARE3)*SQRT(VARE4)*RHO;

Table 1-1 (Continued)

Structure ($\mathbf{\Theta}_\varepsilon$) and ECM	Statements in PROC CALIS
TOEPH (heterogeneous Toeplitz): $\begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho_1 & \sigma^2_{\varepsilon_2} & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho_2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho_1 & \sigma^2_{\varepsilon_3} & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho_3 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho_2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho_1 & \sigma^2_{\varepsilon_4} \end{bmatrix}$	**STD** E1=VARE1, E2=VARE2, E3=VARE3, E4=VARE4, D0=VARD0, D1=VARD1; **COV** E1 E2=COVE1E2, E1 E3=COVE1E3, E1 E4=COVE1E4, E2 E3=COVE2E3, E2 E4=COVE2E4, E3 E4=COVE3E4, D0 D1=CD0D1; **PARAMETERS** RHO1 RHO2 RHO3; COVE1E2=SQRT(VARE1)*SQRT(VARE2)*RHO1; COVE1E3=SQRT(VARE1)*SQRT(VARE3)*RHO2; COVE1E4=SQRT(VARE1)*SQRT(VARE4)*RHO3; COVE2E3=SQRT(VARE2)*SQRT(VARE3)*RHO1; COVE2E4=SQRT(VARE2)*SQRT(VARE4)*RHO2; COVE3E4=SQRT(VARE3)*SQRT(VARE4)*RHO1;
UN: $\begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2\varepsilon_1} & \sigma^2_{\varepsilon_2} & & \\ \sigma_{\varepsilon_3\varepsilon_1} & \sigma_{\varepsilon_3\varepsilon_2} & \sigma^2_{\varepsilon_3} & \\ \sigma_{\varepsilon_4\varepsilon_1} & \sigma_{\varepsilon_4\varepsilon_2} & \sigma_{\varepsilon_4\varepsilon_3} & \sigma^2_{\varepsilon_4} \end{bmatrix}$	**STD** E1=VARE1, E2=VARE2, E3=VARE3, E4=VARE4, D0=VARD0, D1=VARD1; **COV** E1 E2=COVE1E2, E1 E3=COVE1E3, E1 E4=COVE1E4, E2 E3=COVE2E3, E2 E4=COVE2E4, E3 E4=COVE3E4, D0 D1=CD0D1;

Note. The level-2 ECM, $\mathbf{\Psi}_{\zeta_\eta} = \begin{bmatrix} \sigma^2_{\zeta_{\eta\alpha}} & \\ \sigma_{\zeta_{\eta\alpha}\zeta_{\eta\beta}} & \sigma^2_{\zeta_{\eta\beta}} \end{bmatrix}$, is estimated with type = UN. ρ_k denotes the autocorrelation coefficient at lag *k*. SAS PROC MIXED provides only the options of ARMA(1,1) and AR(1) for the ARMA family.

Table 1-2
Population Parameters of the Model in Figure 1-1 with the Level-1 Error Covariance Structure of ARH(1) and the Sample Covariance Matrix of y_1–y_4 and X Resulting from a Dataset of Size 300 Generated from the Model

$$\mathbf{\Lambda}_{\mathbf{y}}^{*} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \mathbf{\Gamma}_{\mathbf{x}} = \begin{bmatrix} \gamma_{10} \\ \gamma_{11} \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \ \mathbf{\Sigma}_{\mathbf{xx}} = \sigma_X^2 = 1,$$

$$\boldsymbol{\mu}_{\mathbf{x}} = \mu_X = 0, \ \mathbf{\Gamma}_0 = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \end{bmatrix},$$

$$\mathbf{\Psi}_{\zeta_\eta} = \begin{bmatrix} \sigma^2_{\zeta_{\eta\alpha}} & \\ \sigma_{\zeta_{\eta\alpha}\zeta_{\eta\beta}} & \sigma^2_{\zeta_{\eta\beta}} \end{bmatrix} = \begin{bmatrix} 15 & \\ 7 & 10 \end{bmatrix},$$

$$\mathbf{\Theta}_{\boldsymbol{\varepsilon}} = \begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho & \sigma^2_{\varepsilon_2} & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho^2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho & \sigma^2_{\varepsilon_3} & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho^3 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho^2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho & \sigma^2_{\varepsilon_4} \end{bmatrix}$$

$\sigma^2_{\varepsilon_1} = 36, \sigma^2_{\varepsilon_2} = 25, \sigma^2_{\varepsilon_3} = 49, \sigma^2_{\varepsilon_4} = 64$, $\rho = .7$

Population Model-Implied Mean Vector and Covariance Matrix (Equations 4 and 5)

	y_1	y_2	y_3	y_4	X
y_1	67.000				
y_2	83.000	164.000			
y_3	113.580	240.500	388.000		
y_4	140.464	312.600	501.200	695.000	
X	4.000	10.000	16.000	22.000	1.000
Mean	10.000	14.000	18.000	22.000	.000

Sample Mean Vector and Covariance Matrix

	y_1	y_2	y_3	y_4	X
y_1	66.557				
y_2	80.505	157.888			
y_3	109.910	233.411	385.350		
y_4	140.643	307.945	501.530	703.510	
X	3.103	8.685	14.137	19.714	.855
Mean	9.834	14.098	17.524	21.167	–.001

Table 1-3

Summary of the Results by Fitting ARH(1) for Level-1 Errors and UN for Level-2 Errors Based on the Sample Covariance Matrix Shown in Table 1-2 by Using PROC CALIS and PROC MIXED

Assessment of model fit by							
PROC CALIS						PROC MIXED	
χ^2	*df*	$P_r > \chi^2$	CFI	NNFI	RMSEA	AIC	BIC
11.076	6	.086	.998	.996	.050	7825.5	7870.0

Parameter estimates by fitting ARH(1) for $\mathbf{\Theta}_\varepsilon$ and UN for $\mathbf{\Psi}_{\zeta_\eta}$

Parameters	Estimates by using PROC CALIS	Estimates by using PROC MIXED
$\mathbf{\Theta}_\varepsilon = \begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho & \sigma^2_{\varepsilon_2} & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho^2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho & \sigma^2_{\varepsilon_3} & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho^3 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho^2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho & \sigma^2_{\varepsilon_4} \end{bmatrix}$	$\begin{bmatrix} 40.56^{**} & & & \\ 25.47^{*} & 29.37^{*} & & \\ 27.68^{*} & 31.93^{*} & 63.84^{**} & \\ 24.48 & 28.24 & 56.46^{*} & 91.79^{**} \end{bmatrix}$ $\hat{\rho} = .74^{***}$.	$\begin{bmatrix} 40.43^{**} & & & \\ 25.37^{a} & 29.27^{*} & & \\ 27.59^{a} & 31.83^{a} & 63.63^{**} & \\ 24.40^{a} & 28.15^{a} & 56.27^{a} & 91.49^{**} \end{bmatrix}$ $\hat{\rho} = .74^{***}$.
$\mathbf{\Psi}_{\zeta_\eta} = \begin{bmatrix} \sigma^2_{\zeta_{\eta\alpha}} & \\ \sigma_{\zeta_{\eta\alpha}\zeta_{\eta\beta}} & \sigma^2_{\zeta_{\eta\beta}} \end{bmatrix}$	$\begin{bmatrix} 14.53 & \\ 9.21^{***} & 8.85^{***} \end{bmatrix}$	$\begin{bmatrix} 14.48 & \\ 9.18^{***} & 8.82^{***} \end{bmatrix}$
$\begin{bmatrix} \gamma_{00} & \gamma_{01} \\ \gamma_{10} & \gamma_{11} \end{bmatrix}$	$\begin{bmatrix} 10.18^{***} & 3.86^{***} \\ 3.68^{***} & 6.48^{***} \end{bmatrix}$	$\begin{bmatrix} 10.17^{***} & 3.86^{***} \\ 3.68^{***} & 6.47^{***} \end{bmatrix}$

[a] Test for significance cannot be achieved. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 1-4

Population Parameters of the Model in Figure 1-2 with the Level-1 Error Covariance Structure of AR(1) and the Sample Covariance Matrix Resulting from a Dataset of Size 300 Generated from the Model

$$\mathbf{\Lambda_y} = \begin{bmatrix} \lambda_{y_{11}} & 0 & 0 & 0 \\ \lambda_{y_{21}} & 0 & 0 & 0 \\ \lambda_{y_{31}} & 0 & 0 & 0 \\ 0 & \lambda_{y_{12}} & 0 & 0 \\ 0 & \lambda_{y_{22}} & 0 & 0 \\ 0 & \lambda_{y_{32}} & 0 & 0 \\ 0 & 0 & \lambda_{y_{13}} & 0 \\ 0 & 0 & \lambda_{y_{23}} & 0 \\ 0 & 0 & \lambda_{y_{33}} & 0 \\ 0 & 0 & 0 & \lambda_{y_{14}} \\ 0 & 0 & 0 & \lambda_{y_{24}} \\ 0 & 0 & 0 & \lambda_{y_{34}} \end{bmatrix} = \begin{bmatrix} 1.00 & 0 & 0 & 0 \\ .75 & 0 & 0 & 0 \\ .85 & 0 & 0 & 0 \\ 0 & 1.00 & 0 & 0 \\ 0 & .75 & 0 & 0 \\ 0 & .85 & 0 & 0 \\ 0 & 0 & 1.00 & 0 \\ 0 & 0 & .75 & 0 \\ 0 & 0 & .85 & 0 \\ 0 & 0 & 0 & 1.00 \\ 0 & 0 & 0 & .75 \\ 0 & 0 & 0 & .85 \end{bmatrix},$$

$$\mathbf{\Gamma}_0 = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \end{bmatrix} = \begin{bmatrix} 12 \\ 1 \end{bmatrix}, \quad \mathbf{\Gamma}_\xi = \begin{bmatrix} \gamma_{10} \\ \gamma_{11} \end{bmatrix} = \begin{bmatrix} 6 \\ .5 \end{bmatrix},$$

$$\mathbf{\Phi}_\xi = \sigma^2_\xi = 4\,, \quad \boldsymbol{\mu}_\xi = 13\,,$$

$$\mathbf{\Theta}_{\boldsymbol{\delta}} = Diag[\sigma^2_{\delta_1} \;\; \sigma^2_{\delta_2} \;\; \sigma^2_{\delta_3}] = Diag[.81 \;\; .36 \;\; 1.00]$$

$$\mathbf{\Psi}_{\zeta_{\boldsymbol{\eta}}} = Cov[\zeta_{\eta_\alpha} \;\; \zeta_{\eta_\beta}]' = \begin{bmatrix} \sigma^2_{\zeta_{\eta_\alpha}} & \\ \sigma_{\zeta_{\eta_\alpha}\zeta_{\eta_\beta}} & \sigma^2_{\zeta_{\eta_\beta}} \end{bmatrix} = \begin{bmatrix} .80 & \\ .25 & .60 \end{bmatrix}$$

$$\mathbf{\Psi}_{\zeta_F} = Diag[\sigma^2_{\zeta_{F_1}} \;\; \sigma^2_{\zeta_{F_2}} \;\; \sigma^2_{\zeta_{F_3}} \;\; \sigma^2_{\zeta_{F_4}}] = Diag[.25 \;\; .36 \;\; .49 \;\; .64],$$

$$\mathbf{\Lambda}^*_{\mathbf{y}} = \begin{bmatrix} 1 & T_1 - 1 \\ 1 & T_2 - 1 \\ 1 & T_3 - 1 \\ 1 & T_4 - 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \mathbf{\Lambda}_x = \begin{bmatrix} \lambda_{x_1} \\ \lambda_{x_2} \\ \lambda_{x_3} \end{bmatrix} = \begin{bmatrix} 1 \\ .75 \\ .70 \end{bmatrix},$$

