Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2010 Proceedings

International Conference on Information Systems (ICIS)

2010

Latent Growth Modeling in IS Research: Basic Tenets, Illustration, and Practical Guidelines

Paul A. Pavlou
Temple University, pavlou@temple.edu

Eric Zheng
University of Texas at Dallas, ericz@utdallas.edu

Bin Gu
University of Texas at Austin, bin.gu@mccombs.utexas.edu

Latent Growth Modeling in IS Research: Basic Tenets, Illustration, and Practical Guidelines

Completed Research Paper

Paul A. Pavlou
Temple University
Philadelphia, Pennsylvania, USA
pavlou@temple.edu

Eric Zheng
University of Texas at Dallas
Texas, Dallas, USA
ericz@utdallas.edu

Bin Gu
University of Texas at Austin
Austin, Texas, USA
bin.gu@mccombs.utexas.edu

Abstract

This paper introduces Latent Growth Modeling (LGM) as a feasible method for analyzing longitudinal data to understand the process of change over time. Given the need to go beyond cross-sectional models, explore longitudinal Information Systems (IS) phenomena, and test IS theories over time, LGM is proposed as a complementary method to help IS researchers propose and evaluate time-centric hypotheses and make longitudinal inferences.

The paper first describes the basic tenets of LGM and offers guidelines for using LGM in IS research, including framing hypotheses with time as a central component and implementing LGM models to test these hypotheses. The application of LGM in IS research is illustrated by modeling the longitudinal relationship between two IT variables (IT infrastructure and IT labor) and firm performance with 2001-2004 data from Fortune 1000 firms. Comparisons with other methods for analyzing longitudinal data reveal the advantages of LGM for studying time-dependent relationships and growth patterns.

Keywords: Latent growth modeling, LGM, Longitudinal data, Time-dependent hypotheses.

Latent Growth Modeling in IS Research: Basic Tenets, Illustration, and Practical Guidelines

Completed Research Paper

Introduction

Despite the call for going beyond cross-sectional models and capturing longitudinal effects in Information Systems (IS) research, empirical IS studies are dominated by cross-sectional data and static analysis. Notably, out of 286 published empirical studies in MISQ and ISR over the last decade (2000-2010), 89% relied on cross-sectional data[1]. Besides, most hypotheses are proposed in simple static terms, such as "X is associated with Y." Hypotheses stated in a cross-sectional manner are difficult to falsify, and they may lead to overestimates of the amount of support for a hypothesis (Mitchell and James, 2001). Even when the data are indeed longitudinal, hypotheses are often presented in a static relationship without specifying how the relationship changes over time, and the longitudinal aspect of the data is mainly used to better identify the static relationship. This is evident, for example, in the commonly-used fixed-effects model in panel data analysis, which leverages within-variations to estimate the relationship between the dependent and independent variables (Greene, 2007). This treatment controls for individual variations and common time variations, but assumes that the underlying relationship between the dependent and independent variables is individual-specific and time invariant (i.e., static). Many other approaches directly incorporate time as an independent variable. However, these approaches only acknowledge that the value of the variable of interest is different at different times, they do not capture the underlying forces that affect a variable's longitudinal trajectory and how the relationship between the two variables changes over time. As noted by Massey and Montoya (2006, p. 111) "merely establishing that time matters or that the relationship possesses a temporal process is not sufficient to advance our understanding." Furthermore, typical IS studies using SEM tools (e.g., LISREL, PLS) also do not substantively account for longitudinal data, and virtually all these studies are based on cross-sectional analyses. Among the 181 papers that employ SEM tools in ISR and MISQ from 2000-2010 (representing 63% of all IS empirical papers), only 8 papers (about 4%) have any longitudinal analyses.

From a theoretical perspective, the proper analysis of change patterns over time is quintessential for understanding the dynamic relationship among IS variables. Dynamic relationship refers to the relationship between the current state and the rate of change of a given variable or those with other variables. In fact, most IS theories are intrinsically rooted in the change of variables and the relationships among the variables over time. For example, IT diffusion theories (Fichman and Kemerer, 1999; Straub, 1995) prescribe how IT products (e.g. e-mail) penetrate the market over time. Information processing theory posits that firms must change their information processing capabilities to match their information processing needs, a process that persists over time (e.g., Daft and Lengel, 1986; Tushman and Nadler, 1978). Economic growth theory, which has been extensively used in IS research, suggests that the fundamental driving factor of economic growth is cumulative technology changes (Findlay 1978). Also, IT adoption theories (e.g., Davis, 1989; Venkatesh and Davis, 2000; Zhu and Weyant, 2003) focus on how users and organizations gradually adopt IT tools over time. Similarly, trust building theories (e.g., Gefen et al., 2003; Pavlou and Gefen, 2004; Pitariu and Ployhart, 2010) examine the longitudinal development of trust at different stages, albeit research has mostly viewed trust as a static concept, largely due to methodological limitations (Lewicki et al., 2006). Research in organizational learning (e.g., March 1991, Levinthal and March 1993) also examines how firms react to changes in the environment, such as that knowledge exploitation leads to superior performance in stable environments while the pursuit of knowledge exploration leads to superior performance in more turbulent environments (Jansen et. al., 2006). In sum, since IS theories do focus on the change of variables and relationships over time, it is important to enhance the portfolio of IS researchers with data analysis tools that explicitly model change across such variables over time.