$$\mathbf{\Theta}_{\boldsymbol{\varepsilon}} = Cov[\varepsilon_{11} \;\; \varepsilon_{12} \;\; \varepsilon_{13} \;\; \varepsilon_{14} \;\; \varepsilon_{21} \;\; \varepsilon_{22} \;\; \varepsilon_{23} \;\; \varepsilon_{24} \;\; \varepsilon_{31} \;\; \varepsilon_{32} \;\; \varepsilon_{33} \;\; \varepsilon_{34}]'$$

$$= \begin{bmatrix}
\sigma^2_{\varepsilon_1} \\
\phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \sigma^2_{\varepsilon_1} \\
\phi^2_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \sigma^2_{\varepsilon_1} \\
\phi^3_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \phi^2_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \phi_{1\varepsilon_1}\sigma^2_{\varepsilon_1} & \sigma^2_{\varepsilon_1} \\
0 & 0 & 0 & 0 & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & \phi^2_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & \phi^3_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \phi^2_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_2}\sigma^2_{\varepsilon_2} & \sigma^2_{\varepsilon_2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{\varepsilon_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \sigma^2_{\varepsilon_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi^2_{1\varepsilon_3}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \sigma^2_{\varepsilon_3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi^3_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \phi^2_{1\varepsilon_3}\sigma^2_{\varepsilon_2} & \phi_{1\varepsilon_3}\sigma^2_{\varepsilon_3} & \sigma^2_{\varepsilon_3}
\end{bmatrix}.$$

$$\phi_{1\varepsilon_1} = .5, \;\; \sigma^2_{\varepsilon_1} = .25, \qquad \phi_{1\varepsilon_2} = .7, \;\; \sigma^2_{\varepsilon_2} = .36, \qquad \phi_{1\varepsilon_3} = .6, \;\; \sigma^2_{\varepsilon_3} = .40\,.$$

Table 1-4 (Continued)

Population Model-Implied Mean Vector and Covariance Matrix

	y_{11}	y_{21}	y_{31}	y_{12}	y_{22}	y_{32}	y_{13}	y_{23}	y_{33}	y_{14}	y_{24}	y_{34}	x_1	x_2	x_3
y_{11}	145.30														
y_{21}	108.79	81.95													
y_{31}	123.29	92.47	105.20												
y_{12}	157.18	117.79	133.49	171.51											
y_{22}	117.79	88.59	100.12	128.45	96.69										
y_{32}	133.49	100.12	113.71	145.57	109.18	124.14									
y_{13}	169.36	126.98	143.91	184.88	138.56	157.04	200.94								
y_{23}	126.98	95.41	107.93	138.56	104.17	117.78	150.52	113.25							
y_{33}	143.91	107.93	122.46	157.04	117.78	133.72	170.59	127.94	145.40						
y_{14}	181.58	136.16	154.32	198.66	148.95	168.81	215.78	161.74	183.30	233.59					
y_{24}	136.16	102.25	115.74	148.95	111.89	126.61	161.74	121.56	137.48	175.01	131.61				
y_{34}	154.32	115.74	131.26	168.81	126.61	143.63	183.30	137.48	156.05	198.34	148.75	168.99			
x_1	24.00	18.00	20.40	26.00	19.50	22.10	28.00	21.00	23.80	30.00	22.50	25.50	4.81		
x_2	18.00	13.50	15.30	19.50	14.63	16.58	21.00	15.75	17.85	22.50	16.88	19.13	3.00	2.61	
x_3	16.80	12.60	14.28	18.20	13.65	15.47	19.60	14.70	16.66	21.00	15.75	17.85	2.80	2.10	2.96
Mean	90.00	67.50	76.50	97.10	72.83	82.54	104.20	78.15	88.57	111.30	83.48	94.61	13.00	9.75	9.10

Sample Mean Vector and Covariance Matrix

	y_{11}	y_{21}	y_{31}	y_{12}	y_{22}	y_{32}	y_{13}	y_{23}	y_{33}	y_{14}	y_{24}	y_{34}	x_1	x_2	x_3
y_{11}	147.01														
y_{21}	109.81	82.57													
y_{31}	124.47	93.16	105.99												
y_{12}	158.66	118.63	134.42	172.82											
y_{22}	118.87	89.22	100.79	129.40	97.42										
y_{32}	134.12	100.36	113.97	145.95	109.46	123.87									
y_{13}	171.53	128.31	145.36	186.96	140.12	157.94	203.88								
y_{23}	127.89	95.91	108.45	139.39	104.82	117.87	152.01	113.85							
y_{33}	144.92	108.44	123.04	157.91	118.44	133.76	172.06	128.45	145.84						
y_{14}	182.10	136.18	154.29	198.88	149.11	168.06	216.81	161.76	183.06	232.49					
y_{24}	137.16	102.74	116.25	149.83	112.54	126.66	163.35	122.20	137.99	175.04	132.28				
y_{34}	154.50	115.60	131.06	168.81	126.64	142.83	183.98	137.38	155.67	197.21	148.66	167.89			
x_1	24.39	18.28	20.68	26.40	19.74	22.29	28.51	21.27	24.07	30.08	22.68	25.53	4.94		
x_2	18.66	14.02	15.81	20.15	15.14	17.04	21.80	16.29	18.42	23.11	17.45	19.60	3.10	2.77	
x_3	18.01	13.47	15.28	19.36	14.55	16.40	20.93	15.60	17.69	22.20	16.73	18.85	3.07	2.34	3.22
Mean	91.24	68.43	77.64	98.60	73.96	83.90	105.81	79.37	89.96	113.04	84.81	96.12	13.21	9.92	9.30

Table 1-5

Summary of the Results by Fitting AR(1) for Level-1 Errors ε_{1t}, ε_{2t}, and ε_{3t} and UN for Level-2 Errors Based on the Sample Covariance Matrix Shown in Table 1-4 by Using PROC CALIS

Assessment of model fit					
χ^2	*df*	$P_r > \chi^2$	CFI	NNFI	RMSEA
90.49	109	.9009	1.0	1.0	<.0001

Parameter estimates by fitting AR(1) for $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, and $\boldsymbol{\varepsilon}_3$

$$\hat{\mathbf{\Lambda}}_y = \begin{bmatrix} 1.00 & 0 & 0 & 0 \\ .75^{***} & 0 & 0 & 0 \\ .85^{***} & 0 & 0 & 0 \\ 0 & 1.00 & 0 & 0 \\ 0 & .75^{***} & 0 & 0 \\ 0 & .85^{***} & 0 & 0 \\ 0 & 0 & 1.00 & 0 \\ 0 & 0 & .75^{***} & 0 \\ 0 & 0 & .85^{***} & 0 \\ 0 & 0 & 0 & 1.00 \\ 0 & 0 & 0 & .75^{***} \\ 0 & 0 & 0 & .85^{***} \end{bmatrix}, \hat{\mathbf{\Lambda}}_x = \begin{bmatrix} 1.00 \\ .75^{***} \\ .70^{***} \end{bmatrix},$$

$$\hat{\mathbf{\Gamma}}_0 = \begin{bmatrix} 13.08^{***} \\ 1.05^{***} \end{bmatrix}, \ \hat{\mathbf{\Gamma}}_\xi = \begin{bmatrix} 5.86^{***} \\ .47^{***} \end{bmatrix},$$

$$\hat{\mathbf{\Phi}}_\xi = 4.22^{***}, \ \hat{\boldsymbol{\mu}}_\xi = 13.20^{***},$$

$$\hat{\mathbf{\Theta}}_\delta = Diag[.847^{***} \ \ .375^{***} \ \ .988^{***}],$$

$$\hat{\mathbf{\Psi}}_{\zeta_\eta} = \begin{bmatrix} .88 & \\ .40^{**} & .67^{***} \end{bmatrix},$$

$$\hat{\mathbf{\Psi}}_{\zeta_F} = Diag\left[.300^{**} \ \ .361^{***} \ \ .466^{***} \ \ .699^{***}\right]$$

$$\hat{\mathbf{\Theta}}_\varepsilon = \begin{bmatrix}
.271^{***} \\
.137^{***} & .271^{***} \\
.069^{**} & .137^{***} & .271^{***} \\
.036^{*} & .069^{**} & .137^{***} & .271^{***} \\
0 & 0 & 0 & 0 & .367^{***} \\
0 & 0 & 0 & 0 & .254^{***} & .367^{***} \\
0 & 0 & 0 & 0 & .176^{***} & .254^{***} & .367^{***} \\
0 & 0 & 0 & 0 & .122^{***} & .176^{***} & .254^{***} & .367^{***} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .422^{***} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .243^{***} & .422^{***} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .140^{***} & .243^{***} & .422^{***} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .081^{***} & .140^{***} & .243^{***} & .422^{***}
\end{bmatrix}$$

$\hat{\phi}_{1\varepsilon_1} = .508^{***}$, $\hat{\sigma}^2_{\varepsilon_1} = .271^{***}$, $\hat{\phi}_{1\varepsilon_2} = .693^{***}$, $\hat{\sigma}^2_{\varepsilon_2} = .367^{***}$, $\hat{\phi}_{1\varepsilon_3} = .576^{***}$, $\hat{\sigma}^2_{\varepsilon_3} = .422^{***}$

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 2-1
Identification of the Orders p and q for AR(p) and MA(q) Processes for Level-1 Errors Based on ACF and PACF [a]

Sample ACF	Sample PACF	Process Identified
cut-off after lag 0	cut-off after lag 0	TOEP(1)
without cut-off	cut-off after lag p	AR(p)
cut-off after lag q	without cut-off	MA(q)
cut-off after lag q	cut-off after lag p	AR(p) or MA(q)
without cut-off	without cut-off	ARMA(p, q)

[a] $p \leq T–1$, $q \leq T–1$ (T is the total number of time points). A cut-off is determined by the occurrence of non-significance of a sample autocorrelation / partial autocorrelation.

Table 2-2

Population Parameters of the Model in Figure 1-1 with the Level-1 Error Covariance Structure of ARH(1) and the Results of Identifying an "Optimal" Level-1 Error Covariance Structure Based on a Dataset with N =300 Generated from the Population Model

$$\boldsymbol{\Lambda}_{\mathbf{y}}^{*}=\begin{bmatrix}1&0\\1&1\\1&2\\1&3\end{bmatrix},\boldsymbol{\Gamma}_{\mathbf{x}}=\begin{bmatrix}\gamma_{10}\\\gamma_{11}\end{bmatrix}=\begin{bmatrix}4\\6\end{bmatrix},\ \boldsymbol{\Sigma}_{\mathbf{xx}}=\sigma_X^2=1,\ \boldsymbol{\mu}_{\mathbf{x}}=\mu_X=0,\quad \boldsymbol{\Theta}_{\varepsilon}=\begin{bmatrix}\sigma_{\varepsilon_1}^2&&&\\\sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho&\sigma_{\varepsilon_2}^2&&\\\sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho^2&\sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho&\sigma_{\varepsilon_3}^2&\\\sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho^3&\sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho^2&\sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho&\sigma_{\varepsilon_4}^2\end{bmatrix}$$

$$\boldsymbol{\Gamma}_0=\begin{bmatrix}\gamma_{00}\\\gamma_{01}\end{bmatrix}=\begin{bmatrix}10\\4\end{bmatrix},\boldsymbol{\Psi}_{\zeta_\eta}=\begin{bmatrix}\sigma_{\zeta_{\eta\alpha}}^2&\\\sigma_{\zeta_{\eta\alpha}\zeta_{\eta\beta}}&\sigma_{\zeta_{\eta\beta}}^2\end{bmatrix}=\begin{bmatrix}15&\\7&10\end{bmatrix},\quad \sigma_{\varepsilon_1}^2=36,\sigma_{\varepsilon_2}^2=25,\sigma_{\varepsilon_3}^2=49,\sigma_{\varepsilon_4}^2=64,\ \rho=.7$$

	Population covariance matrix (Equation 5)					Sample Covariance Matrix				
	y_1	y_2	y_3	y_4	X	y_1	y_2	y_3	y_4	X
y_1	67.00					66.56				
y_2	83.00	164.00				80.51	157.89			
y_3	113.58	240.50	388.00			109.91	233.41	385.35		
y_4	140.45	312.60	501.20	695.00		140.64	307.95	501.53	703.51	
X	4.00	10.00	16.00	22.00	1.00	3.10	8.69	14.14	19.71	.86
Mean	10.00	14.00	18.00	22.00	.00	9.83	14.10	17.52	21.17	.00

Test for stationarity

Structure	#Para.	χ^2	df	Δdf	$\Delta\chi^2$	$P_r > \Delta\chi^2_{\Delta df}$
UN (unconstrained)	10	5.170	1	--	--	--
TOEP (constrained by stationarity)	4	58.411	7	6	53.241	<.0001

SCDT

		$H_0: M_T$ = UN						Model fit	
Step	Structure	#Para.	χ^2	df	Δdf	$\Delta\chi^2$	$P_r > \Delta\chi^2_{\Delta df}$	$P_r > \chi^2$	AIC
0	UN	10	5.170	1	--	--	--	.023	3.170
1	TOEPH(1)	4	30.960	7	6	25.789	< .0001	.0002	16.959
2	TOEPH(2)	5	19.941	6	5	14.771	.011	.003	7.941
	CSH	5	18.376	6	5	13.206	.022	.005	6.376
	ARH(1)	5	11.076	6	5	5.906	.315	.086	–.924

Parameter estimates by fitting ARH(1) for $\boldsymbol{\Theta}_\varepsilon$ and UN for $\boldsymbol{\Psi}_{\zeta_\eta}$

$\hat{\gamma}_{00}$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{11}$	$\hat{\sigma}^2_{\zeta_{\eta\alpha}}$	$\hat{\sigma}^2_{\zeta_{\eta\beta}}$	$\hat{\sigma}_{\zeta_{\eta\alpha}\zeta_{\eta\beta}}$	$\hat{\rho}$	$\hat{\sigma}^2_{\varepsilon_1}$	$\hat{\sigma}^2_{\varepsilon_2}$	$\hat{\sigma}^2_{\varepsilon_3}$	$\hat{\sigma}^2_{\varepsilon_4}$
10.18***	3.68***	6.48***	14.53	8.85***	9.21***	.74***	40.56**	29.37*	63.84**	91.79**

Note. M_T = {TOEPH(1), (TOEPH(2), CSH, ARH(1))}, $\Delta\chi^2$ = the chi-square difference between the constrained model and the unconstrained model, Δdf = the difference of degrees of freedom, $P_r > \Delta\chi^2_{\Delta df}$ denotes the p-value of the chi-square difference test.

$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 2-3
Population Parameters of the LGM in Figure 1-1 with Four Cases of ARMA(p, q) for the Level-1 Error Covariance Structure

$$\mathbf{\Lambda}_{\mathbf{y}}^{*}=\begin{bmatrix}1&0\\1&1\\1&2\\1&3\end{bmatrix},\ \mathbf{\Gamma}_{\mathbf{x}}=\begin{bmatrix}\gamma_{10}\\\gamma_{11}\end{bmatrix}=\begin{bmatrix}4\\6\end{bmatrix},\ \ \mathbf{\Sigma}_{ww}=\sigma_{w}^{2}=1,\ \ \boldsymbol{\mu}_{\mathbf{x}}=\mu_{X}=0,\ \ \mathbf{\Gamma}_{0}=\begin{bmatrix}\gamma_{00}\\\gamma_{01}\end{bmatrix}=\begin{bmatrix}10\\4\end{bmatrix},$$

$$\mathbf{\Psi}_{\zeta_{\eta}}=\begin{bmatrix}\sigma_{\zeta_{\eta\alpha}}^{2}&\\\sigma_{\zeta_{\eta\alpha}\zeta_{\eta\beta}}&\sigma_{\zeta_{\eta\beta}}^{2}\end{bmatrix}=\begin{bmatrix}1&\\-.5&1\end{bmatrix},\ \ \mathbf{\Theta}_{\varepsilon}=\sigma_{\varepsilon}^{2}\begin{bmatrix}1&&&\\\rho_{1}&1&&\\\rho_{2}&\rho_{1}&1&\\\rho_{3}&\rho_{2}&\rho_{1}&1\end{bmatrix},\ \sigma_{\varepsilon}^{2}=4.$$

Case	ARMA(p, q)	ϕ_1	ϕ_2	θ_1	θ_2
1	AR(1)	–.4	--	--	--
2	AR(2)	.4	–.7	--	--
3	MA(2)	--	--	–.4	.–.5
4	ARMA(1,1)	–.5	--	.5	--

Table 2-4

Results of Identifying an “Optimal” Level-1 Error Covariance Structure Based on Each of the Four Datasets with N =300 Generated from the LGM with Population Parameters Given in Table 2-3

Case 1: AR(1)

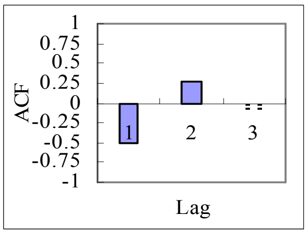
	Population covariance matrix (Equation 5)					Sample covariance matrix				
	y_1	y_2	y_3	y_4	X	y_1	y_2	y_3	y_4	X
y_1	21.00	38.50	65.00	87.00	4.00	20.95	38.71	64.25	86.68	3.98
y_2	38.50	105.00	159.50	223.00	10.00	38.71	102.74	155.02	218.26	9.77
y_3	65.00	159.50	263.00	354.50	16.00	64.25	155.02	253.91	344.43	15.55
y_4	87.00	223.00	354.50	495.00	22.00	86.68	218.26	344.43	484.26	21.49
X	4.00	10.00	16.00	22.00	1.00	3.98	9.77	15.55	21.49	0.98
Mean	10.00	14.00	18.00	22.00	.00	9.86	13.87	18.10	21.82	–.01

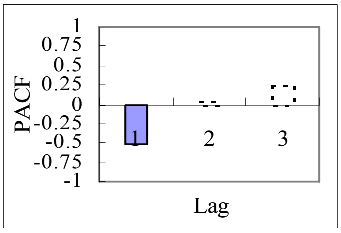
Test for stationarity

Structure	#Para.	χ^2	df	Δdf	$\Delta\chi^2$	$P_r > \Delta\chi^2_{\Delta df}$
UN (unconstrained)	10	3.292	1	--	--	--
TOEP (constrained by stationarity)	4	6.321	7	6	3.029	.805

ACF / PACF [a]

	Lag-1	Lag-2	Lag-3
ACF	$-.499^{***}$	$.278^{**}$	–.069
PACF	$-.508^{***}$	.037	.265





SCDT

			$H_0 : M_T = M_S$					Model fit	
Step	Structure	#para	χ^2	df	Δdf	$\Delta\chi^2$	$p_r > \Delta\chi^2_{\Delta df}$	$p_r > \chi^2$	AIC
0	TOEP	4	6.321	7	--	--	--	.503	–7.679
1	TOEP(1)	1	91.356	10	3	85.035	< .0001	<.0001	71.359
2	CS	2	78.358	9	2	72.067	< .0001	<.0001	60.358
3	AR(1)	2	6.466	9	2	.145	.930	.693	–11.533

Parameter estimates based on the LGM with AR(1) for $\boldsymbol{\varepsilon}$

$\hat{\gamma}_{00}$	$\hat{\gamma}_{01}$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{11}$	$\hat{\sigma}^2_{\zeta_{\eta_\alpha}}$	$\hat{\sigma}^2_{\zeta_{\eta_\beta}}$	$\hat{\sigma}_{\zeta_{\eta_\alpha}\zeta_{\eta_\beta}}$	$\hat{\phi}_1$	$\hat{\sigma}^2_{\varepsilon}$
9.92^{***}	4.07^{***}	4.05^{***}	5.97^{***}	1.16^{***}	0.97^{***}	$-.58^{***}$	$-.51^{***}$	3.80^{***}

$^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

[a] The boxes with solid line and dash line in plots represent significance and nonsignificance, respectively.

Table 2-4 (Continued)

Case 2: AR(2)

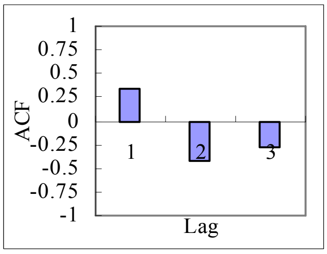
	Population covariance matrix (Equation 5)					Sample covariance matrix				
	y_1	y_2	y_3	y_4	X	y_1	y_2	y_3	y_4	X
y_1	21.00	41.44	61.58	85.87	4.00	19.95	40.03	59.21	82.23	3.84
y_2	41.44	105.00	162.44	219.58	10.00	40.03	101.16	154.81	208.63	9.55
y_3	61.58	162.44	263.00	357.44	16.00	59.21	154.81	247.72	336.54	15.10
y_4	85.87	219.58	357.44	495.00	22.00	82.23	208.63	336.54	465.81	20.75
X	4.00	10.00	16.00	22.00	1.00	3.84	9.55	15.10	20.75	0.95
Mean	10.00	14.00	18.00	22.00	.00	9.64	13.42	17.15	20.78	–.06

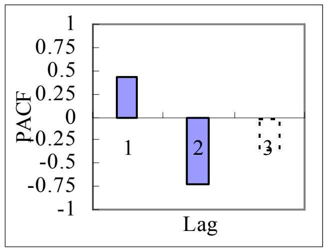
Test for stationarity

Structure	#Para.	χ^2	df	Δdf	$\Delta\chi^2$	$P_r > \Delta\chi^2_{\Delta df}$
UN (unconstrained)	10	1.192	1	--	--	--
TOEP (constrained by stationarity)	4	4.765	7	6	3.573	.734

ACF / PACF [a]

	Lag-1	Lag-2	Lag-3
ACF	.339***	–.430**	–.285†
PACF	.434***	–.730***	–.351





SCDT

			$H_0 : M_T = M_S$					Model fit	
Step	Structure	#para	χ^2	df	Δdf	$\Delta\chi^2$	$p_r > \Delta\chi^2_{\Delta df}$	$p_r > \chi^2$	AIC
0	TOEP	4	4.765	7	--	--	--	.677	–9.235
1	TOEP(1)	1	288.699	10	3	283.894	< .0001	< .0001	268.699
2	CS	2	286.899	9	2	282.134	< .0001	< .0001	268.899
3	AR(2)	3	4.853	8	1	.088	.766	.773	–11.147

Parameter estimates based on the LGM with AR(2) for $\boldsymbol{\varepsilon}$

$\hat{\gamma}_{00}$	$\hat{\gamma}_{01}$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{11}$	$\hat{\sigma}^2_{\zeta_{\eta\alpha}}$	$\hat{\sigma}^2_{\zeta_{\eta\beta}}$	$\hat{\sigma}_{\zeta_{\eta\alpha}\zeta_{\eta\beta}}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}^2_{\varepsilon}$
9.89***	4.04***	4.02***	5.98***	.0.77*	.91**	–.31	.43***	–.72***	3.70*

† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

[a] The solid-line and dash-line boxes in plots represent significance and nonsignificance, respectively.