To go beyond static hypotheses and data analysis methods, and to illustrate the value of exploring dynamic IS phenomena and testing IS theories over time, this paper proposes *Latent Growth Modeling* (*LGM*) as a viable data

[1] These statistics are based on the two leading IS journals: ISR from 2000- December 2010 and MISQ from 2000-March 2010.

analysis tool for analyzing dynamic relationships using longitudinal data, particularly for IS theories that focus on modeling change over time and making dynamic inferences about IS phenomena. LGM is a relatively recent data analysis method that aims to model dynamic relationships both between variables and within individual variables, focusing on the pattern of change over time (e.g., Duncan and Duncan, 1995; Duncan, Duncan and Strycker, 2006; Curran et al. 2004; Muthen, 1993). The pattern (or shape) of change renders information on latent trajectories, the underlying longitudinal process that is not directly observed (Bollen and Curran, 2006, p. 2). In general, latent growth models (Whiteman and Mroczek 2007, p. 78-80) advance our understanding of change over time by breaking down the variance into two major components: *within-individual* and *between-individual* variation. The within-individual variation is the degree of change within individuals, such as firms in our study, also referred to as *Level 1* model. Individuals may also differ significantly in terms of their trajectory over time. A random effects model, referred to as the *Level 2* model, is used to model the heterogeneity between individuals. Notably, Curran and Willoughby (2003, p. 603) make an important observation that "latent growth models might be viewed as residing at an intersection between variable-centered and individual-centered analysis".

The application of LGM in this study is illustrated in the context of modeling the dynamic relationship between information technology (IT) and firm performance over time. Despite the importance of specifying the role of IT in firm performance over time to show the long-term performance effects of IT, the dynamic relationship between IT and firm performance is still not very well understood (Bryjolfsson and Hitt, 2003). IS researchers have noted that it takes a significant amount of time for IT investments to affect a firm's bottom line, and these investments often need time to materialize. This time lag is potentially due to implementation and adoption issues, learning, and the need to invest in complementary IT and other assets (e.g., Bresnahan et al., 2002; Brynjolfsson, 1993; Devaraj and Kohli, 2000). As such, the impact of IT investment on firm performance is not instant or static. Rather, IT investments influence a firm's long-term performance. Traditional longitudinal analysis tools face several challenges in modeling such a long-term dynamic relationship. To overcome this void, we illustrate the use of LGM to model the dynamic role of IT (IT infrastructure and IT labor) in firm performance using 2001-2004 data from publicly-traded Fortune 1000 firms taken from the Harte-Hanks CI dataset, matched with Compustat data on firm performance.

Our results reveal several interesting longitudinal patterns that have not been addressed in the IS literature.[2] First, the initial level of IT infrastructure (in year 2001) does not have a significant impact on the initial level of firm performance, but it has a marginally significant positive effect (at the $p<0.1$ significance level) on the longitudinal growth rate of firm performance. This implies that a higher level of initial IT infrastructure can ultimately help improve firm performance down the road. Second, the slope (or growth rate) of IT infrastructure is *not* significantly associated with the slope of firm performance over time, implying that the incremental investment in the IT infrastructure over the years may not further accelerate firm performance. Third, the initial level of IT labor, on the other hand, is found to be significantly associated with the initial level of firm performance, but it has a negative impact on the slope of firm performance. This suggests that a higher level of initial IT labor constrains firm growth, potentially due to the lack of scalability in labor and the law of diminishing marginal returns. Fourth, the slope of IT labor was *not* found to have positive impact on the slope of firm performance, which is not surprising at a period that experienced large-scale IT downsizing (2001-2004). Finally, as expected, the initial levels of IT infrastructure and IT labor were found to negatively impact the longitudinal slope of IT infrastructure and IT labor, respectively.

This paper also compares LGM with other methods for longitudinal and structural analysis, such as panel data analysis and Structural Equation Modeling (SEM). Our results reveal the added advantages of LGM, (1) directly modeling change over time, (2) examining the effect of initial conditions on change (trajectory) over time, (3) identifying the factors that influence growth, and (4) examining the interplay among the change across variables.