Table 2-4 (Continued)

Case 3: MA(2)

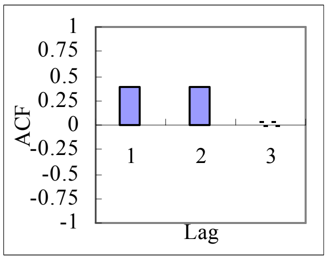
	Population covariance matrix (Equation 5)					Sample covariance matrix				
	y_1	y_2	y_3	y_4	X	y_1	y_2	y_3	y_4	X
y_1	21.00	42.20	65.42	87.50	4.00	21.39	43.49	68.41	92.15	4.20
y_2	42.20	105.00	163.20	223.42	10.00	43.49	109.43	172.26	237.71	10.65
y_3	65.42	163.20	263.00	358.20	16.00	68.41	172.26	281.29	385.47	17.29
y_4	87.50	223.42	358.20	495.00	22.00	92.15	237.71	385.47	536.50	23.92
X	4.00	10.00	16.00	22.00	1.00	4.20	10.65	17.29	23.92	1.09
Mean	10.00	14.00	18.00	22.00	.00	9.54	13.30	16.89	20.59	–.07

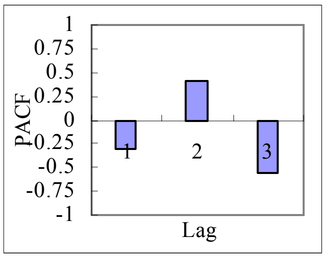
Test for stationarity

Structure	#Para.	χ^2	df	Δdf	$\Delta\chi^2$	$P_r > \Delta\chi^2_{\Delta df}$
UN (unconstrained)	10	2.345	1	--	--	--
TOEP (constrained by stationarity)	4	5.492	7	6	3.142	.791

ACF / PACF [a]

	Lag-1	Lag-2	Lag-3
ACF	$.386^{***}$	$.381^{***}$	.033
PACF	$-.296^{***}$	$.409^{*}$	$-.561^{*}$





SCDT

Step	Structure	#para	χ^2	df	Δdf	$\Delta\chi^2$	$p_r > \Delta\chi^2_{\Delta df}$	$p_r > \chi^2$	AIC
	$H_0 : M_T = M_S$							Model fit	
0	TOEP	4	5.492	7	--	--	--	.600	–8.508
1	TOEP(1)	1	26.137	10	3	20.645	< .0001	.004	6.137
2	CS	2	26.137	9	2	20.645	< .0001	.002	8.137
3	MA(2)	3	5.510	8	1	.018	.893	.702	–10.508

Parameter estimates based on the LGM with MA(2) for $\boldsymbol{\varepsilon}$

$\hat{\gamma}_{00}$	$\hat{\gamma}_{01}$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{11}$	$\hat{\sigma}^2_{\zeta_{\eta_\alpha}}$	$\hat{\sigma}^2_{\zeta_{\eta_\beta}}$	$\hat{\sigma}_{\zeta_{\eta_\alpha}\zeta_{\eta_\beta}}$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\sigma}^2_{\varepsilon}$
9.84^{***}	4.10^{***}	3.82^{***}	6.06^{***}	$1.30^{\dagger}$	$.85^{***}$	$-.53^{\dagger}$	$-.45^{***}$	$-.52^{***}$	3.94^{***}

$^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

[a] The solid-line and dash-line boxes in plots represent significance and nonsignificance, respectively.

Table 2-4 (Continued)

Case 4: ARMA(1,1)

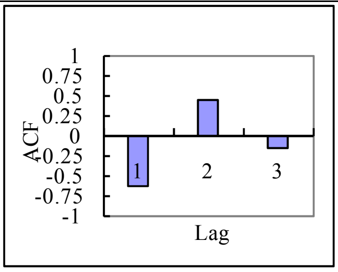
	Population covariance matrix (Equation 5)					Sample covariance matrix				
	y_1	y_2	y_3	y_4	X	y_1	y_2	y_3	y_4	X
y_1	21.00	37.64	65.43	86.79	4.00	22.26	40.17	68.79	91.51	4.24
y_2	37.64	105.00	158.64	223.43	10.00	40.17	111.57	167.48	237.19	10.60
y_3	65.43	158.64	263.00	353.64	16.00	68.79	167.48	272.55	369.92	16.73
y_4	86.79	223.43	353.64	495.00	22.00	91.51	237.19	369.92	521.52	23.16
X	4.00	10.00	16.00	22.00	1.00	4.24	10.60	16.73	23.16	1.05
Mean	10.00	14.00	18.00	22.00	0.00	10.15	14.41	18.61	22.76	0.05

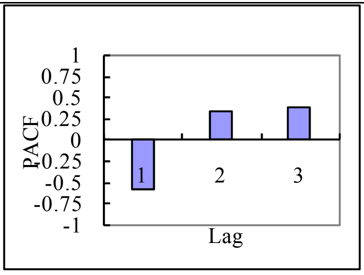
Test for stationarity

Structure	#Para.	χ^2	df	Δdf	$\Delta\chi^2$	$P_r > \Delta\chi^2_{\Delta df}$
UN (unconstrained)	10	0.819	1	--	--	--
TOEP (constrained by stationarity)	4	3.541	7	6	2.722	.842

ACF / PACF [a]

	Lag-1	Lag-2	Lag-3
ACF	$-.636^{***}$	$.448^{**}$	$-.152^{*}$
PACF	$-.592^{***}$	$.334^{***}$	$.376^{*}$





SCDT

			$H_0: M_T = M_S$					Model fit	
Step	Structure	#para	χ^2	df	Δdf	$\Delta\chi^2$	$p_r > \Delta\chi^2_{\Delta df}$	$p_r > \chi^2$	AIC
0	TOEP	4	3.541	7	--	--	--	.831	–10.459
1	TOEP(1)	1	216.950	10	3	213.409	<.0001	<.0001	196.950
2	CS	2	172.820	9	2	169.279	<.0001	<.0001	154.820
3	ARMA(1,1)	3	3.600	8	1	.059	.808	.892	–12.400

Parameter estimates based on the LGM with ARMA(1,1) for $\boldsymbol{\varepsilon}$

$\hat{\gamma}_{00}$	$\hat{\gamma}_{01}$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{11}$	$\hat{\sigma}^2_{\zeta_{\eta\alpha}}$	$\hat{\sigma}^2_{\zeta_{\eta\beta}}$	$\hat{\sigma}_{\zeta_{\eta\alpha}\zeta_{\eta\beta}}$	$\hat{\phi}_1$	$\hat{\theta}_1$	$\hat{\sigma}^2_{\varepsilon}$
9.98^{***}	3.89^{***}	4.06^{***}	5.96^{***}	1.02^{*}	$.1.09^{***}$	$-.56^{**}$	$-.55^{***}$	$.49^{*}$	3.97^{***}

$^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

[a] The solid-line and dash-line boxes in plots represent significance and nonsignificance, respectively.

Table 3-1
Results of Fitting Growth Models

	Parameter estimate				
Parameter	Model (1-A)	Model (1-B)	Model (1-C)	Model (1-D)	Model (1-E)
ϕ	0.5285**	0.2922	0.5369***	0.3960*	0.2198
γ_{00}	16.5027***	15.7508***	17.5647***	16.9042***	16.0873***
γ_{01}		16.1630***	–1.3485***	–0.4882	21.9911***
γ_{02}					–0.8883
γ_{03}					–10.3162**
γ_{10}	–1.7980***	–2.0675***	–1.9733***	–2.3632***	–2.4320***
γ_{11}		3.6353***	0.2206***	2.1771***	1.6625**
γ_{12}					1.5469***
γ_{13}					2.9072**
$\sigma^2_{\varepsilon_1}$	75.6618***	74.2979***	73.7251***	73.8213***	73.5453***
$\sigma^2_{\varepsilon_2}$	95.8847***	96.0294***	95.3825***	97.5293***	98.0433***
$\sigma^2_{\varepsilon_3}$	33.4317***	35.3216***	33.5084***	33.8775***	35.8736***
$\sigma^2_{\varepsilon_4}$	40.4354***	36.4832***	40.1770***	41.8058***	37.7967***
ρ	0.2895***	0.2923***	0.2857***	0.2985***	0.3038***
$\sigma^2_{\zeta_0}$	224.1011***	204.7596***	218.4588***	224.5650***	201.8265***
$\sigma^2_{\zeta_1}$	5.1383**	4.1179**	5.1481**	4.2153**	3.8823*
$\sigma_{\zeta_0\zeta_1}$	–9.2642*	–13.7503***	–8.6261	–9.1876**	–12.4987***
CFI	0.9904	0.9893	0.9905	0.9934	0.9901
SRMR	0.0361	0.0326	0.0327	0.0258	0.0250

Note. Dummy variable D_1 = 1 for high-tech stocks and D_1 = 0 for non-high-tech stocks. Dummy variable D_2 = 1 for stocks with ease of margin trading and D_2 = 0 for stocks without ease of margin trading. Level-1 submodel is given by

$$Y_{it} = \beta_{0i} + \beta_{1i} TIME_t + \phi EPS_t + \varepsilon_{it},\ t = 1, 2, ..., T. \quad (1)$$

Level-2 submodels are given by

$$\beta_{0i} = \gamma_{00} + \zeta_{0i},\ \beta_{1i} = \gamma_{10} + \zeta_{1i}, \quad (A)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\ D_{1i} + \zeta_{0i},\ \beta_{1i} = \gamma_{10} + \gamma_{11}\ D_{1i} + \zeta_{1i}, \quad (B)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\ size_i + \zeta_{0i}, \beta_{1i} = \gamma_{10} + \gamma_{11}\ size_i + \zeta_{1i}, \quad (C)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\ D_{2i} + \zeta_{0i}, \beta_{1i} = \gamma_{10} + \gamma_{11}\ D_{2i} + \zeta_{1i}, \quad (D)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01} D_{1i} + \gamma_{02} D_{2i} + \gamma_{03} D_{1i} D_{2i} + \zeta_{0i}, \beta_{1i} = \gamma_{10} + \gamma_{11}\ D_{1i} + \gamma_{12} D_{2i} + \gamma_{13} D_{1i} D_{2i} + \zeta_{1i}. \quad (E)$$

One-sided tests were conducted for γ_{11} in Models (1-B), (1-C), and (1-E), for γ_{13} in Model (1-E), and for error variances.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4-1
Correlation Matrices of the IS Indicators for Each Time Level

	IS indicator		
	x_1	x_2	x_3
2007/09			
x_1	1.0000	0.9980***	0.9290***
x_2		1.0000	0.9308***
x_3			1.0000
2007/10			
x_1	1.0000	0.9935***	0.8890***
x_2		1.0000	0.9033***
x_3			1.0000
2007/11			
x_1	1.0000	0.9936***	0.9436***
x_2		1.0000	0.9352***
x_3			1.0000
2007/12			
x_1	1.0000	0.9981***	0.9559***
x_2		1.0000	0.9536***
x_3			1.0000

Note. x_1, x_2 and x_3 denote, respectively, during each time period of three IS indicators (turnover rate of trading volume, turnover rate of trading value and number of transaction / outstanding shares).

Table 4-2
Results of the Exploratory Factor Analysis for the Three IS Indicators for Individual Months.

IS indicator	2007/09	2007/10	2007/11	2007/12
x_1	0.9917	0.9851	0.9934	0.9950
x_2	0.9923	0.9899	0.9906	0.9943
x_3	0.9682	0.9526	0.9732	0.9798
Eigenvalue	2.9056	2.8579	2.9152	2.9386
Proportion of total variance explained	0.9685	0.9526	0.9717	0.9795

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. One factor was extracted for x_1, x_2 and x_3 at each individual month, based on the criterion of 'Eigenvalue > 1'.

Table 4-3
Results of the Confirmatory Factor Analysis for the Three IS Indicators for Individual Months

IS indicator	2007/9	2007/10	2007/11	2007/12
x_1[a]	0.9980***	0.9934***	1.0000***	1.0000***
x_2[a]	1.0000***	1.0000***	0.9936***	0.9981***
x_3[a]	0.9308***	0.9033***	0.9436***	0.9559***
Model fit:				
CFI	1.0000	0.9954	0.9996	1.0000
SRMR	0.0000	0.0020	0.0011	0.0005
Reliability:				
Coefficient α	0.9837	0.9750	0.9837	0.9895
Composite reliability	0.9842	0.9770	0.9860	0.9898

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. Standardized coefficient α and composite reliability are reported because of different measurement units of the indicators.

[a] standardized factor loadings were presented.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4-4
Correlation Matrix of the IS Linear Trends for the Three Indicators

	Trend of x_1	Trend of x_2	Trend of x_3
Trend of x_1	1.0000	0.9897***	0.8995***
Trend of x_2		1.0000	0.8940***
Trend of x_3			1.0000

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. Standardized coefficient α and composite reliability are reported because of different measurement units of the indicators.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4-5
Results of the Exploratory Factor Analysis for the Trends of the Three IS Indicators

Variable	Factor 1
Trend of x_1	0.9874
Trend of x_2	0.9859
Trend of x_3	0.9532
Eigenvalue	2.8555
Proportion of total variance explained	0.9518

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. One factor was extracted, based on the criterion of 'Eigenvalue > 1'.

Table 4-6
Results of the Confirmatory Factor Analysis for the Trends of the Three IS Indicators

Variable	F_{β_1}
Trend of x_1[a]	0.9973***
Trend of x_2[a]	0.9923***
Trend of x_3[a]	0.9009***
Model fit:	
CFI	1.0000
SRMR	0.0000
Reliability:	
Coefficient α	0.9746
Composite reliability	0.9756

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. F_{β_1} denotes the latent factor for the trends of the three IS indicators. Standardized coefficient α and composite reliability are reported because of different measurement units of the indicators.

[a] standardized factor loadings were presented.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4-7
Results of the Convergent Validity Analysis for the Trends of the Three IS Indicators and Those of PBR and SSTR

Variable	F_{β_1}
Trend of x_1[a]	0.9960***
Trend of x_2[a]	0.9972***
Trend of x_3[a]	0.8988***
Trend of PBR	–0.0494
Trend of SSTR	–0.0214
Model fit:	
CFI	0.9958
SRMR	0.0034
Reliability:	
Coefficient α	0.6654
Composite reliability	0.7832

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. PBR represents Price / book value. SSTR represents the short-selling turnover ratio. F_{β_1} denotes the latent factor for the trends of the three IS indicators. Standardized coefficient α and composite reliability are reported because of different measurement units of the indicators.

[a] standardized factor loadings were presented.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4-8
Results of the Discriminant Validity Analysis between the IS Trend and the Trends of PBR and SSTR

Variable	Trend of PBR	Trend of SSTR	F_{β_1}
Trend of PBR	1.0000		
Trend of SSTR	0.0893*	1.0000	
F_{β_1}	–0.0132	–0.0191	1.0000

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. PBR represents Price / Earnings Ratio. SSTR represents the short-selling turnover ratio. F_{β_1} denotes the latent factor for the trends of the three IS indicators. Standardized coefficient α and composite reliability are reported because of different measurement units of the indicators. Model fit results: CFI = 0.9986; SRMR = 0.0034.
[a] standardized factor loadings were presented.
* $p < 0.1$.

Table 4-9
Results of the Convergent Validity Analysis for the IS Trends by the Three Liquidity Indicators and PBR and SSTR for the Optoelectronic Stocks

	F_{β_1}	F_{β_1}
IS trend based on		
x_1	0.9996***	0.9996***
x_2	0.9926***	0.9926***
x_3	0.9578***	0.9578***
PBR	0.3964***	0.3964***
SSTR	–0.1037	
Model fit		
CFI	0.9661	0.9661
SRMR	0.0295	0.0295
Reliability index		
Coefficient α	0.6859	0.7117
Composite reliability	0.8449	0.9225

Note. x_1, x_2 and x_3, three IS indicators, denote, respectively turnover rate of trading volume, turnover rate of trading value, and number of transaction / outstanding shares. PBR represents Price / Earnings Ratio. SSTR represents the short-selling turnover ratio. F_{β_1} denotes the latent factor for the trend based on IS indicators. Standardized coefficient α and composite reliability are reported because of different measurement units of the indicators. The stocks belonging to the optoelectronic industry are selected because they are high-tech stocks and mainly owned by noise traders (with 64.66% individual holdings). Standardized factor loadings are presented. For the model with the indicators except SSTR, SSTR is used as a control variable, and the covariance of IS trend based on SSTR and F_{β_1} = −0.1037.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4-10
The Regression of Return on IS Trend for Different Types of Stocks

Type	Regression Coefficient
Entire Market	0.0206***
Type of traders	
NT	0.0367***
RT	–0.0032
Type of traders and margin trading	
NT-EMT	0.1211**
NT-NEMT	0.0405**
RT-EMT	–0.1007
RT-NEMT	0.0832

Note. NT denotes Optoelectronics stocks, representative of those held more by individual investors (noise traders). RT denotes financial and insurance stocks, representative of those held more by judicial investors (rational investors). EMT and NEMT, denote, respectively, stocks with and without ease of margin trading. NT-EMT denotes the type with NT and EMT. NT-NEMT, RT-EMT, and RT-NEMT are defined similarly.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

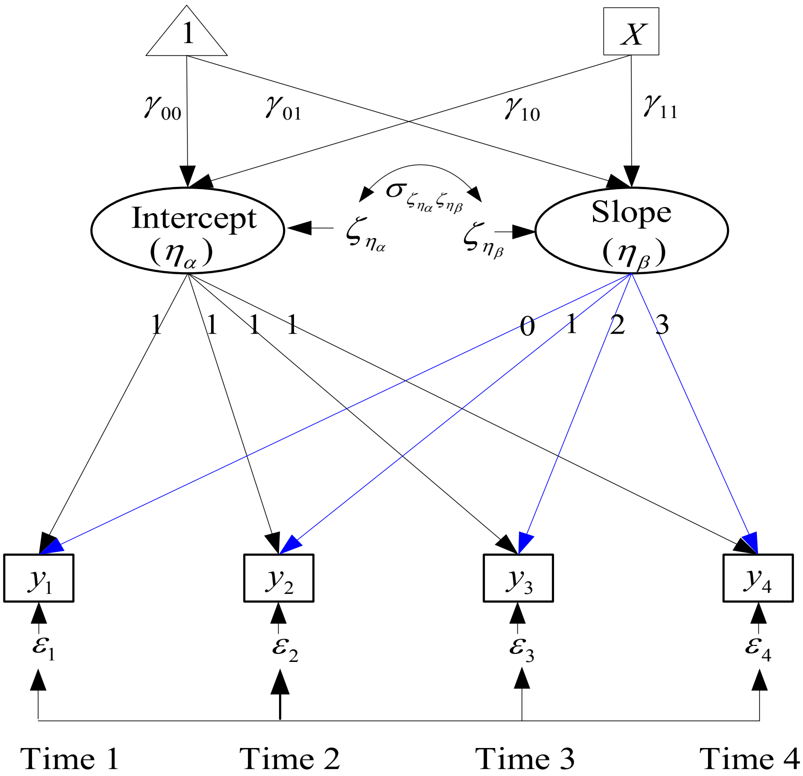


Figure 1-1. Linear latent growth model with four repeated measures and a predictor *X* (adapted from Bollen and Curran (2006, p. 128) and Preachers et al. (2008, p. 29)).

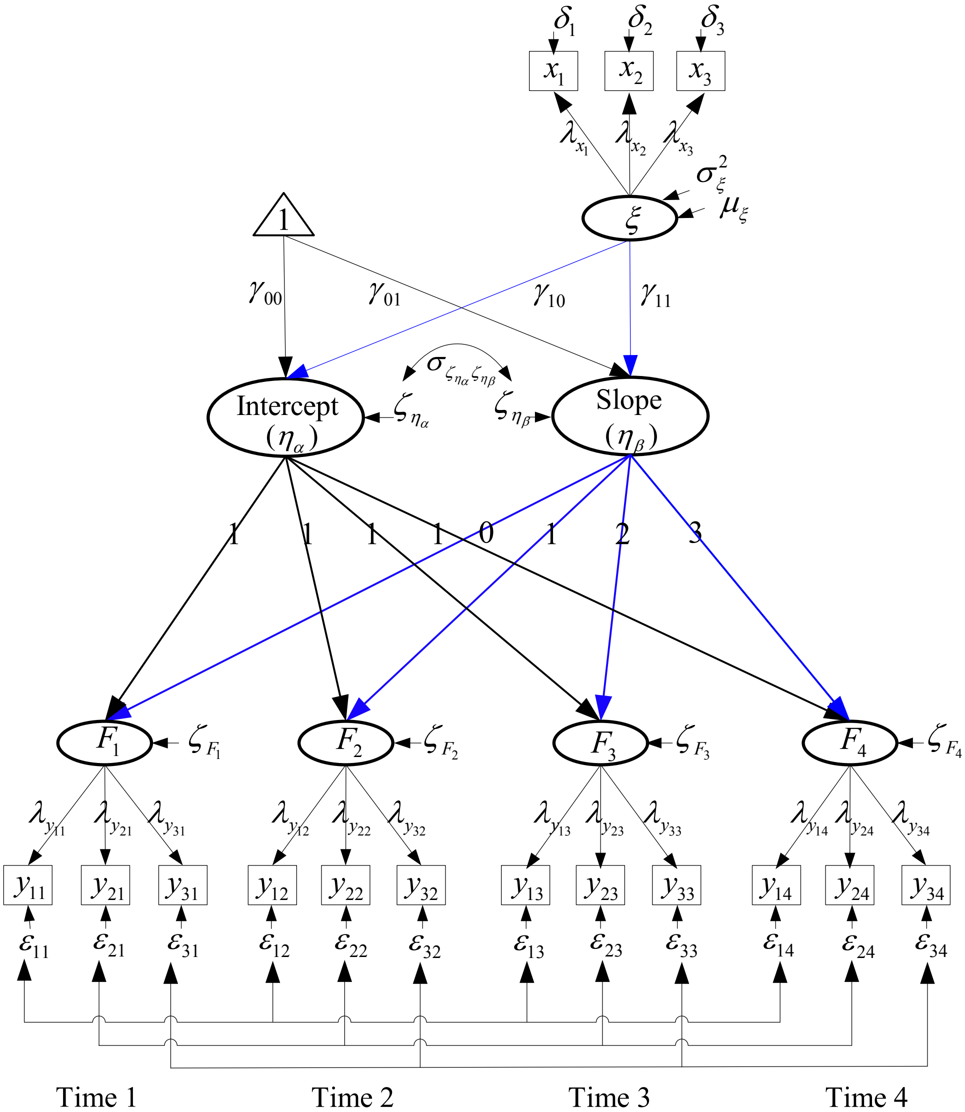


Figure 1-2. A second-order linear latent growth model with one time-invariant latent predictor and four repeated latent constructs, each measured by three indicators (adapted from Chan (1998) and Preachers et al. (2008, p. 63)).