Taken together, the paper makes four major contributions: First, it describes LGM as a complementary new method to model change patterns over time for individual variables and relationships among variables. Second, it illustrates the LGM method by modeling the relationship between IT and firm performance over time. Third, it compares LGM with other data analysis methods commonly used in IS research and identifies its advantages and disadvantages. Finally, it provides practical guidelines for IS researchers to properly use different variations of LGM and how to implement these variations in commercial software, such as SAS, by offering guidelines on effectively using LGM.

[2] It is important to note that these results from the 2001-2004 period is shown as an illustration of the proposed LGM method, and it is beyond the scope of this study to develop a generalizable theory on the longitudinal effects of IT on firm performance. Future research (potentially using LGM) could develop formal IS theories on this.

The rest of the paper is organized as follows. Section 2 discusses conceptual modeling associated with longitudinal models and discusses how to propose longitudinal hypotheses. Section 3 offers an illustrative example by applying various LGM models to examine the relationship between IT and firm performance and presents the study's results. Finally, Section 4 discusses the study's contributions and implications for conducting longitudinal research in IS.

Latent Growth Modeling

The Need of Latent Growth Models

Mitchell and James (2001) point out that three facets of time need to be accounted for to precisely describe a longitudinal relationship (between *X* and *Y*) - *time lag*, *duration*, and *rate of change*. In other words, how long *Y* occurs after *X* occurs, how long does the relationship last, and what is the rate of change? Pitariu and Ployhart (2010) further argue that understanding the *shape*[3] (e.g. linear or non-linear) of the longitudinal relationship is also necessary because it is unlikely that any two variables would have an identical relationship over time. Therefore, hypotheses stated in a longitudinal form using these parameters are more falsifiable and offer a more rigorous and informative tests of theories than simple "X is associated with Y" hypotheses (Mitchell and James, 2001).

Traditionally, longitudinal processes of change were analyzed based on repeated measures analysis of variance (ANOVA) and regression and analysis of covariance (ANCOVA) frameworks where an individual variable is observed repeatedly over time. For example, panel data analysis always starts with ANOVA to understand how much of the variation in *y* comes from cross sectional or longitudinal variations (Jackman, 2009, p. 317). However, these methods have important limitations resulting from assumptions that are often violated in empirical research. According to Raykov and Marcoulides (2008), these are: (1) *homogeneity of variance*, that is covariance matrices across levels of between-subject variables; (2) *sphericity*, which implies the same inter-correlation among the repeatedly-measured variables; and (3) *perfectly measured covariate*(*s*). Assumption (3) is often hard to attain in most empirical research since it is rare to have access to measures that do not generally contain measurement error.

Besides, repeated-measure ANOVA methods are essentially indifferent to time, in the sense that they produce the same statistical test results even if one were to 'reshuffle the order of the assessment occasions' (Gottfried et al. 2007, 2009). For example, if a researcher were to analyze the study by letting time "run backward." This becomes evident in the case of the commonly-used fixed-effects model in panel data analysis when time itself is treated as a fixed effect, essentially that time merely serves as the occasion where repeated measures are observed. The possible change effects, such as the latent growth rate of the variable are overlooked in panel data analysis. Another approach taken in longitudinal studies uses lag variables or auto-regression to deal with the effect of time (Greene, 2007). However, the mechanism governing the change in this approach is the past value and the values of the covariates (Bollen and Curran, 2006, p. 2), and thus the underlying trajectory of change is still not fully captured.

These limitations for modeling change over time can be overcome by using an alternative method – LGM. In LGM, a latent variable can be viewed as a random variable with individual realizations in a given sample (or population, for that matter), which are not observed. Models used in applications of LGM are typically developed in terms of latent variables, and they can be designed to reflect important aspects of a wide variety of longitudinal processes (Qureshi et al. 2008, Kher et al. 2009). LGM models are also less restrictive than repeated measures ANOVA methods because they do not make any of the above three assumptions. For example, similar to SEM tools, the use of latent variables in LGM directly models measurement errors in variables, addressing limitation (3). Additionally, methods based on repeated-measures ANOVA capture change at the aggregate level, while the notion of individual difference in change is a core strength of the LGM method (Whiteman and Mroczek, 2007). Notably, LGM directly models individual change over time by allowing different initial state and rate of change for each individual entity (e.g., Fortune 1000 firms in our analysis).