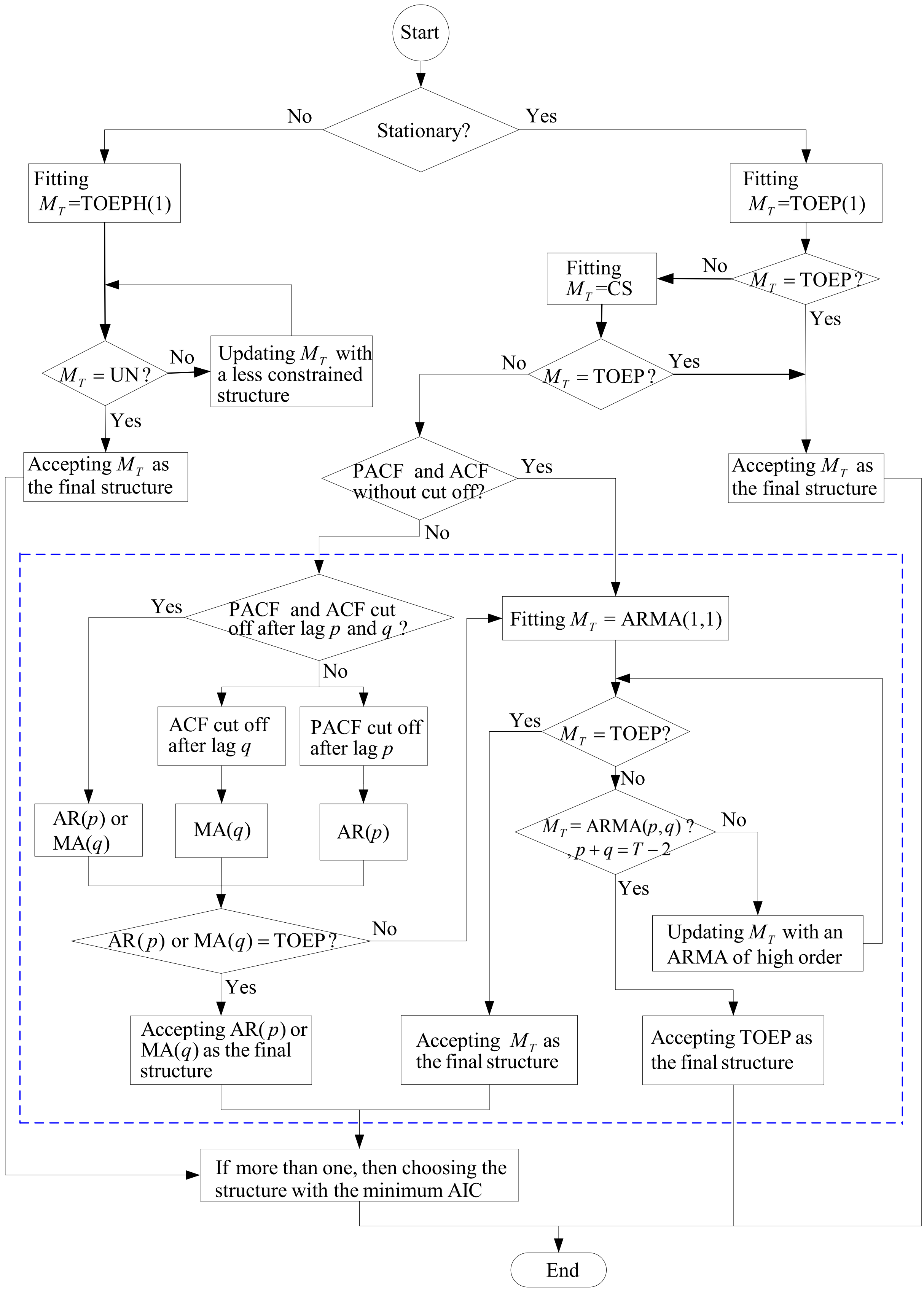


Figure 2-1. A flow chart for identifying an "optimal" level-1 error covariance structure. TOEP is the saturated stationary structure and UN is the saturated nonstationary model.

M_T denotes a temporary structure, nested within the saturated structure. M_T is updated sequentially according to the following order:

(a) for stationary structures: TOEP(1) $\rightarrow$ CS $\rightarrow$ AR(p)/MA(q) $\rightarrow$ ARMA(1,1) $\rightarrow$ (ARMA(2,1), ARMA(1,2)) $\rightarrow$ (ARMA(3,1), ARMA(2,2), ARMA(1,3)) $\rightarrow$ . . . . $\rightarrow$ (ARMA(p,q), p+q = T–2). The flow to determine the order of ARMA(p,q) is contained in the dash-line box.

(b) for nonstationary structures:
TOEPH(1) $\rightarrow$ (TOEPH(2), CSH, ARH(1)) $\rightarrow$ TOEPH(3) $\rightarrow$ . . . . $\rightarrow$ (TOEPH(T), UN(2)) $\rightarrow$ . . . . $\rightarrow$ UN(T–1).

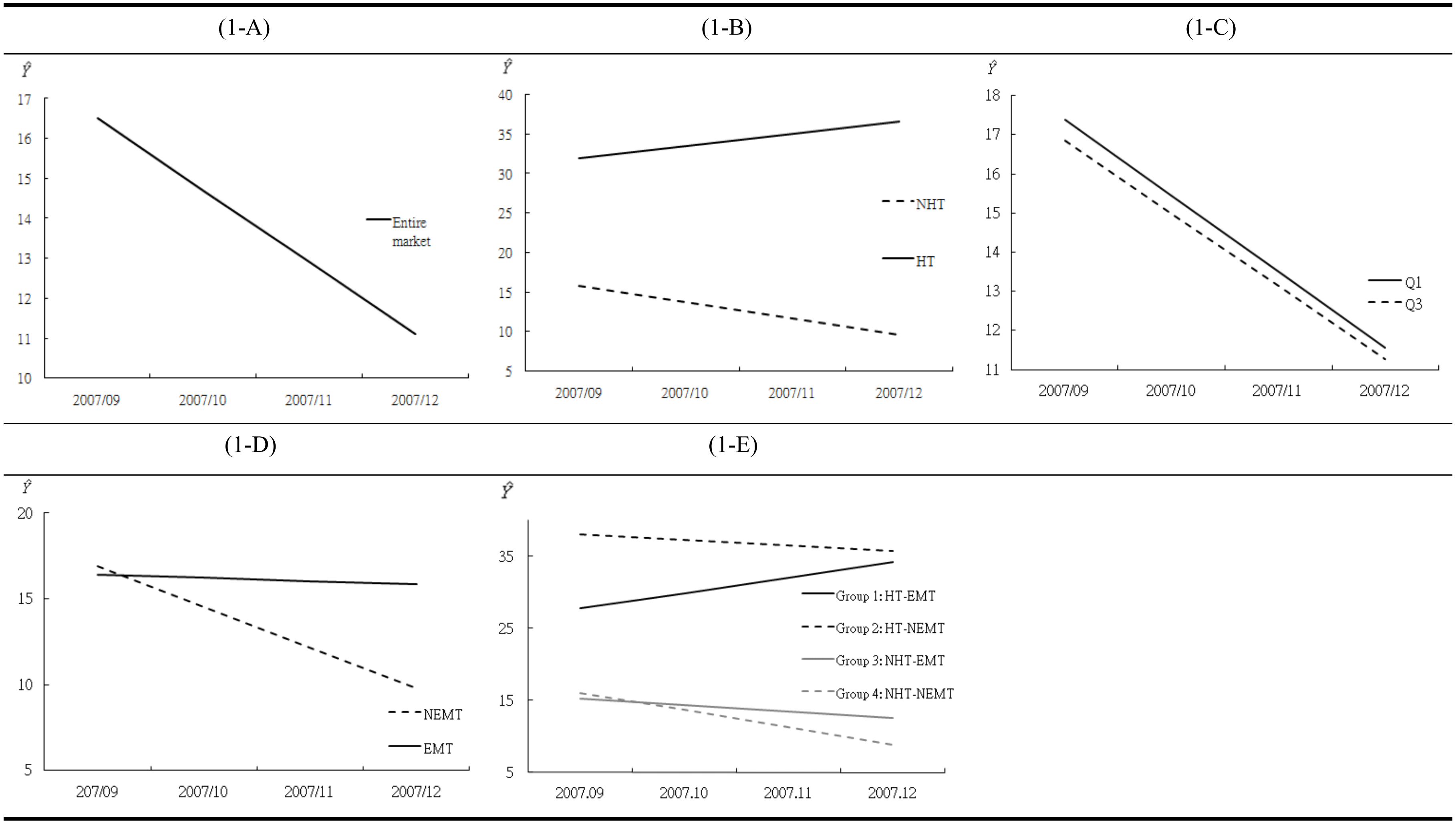


Figure 3-1. Average linear growth trend of sentiment by group. $\hat{Y}$ denotes the predicted turnover ratio after controlling for EPS.

Appendix 1-A
Sample SAS Programs for LGM by Using PROC CALIS

Program 1
A SAS Program for the LGM Shown in Figure 1-1 by Fitting ARMA(1,1) for Level-1 Error Covariance Structure

```
/* The dataset used for PROC CALIS should be a multi-variable dataset rather than a
   multi-record dataset (Singer, 1998) */

PROC CALIS UCOV AUG;
    LINEQS
        Y1 = 1 F_Alpha + 0 F_Beta + E1,
        Y2 = 1 F_Alpha + 1 F_Beta + E2,
        Y3 = 1 F_Alpha + 2 F_Beta + E3,
        Y4 = 1 F_Alpha + 3 F_Beta + E4,
        F_Alpha = GA00 INTERCEPT + GA01 X + D0,
        F_Beta = GA10 INTERCEPT + GA11 X + D1;
    STD
        E1=VARE, E2=VARE, E3=VARE, E4=VARE, X=VARX,
        D0=VARD0, D1=VARD1;
    COV
        E1 E2=COV_lag1, E2 E3=COV_lag1, E3 E4=COV_lag1,
        E1 E3=COV_lag2, E2 E4=COV_lag2, E1 E4=COV_lag3,
        D0 D1=COVD0D1;
    PARAMETERS PHI1 RHO1;
        COV_lag1=RHO1*VARE;
        COV_lag2=PHI1* COV_lag1;
        COV_lag3=PHI1* COV_lag2;
    BOUNDS
        –1. < PHI1 < 1.;
    VAR Y1 Y2 Y3 Y4 X;
    TITLE ‘Linear Growth Modeling with Four Occasions by Specifying’;
    TITLE2 ‘ARMA(1,1) for Level-1 Error Covariance Structure’;
RUN;
```

Program 2
A SAS Program for the Second-Order LGM Shown in Figure 1-2 by Fitting AR(1) for the Level-1 Error Covariance Structure Associated with Each Indicator

```
PROC CALIS UCOV AUG;
    LINEQS
       Y11 = 1 F1 + EY11,    Y21 = LY21F1 F1 + EY21,    Y31 = LY31F1 F1 + EY31,
       Y12 = 1 F2 + EY12,    Y22 = LY22F2 F2 + EY22,    Y32 = LY32F2 F2 + EY32,
       Y13 = 1 F3 + EY13,    Y23 = LY23F3 F3 + EY23,    Y33 = LY33F3 F3 + EY33,
       Y14 = 1 F4 + EY14,    Y24 = LY24F4 F4 + EY24,    Y34 = LY34F4 F4 + EY34,
       X1 = 1 F7 + EX1,      X2 = LX2F7 F7 + EX2,       X3 = LX3F7 F7 + EX3,
       F1 = 1 F_Alpha + 0 F_Beta + EZF1,
       F2 = 1 F_Alpha + 1 F_Beta + EZF2,
```

```
      F3 = 1 F_Alpha + 2 F_Beta + EZF3,
      F4 = 1 F_Alpha + 3 F_Beta + EZF4,
      F_Alpha = GA00 INTERCEPT + GA10 F7 + EZF5,
      F_Beta = GA01 INTERCEPT + GA11 F7 + EZF6,
      F7 = F7_int INTERCEPT + EZF7;
   STD
      EY11−EY14=4*VARE1, EY21−EY24=4*VARE2, EY31−EY34=4*VARE3,
      EX1=VAREX1, EX2=VAREX2, EX3=VAREX3,
      EZF1=VARZF1, EZF2=VARZF2, EZF3=VARZF3, EZF4=VARZF4,
      EZF5=VARE_Intercept, EZF6=VARE_Slope, EZF7=VARZF7;
   COV
      /* for the level-1 errors associated with indicator 1 */
      EY11 EY12=COV1_lag1, EY12 EY13=COV1_lag1, EY13 EY14=COV1_lag1,
      EY11 EY13=COV1_lag2, EY12 EY14=COV1_lag2, EY11 EY14=COV1_lag3,
      /* for the level-1 errors associated with indicator 2 */
      EY21 EY22=COV2_lag1, EY22 EY23=COV2_lag1, EY23 EY24=COV2_lag1,
      EY21 EY23=COV2_lag2, EY22 EY24=COV2_lag2, EY21 EY24=COV2_lag3,
      /* for the level-1 errors associated with indicator 3 */
      EY31 EY32=COV3_lag1, EY32 EY33=COV3_lag1, EY33 EY34=COV3_lag1,
      EY31 EY33=COV3_lag2, EY32 EY34=COV3_lag2, EY31 EY34=COV3_lag3,
      /* for the level-2 errors associated with growth factors */
      EZF5 EZF6=CZF5ZF6;
   PARAMETERS   PHI1   PHI2   PHI3;
      /* for the level-1 errors associated with indicator 1 */
      COV1_lag1=PHI1*VARE1; COV1_lag2= (PHI1**2)*VARE1;
      COV1_lag3=(PHI1**3)*VARE1;
      /* for the level-1 errors associated with indicator 2 */
      COV2_lag1=PHI2*VARE2; COV2_lag2=(PHI2**2)*VARE2;
      COV2_lag3=(PHI2**3)*VARE2;
      /* for the level-1 errors associated with indicator 3 */
      COV3_lag1=PHI3*VARE3; COV3_lag2=(PHI3**2)*VARE3;
      COV3_lag3=(PHI3**3)*VARE3;
   BOUNDS
      –1.< PHI1<1.,   –1.< PHI2<1.,   –1.< PHI3<1. ;
   LINCON   /* Weak factorial invariance across time is assumed */
       LY21F1=LY22F2, LY21F1=LY23F3, LY21F3=LY24F4,
       LY31F1=LY32F2, LY31F1=LY33F3, LY31F3=LY34F4;
   TITLE 'Second-Order Linear Growth Modeling for a Construct Measured by';
   TITLE2 'Three Indicators at Four Occasions by Fitting AR(1) for the';
   TITLE3 'Level-1 Error Covariance Structure Associated with Each Indicator';
   VAR Y11 Y21 Y31 Y12 Y22 Y32 Y13 Y23 Y33 Y14 Y24 Y34 X1 X2 X3;
RUN;
```

Appendix 1-B
More Types of Level-1 Error Covariance Structures with Four Equally Spaced Occasions

Structure ($\mathbf{\Theta}_\varepsilon$)	ECM
ARMA(p, q): $\varepsilon_t = \phi_1\varepsilon_{t-1} + \cdots + \phi_p\varepsilon_{t-p} +$ $v_t - \theta_1 v_{t-1} - \cdots - \theta_q v_{t-q}$	$\sigma_\varepsilon^2\begin{bmatrix} 1 & & & \\ \rho_1 & 1 & & \\ \rho_2 & \rho_1 & 1 & \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}$
ARMA(1,2): $\varepsilon_t = \phi_1\varepsilon_{t-1} + v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2}$, $\lvert\phi_1\rvert < 1$, $\lvert\theta_2\rvert < 1$, $\theta_1 + \theta_2 < 1$, $\theta_2 - \theta_1 < 1$; $\rho_0 = 1$, $\rho_1 = \dfrac{(-\theta_1 + \theta_1\theta_2) + [\phi_1(1+\theta_2^2) + \phi_1(\theta_1^2 - \theta_2) - \phi_1^2\theta_1(1-\theta_2) - \phi_1^3\theta_2]}{(1+\theta_1^2+\theta_2^2) - [2\phi_1(\theta_1 - \theta_1\theta_2 + \phi_1\theta_2)]}$, $\rho_2 = \dfrac{-\theta_2 - [\phi_1\theta_1(1-\theta_2) + \phi_1^2(1+\theta_1^2+\theta_2^2) - \phi_1^3\theta_1(1-\theta_2) - \phi_1^4\theta_2]}{(1+\theta_1^2+\theta_2^2) - [2\phi_1(\theta_1 - \theta_1\theta_2 + \phi_1\theta_2)]}$, $\rho_k = \phi_1\rho_{k-1}$, $k > 2$.	
ARMA(2,1): $\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + v_t - \theta_1 v_{t-1}$, $\lvert\phi_2\rvert < 1$, $\phi_2 + \phi_1 < 1$, $\phi_2 - \phi_1 < 1$, $\lvert\theta_1\rvert < 1$; $\rho_0 = 1$, $\rho_1 = [(\phi_1 - \theta_1)(1 - \phi_1\theta_1) + \phi_2^2\theta_1] / [(1-\phi_2)(1+\theta_1^2) - 2\phi_1\theta_1]$, $\rho_k = (\phi_1\rho_{k-1} + \phi_2\rho_{k-2})$, $k > 1$..	
CS: $\sigma_\varepsilon^2[\rho\ 1(t \neq t') + 1(t = t')]$, $\rho_k = \rho = \sigma_1\ / \sigma_\varepsilon^2$, $k > 0$	$\begin{bmatrix} \sigma_\varepsilon^2 & & & \\ \sigma_1 & \sigma_\varepsilon^2 & & \\ \sigma_1 & \sigma_1 & \sigma_\varepsilon^2 & \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma_\varepsilon^2 \end{bmatrix}$

Appendix 1-B (Continued)

Structure ($\boldsymbol{\Theta}_\varepsilon$)	ECM
TOEP(q) (Toeplitz with q bands, q = 1,…,4): $\sigma_{\lvert t-t'\rvert} 1(\lvert t-t'\rvert < q)$, $\sigma_0 = \sigma_\varepsilon^2, \sigma_k = \rho_k\, \sigma_\varepsilon^2, k > 0$	TOEP(1) $\begin{bmatrix} \sigma_\varepsilon^2 & & & \\ 0 & \sigma_\varepsilon^2 & & \\ 0 & 0 & \sigma_\varepsilon^2 & \\ 0 & 0 & 0 & \sigma_\varepsilon^2 \end{bmatrix}$, TOEP(2) $\begin{bmatrix} \sigma_\varepsilon^2 & & & \\ \sigma_1 & \sigma_\varepsilon^2 & & \\ 0 & \sigma_1 & \sigma_\varepsilon^2 & \\ 0 & 0 & \sigma_1 & \sigma_\varepsilon^2 \end{bmatrix}$, TOEP(3) $\begin{bmatrix} \sigma_\varepsilon^2 & & & \\ \sigma_1 & \sigma_\varepsilon^2 & & \\ \sigma_2 & \sigma_1 & \sigma_\varepsilon^2 & \\ 0 & \sigma_2 & \sigma_1 & \sigma_\varepsilon^2 \end{bmatrix}$, TOEP(4) $\begin{bmatrix} \sigma_\varepsilon^2 & & & \\ \sigma_1 & \sigma_\varepsilon^2 & & \\ \sigma_2 & \sigma_1 & \sigma_\varepsilon^2 & \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma_\varepsilon^2 \end{bmatrix}$
CSH (heterogeneous CS): $\sigma_{\varepsilon_t}\sigma_{\varepsilon_{t'}}[\rho\; 1(t \neq t') + 1(t = t')]$, $\rho_k = \rho, k > 0$	$\begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho & \sigma_{\varepsilon_2}^2 & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho & \sigma_{\varepsilon_3}^2 & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho & \sigma_{\varepsilon_4}^2 \end{bmatrix}$
TOEPH(q) (heterogeneous Toeplitz with q bands, q = 1, …,4): $\sigma_{\varepsilon_t}\sigma_{\varepsilon_{t'}}\rho_{\lvert t-t'\rvert} 1(\lvert t-t'\rvert < q)$	TOEPH(1) $\begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ 0 & \sigma_{\varepsilon_2}^2 & & \\ 0 & 0 & \sigma_{\varepsilon_3}^2 & \\ 0 & 0 & 0 & \sigma_{\varepsilon_4}^2 \end{bmatrix}$, TOEPH(2) $\begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho_1 & \sigma_{\varepsilon_2}^2 & & \\ 0 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho_1 & \sigma_{\varepsilon_3}^2 & \\ 0 & 0 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho_1 & \sigma_{\varepsilon_4}^2 \end{bmatrix}$, TOEPH(3) $\begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho_1 & \sigma_{\varepsilon_2}^2 & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho_2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho_1 & \sigma_{\varepsilon_3}^2 & \\ 0 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho_2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho_1 & \sigma_{\varepsilon_4}^2 \end{bmatrix}$, TOEPH(4) $\begin{bmatrix} \sigma_{\varepsilon_1}^2 & & & \\ \sigma_{\varepsilon_2}\sigma_{\varepsilon_1}\rho_1 & \sigma_{\varepsilon_2}^2 & & \\ \sigma_{\varepsilon_3}\sigma_{\varepsilon_1}\rho_2 & \sigma_{\varepsilon_3}\sigma_{\varepsilon_2}\rho_1 & \sigma_{\varepsilon_3}^2 & \\ \sigma_{\varepsilon_4}\sigma_{\varepsilon_1}\rho_3 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_2}\rho_2 & \sigma_{\varepsilon_4}\sigma_{\varepsilon_3}\rho_1 & \sigma_{\varepsilon_4}^2 \end{bmatrix}$.

Appendix 1-B (Continued)

Structure ($\boldsymbol{\Theta}_{\varepsilon}$)	ECM
UN(q) (UN with q bands, q = 1, …,4): $\sigma_{\varepsilon_t \varepsilon_{t'}} 1(\lvert t-t' \rvert < q)$.	UN(1) $\begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ 0 & \sigma^2_{\varepsilon_2} & & \\ 0 & 0 & \sigma^2_{\varepsilon_3} & \\ 0 & 0 & 0 & \sigma^2_{\varepsilon_4} \end{bmatrix}$, UN(2) $\begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2\varepsilon_1} & \sigma^2_{\varepsilon_2} & & \\ 0 & \sigma_{\varepsilon_3\varepsilon_2} & \sigma^2_{\varepsilon_3} & \\ 0 & 0 & \sigma_{\varepsilon_4\varepsilon_3} & \sigma^2_{\varepsilon_4} \end{bmatrix}$, UN(3) $\begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2\varepsilon_1} & \sigma^2_{\varepsilon_2} & & \\ \sigma_{\varepsilon_3\varepsilon_1} & \sigma_{\varepsilon_3\varepsilon_2} & \sigma^2_{\varepsilon_3} & \\ 0 & \sigma_{\varepsilon_4\varepsilon_2} & \sigma_{\varepsilon_4\varepsilon_3} & \sigma^2_{\varepsilon_4} \end{bmatrix}$, UN(4) $\begin{bmatrix} \sigma^2_{\varepsilon_1} & & & \\ \sigma_{\varepsilon_2\varepsilon_1} & \sigma^2_{\varepsilon_2} & & \\ \sigma_{\varepsilon_3\varepsilon_1} & \sigma_{\varepsilon_3\varepsilon_2} & \sigma^2_{\varepsilon_3} & \\ \sigma_{\varepsilon_4\varepsilon_1} & \sigma_{\varepsilon_4\varepsilon_2} & \sigma_{\varepsilon_4\varepsilon_3} & \sigma^2_{\varepsilon_4} \end{bmatrix}$

Note. 1(A) equal 1 when A is true. For example, $1(\lvert t-t' \rvert < q) = 1$ when $\lvert t-t' \rvert < q$ and 0 otherwise, $q \geq 1$. ρ_k denotes the autocorrelation coefficient at lag *k*. ρ_0 = 1. TOEP(4) = TOEP; TOEPH(4) = TOEPH; UN(4) = UN; TOEPH(1) = UN(1).

Appendix 1-C
Derivation of the Population Model-Implied Mean Vector μ and Variance-Covariance Matrix Σ of y and x for the Second-Order LGM (Figure 1-2)

Based on Equation 11, we have

$$\boldsymbol{F} = \boldsymbol{\Lambda}_y^*(\boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_\xi \xi + \boldsymbol{\zeta}_\eta) + \boldsymbol{\zeta}_F = \boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_0 + \boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi \xi + \boldsymbol{\Lambda}_y^*\boldsymbol{\zeta}_\eta + \boldsymbol{\zeta}_F,$$

$$\begin{aligned}\mathbf{y} &= \boldsymbol{\Lambda}_y \boldsymbol{F} + \boldsymbol{\varepsilon}\\ &= \boldsymbol{\Lambda}_y(\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_0 + \boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi \xi + \boldsymbol{\Lambda}_y^*\boldsymbol{\zeta}_\eta + \boldsymbol{\zeta}_F) + \boldsymbol{\varepsilon}\\ &= \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_0 + \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi \xi + \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\zeta}_\eta + \boldsymbol{\Lambda}_y\boldsymbol{\zeta}_F + \boldsymbol{\varepsilon},\end{aligned}$$

$$\mathbf{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}.$$

Therefore

$$\boldsymbol{\mu}_y = E(\mathbf{y}) = \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_0 + \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi \mu_\xi,$$

$$\boldsymbol{\mu}_x = E(\mathbf{x}) = \boldsymbol{\Lambda}_x \mu_\xi,$$

that is,

$$\boldsymbol{\mu} = \begin{bmatrix}\boldsymbol{\mu}_y\\ \boldsymbol{\mu}_x\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_0 + \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi \mu_\xi\\ \boldsymbol{\Lambda}_x \mu_\xi\end{bmatrix},$$

and

$$\begin{aligned}\boldsymbol{\Sigma}_{yy} = Cov(\mathbf{y}) &= \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi\boldsymbol{\Phi}_\xi\boldsymbol{\Gamma}'_\xi\boldsymbol{\Lambda}_y^{*\prime}\boldsymbol{\Lambda}'_y + \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Psi}_{\zeta_\eta}\boldsymbol{\Lambda}_y^{*\prime}\boldsymbol{\Lambda}'_y + \boldsymbol{\Lambda}_y\boldsymbol{\Psi}_{\zeta_F}\boldsymbol{\Lambda}'_y + \boldsymbol{\Theta}_\varepsilon\\ &= \boldsymbol{\Lambda}_y(\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi\boldsymbol{\Phi}_\xi\boldsymbol{\Gamma}'_\xi\boldsymbol{\Lambda}_y^{*\prime} + \boldsymbol{\Lambda}_y^*\boldsymbol{\Psi}_{\zeta_\eta}\boldsymbol{\Lambda}_y^{*\prime} + \boldsymbol{\Psi}_{\zeta_F})\boldsymbol{\Lambda}'_y + \boldsymbol{\Theta}_\varepsilon,\end{aligned}$$

$$\boldsymbol{\Sigma}_{xx} = Cov(\mathbf{x}) = \boldsymbol{\Lambda}_x\boldsymbol{\Phi}_\xi\boldsymbol{\Lambda}'_x + \boldsymbol{\Theta}_\delta,$$

$$\boldsymbol{\Sigma}_{yx} = Cov(\mathbf{y},\mathbf{x}) = \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi\boldsymbol{\Phi}_\xi\boldsymbol{\Lambda}'_x,$$

$$\boldsymbol{\Sigma}_{xy} = \boldsymbol{\Lambda}_x\boldsymbol{\Phi}'_\xi\boldsymbol{\Gamma}'_\xi\boldsymbol{\Lambda}_y^{*\prime}\boldsymbol{\Lambda}'_y,$$

that is,

$$\boldsymbol{\Sigma} = \begin{bmatrix}\boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx}\\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Lambda}_y(\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi\boldsymbol{\Phi}_\xi\boldsymbol{\Gamma}'_\xi\boldsymbol{\Lambda}_y^{*\prime} + \boldsymbol{\Lambda}_y^*\boldsymbol{\Psi}_{\zeta_\eta}\boldsymbol{\Lambda}_y^{*\prime} + \boldsymbol{\Psi}_{\zeta_F})\boldsymbol{\Lambda}'_y + \boldsymbol{\Theta}_\varepsilon & \boldsymbol{\Lambda}_y\boldsymbol{\Lambda}_y^*\boldsymbol{\Gamma}_\xi\boldsymbol{\Phi}_\xi\boldsymbol{\Lambda}'_x\\ \boldsymbol{\Lambda}_x\boldsymbol{\Phi}'_\xi\boldsymbol{\Gamma}'_\xi\boldsymbol{\Lambda}_y^{*\prime}\boldsymbol{\Lambda}'_y & \boldsymbol{\Lambda}_x\boldsymbol{\Phi}_\xi\boldsymbol{\Lambda}'_x + \boldsymbol{\Theta}_\delta\end{bmatrix}.$$

Appendix 3-A

Investors Structure in Terms of Trading Value in the Taiwan Stock Market

Data source: Securities and Futures Bureau, Financial supervisory Commission, Taiwan. The amounts in the table consist of both buying and selling and do not include auctions and bidding. The unit of trading value is NT billion.

Time	Juridical investors		Individual investors	
	Amount	%	Amount	%
2007/01	1,497.07	28.3	3,792.05	71.7
2007/02	801.27	36.4	1,401.82	63.6
2007/03	1,661.95	32.0	3,531.06	68.0
2007/04	1,273.98	29.9	2,989.35	70.1
2007/05	1,382.98	32.7	2,844.13	67.3
2007/06	1,806.32	30.2	4,174.72	69.8
2007/07	2,479.14	24.8	7,506.08	75.2
2007/08	2,146.47	30.0	4,996.94	69.9
2007/09	1,486.34	29.9	3,481.86	70.1
2007/10	2,187.23	31.2	4,826.65	68.9
2007/11	2,115.48	36.2	3,726.78	63.7
2007/12	1,531.88	34.8	2,867.83	65.2

Appendix 3-B

J. P. Morgan Investor Confidence Indices and Volume of Initial Public Offerings from January, 2005 through December, 2008 in the Taiwan Stock Market

Year	2005		2006		2007		2008	
Month	J. P. ICI	Vol. IPOs	J. P. ICI	Vol. IPOs	J. P. ICI	Vol. IPOs	J. P. ICI	Vol. IPOs
1	99.600	0.201	108.700	0.000	114.500	0.084	104.200	1.360
2	–	0.062	–	0.034	–	0.000	–	0.062
3	115.100	0.038	100.400	2.711	108.900	0.000	119.400	0.174
4	–	0.000	–	0.691	–	0.000	–	0.433
5	105.600	0.078	115.900	0.000	106.100	0.367	132.100	0.000
6	–	0.000	–	0.000	–	0.077	–	0.235
7	106.000	0.000	102.800	0.000	117.800	0.124	106.600	0.164
8	–	3.933	–	0.062	–	0.775	–	0.000
9	94.300	0.000	88.600	0.202	109.900	0.092	81.800	0.000
10	–	0.333	–	2.530	–	1.134	–	0.225
11	93.900	2.058	99.500	0.107	104.500	1.264	–	0.131
12	–	0.186	–	0.208	–	1.701	–	1.203

Note. J. P. Morgan investor confidence indices (J. P. ICI) are available at https://www.jpmrich.com.tw /docs/jf/faith/index.html. J. P. Morgan has stopped publishing the J. P. ICI information since November 2008. The volume of IPOs (in billion shares) for the Taiwan stock market is obtained from the Taiwan Stock Exchange Corporation.

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性)、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

□ 達成目標

√ 未達成目標（請說明，以 100 字為限）

□ 實驗失敗

□ 因故實驗中斷

√ 其他原因

說明：

原計畫係於2008年底提出申請，但第一部分相關理念後來已由 Grim, Pianta, and Konold 以 "Longitudinal Multitrait-Multimethod Models for Developmental Research" 為題發表於 *Multivariate Behavioral Research* (Vol. 44, p.233-258, 2009)；另原計畫第二部分擬利用多屬性多方法(MTMM)模型實證投資人情緒，惟其中P/B ratio及SSTR指標(使用量測方法B)無法通過信效度檢測，故MTMM模型不適用於第二部分實證，因此，計畫須予修正。

2. 研究成果在學術期刊發表或申請專利等情形：

論文：√ 已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 □無

技轉：□已技轉 □洽談中 □無

其他：(以 100 字為限)

本研究第一部分之成果已撰文 "Using SAS PROC CALIS to Fit Level-1 Error Covariance Structures of Latent Growth Models"，獲 *Behavior Research Methods* 接受刊登；另文"Identifying the Order of the ARMA Process for Level-1 Errors in Latent Growth Modeling" 已發表於 2011 年 8 月 15 日在美國德州 San Antonio 舉行之 The Academy of Management 2011 Annual Meeting，該文經修訂後已投稿至 *Psychological Methods*。本研究第二部分之成果已撰二文"Assessing the Change in Investor Sentiment over Time"以及"On the Measurement of the Change in Investor Sentiment over Time Based on Individual Stocks"，前者已投至 *Quantitative Finance*，後者亦將於近期內投至 *Journal of Financial Market*。

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

修正後計畫第一部分研究重點改置於潛伏成長模型中 level-1 誤差共變異結構之鑑定，具體說明及示範如何利用 SAS PROC CALIS 進行 level-1 誤差共變異結構之配適，並提出一有效的誤差共變異結構之鑑定方法；第二部分研究重點改置於利用傳統 LGM 從事投資人情緒之變化趨勢實證研究，並基於個股資料提出投資人情緒水平與情緒趨勢之量測方法,包括評估指標之建立以及信度與效度分析。

潛伏成長模型中 level-1 誤差共變異結構被視為 nuisance parameters，其認定難有理論依據，故需藉助於 specification search，本研究所提鑑定方法係利用卡方差異檢定以及自我相關與偏自我相關檢定，所鑑定結果合乎模型配適與精簡之訴求，並有效降低模型誤設的可能性，在方法上具學術與應用價值；另誤差共變異結構之複雜性常需搭配較細緻的軟體操作指令，為因應使用上彈性所需，並提升 specification search 之執行效率，本研究提供 SAS PROC CALIS 之相關語法，在操作技術上具應用價值，有利於潛伏成長模型之實證研究。本研究第二部分有關投資人情緒變化趨勢實證研究係觀察證券市場投資人情緒波動現象，了解股價變化之前因，投資人情緒之分析有利於掌握證券市場干預時機，在市場管理上具參考價值；另所提投資人情緒評估指標之建立在方法上亦具貢獻。

誌　謝

本研究蒙國科會經費補助 (計畫編號：NSC 98-2410-H-009-010-MY2)，研究過程亦蒙國立交通大學多方行政支援，謹此一併申謝！