LGM Formulation 1: Unconditional Models

In a basic LGM, two new variables are specified to represent aspects of change (Preacher et al., 2008, p. 6): First, the *intercept* factor represents the level of the outcome measure *y*, at the initial time (at which the time variable *t* equals zero). Second, the *slope* factor represents the linear rate at which *Y* changes. Specifically, let y_{it} represent the

[3] *Shape* refers to the specific functional form or temporal trajectory of a relationship over time (Bollen and Curran, 2006).

series of repeated measurements of an individual *i* over different time periods *t*.[4] A simple longitudinal model equation describing an individual *i*'s development over the repeated measures (also called *Level 1* model) can be written separately for each individual (to simplify matters, we only present a single general form equation, but it should be clear that a separate Equation 1 is needed to model the change process for each *i*):

$$y_{it} = \alpha_i + \beta_i \lambda_t + \varepsilon_{it} \tag{1}$$

where $\underline{\alpha_i}$ is the initial status of an individual *i* measured at t_o (i.e., the intercept), and β_i is the slope or the shape of the change trajectory (change in y_{it} between consecutive measurements), λ_t corresponds to the measured time points and a common coding is to have λ_1=0, λ_2=1 … and so on (Bollen and Curran, 2006, p.20), while ε_{it} represents the model residual for each individual. Because α_i and β_i are random variables (coefficients), these model parameters are represented by an overall group mean intercept (μ_α) and mean slope (μ_β), plus the component of individual intercept variation ($\varepsilon_{i\alpha}$) and slope variation ($\varepsilon_{i\beta}$) respectively, as indicated by the following *Level 2* model equations:

$$\alpha_i = \mu_\alpha + \varepsilon_{i\alpha}$$
$$\beta_i = \mu_\beta + \varepsilon_{i\beta} \tag{2}$$

In the simplest form of LGM specified above, no other predictors are assumed to account for the variation in the specific parameters of the trajectories. In this case, the Level 2 model is also called an unconditional model. Figure 1 depicts the simple LGM specified by equations (1) and (2) with four time points. For ease of presentation, the error terms are omitted from Figure 1. Notice that the loadings on α are fixed to be 1 for all four time points, while those of slope β are fixed as λ_t=0, 1, 2, 3 for *t* = 1, 2, 3, 4. This common coding scheme is referred to as the intercept-slope approach. The equally-spaced units reflect equal time passage between assessments, and beginning the coding with zero allows for the intercept factor to reflect the mean value of y at the first period (Bollen and Curran, 2006, p. 36). For this intercept-slope coding of time, LGM needs measurements of at least three time periods.

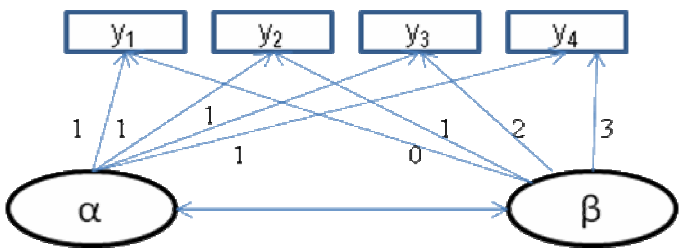

Figure 1: An Unconditional Latent Growth Model with four time points

Moreover, the change *within* individual variables over time may reveal some interesting longitudinal patterns (Marcoulides and Hershberger 1997, Gottfried et al. 2007). For example, as we demonstrate in this study, longitudinal changes in IT investments on IT infrastructure and IT labor may have notable intra-individual variation. However, most MIS research has overlooked such within-variable changes over time.

LGM Formulation 2: Conditional Models

The addition of variables that can potentially be used to subsequently predict the intercept and slope requires the examination of a so-called *conditional* latent growth model (Bollen and Curran, 2006). Figure 2 presents a path diagram of a longitudinal model with two specific predictors, x_1 and x_2. Generally the covariates considered in this manner are time invariant.[5] The general form is:

$$\alpha_i = \mu_\alpha + \gamma_{\alpha 1i} x_{1i} + \gamma_{\alpha 2i} x_{2i} + \varepsilon_{i\alpha}$$
$$\beta_i = \mu_\beta + \gamma_{\beta 1i} x_{1i} + \gamma_{\beta 2i} x_{2i} + \varepsilon_{i\beta} \tag{3}$$

[4] For example, in our study, index *i* corresponds to each observed individual firm and *t* corresponds to the 2001-2004 period.
[5] For including time-variant covariates, please refer to Bollen and Curran (2006, p. 192-197) for a more detailed discussion. The main idea is to add a time-variant covariate Z_{it} to the Level 1 equation, which becomes: $y_{it} = \alpha_i + \beta_i \lambda_t + \gamma_t Z_{it} + \varepsilon_{it}$

where x_{1i} and x_{2i} are the two predictors of the Level and Shape (LS) factors, and $\gamma_{\alpha 1i}$, $\gamma_{\alpha 2i}$, $\gamma_{\beta 1i}$, and $\gamma_{\beta 2i}$ are the coefficients for the predictors of the LS factors. These coefficients can be interpreted as coefficients in regression that reflect the change in the dependent variable for a 1-unit change in a predictor while holding the other predictors constant (Bollen and Curran, 2006; Duncan et al., 2006). Model parameters captured by the group mean intercept (μ_α) and mean slope (μ_β) are the values obtained when x_{1i} and x_{2i} are equal to zero. In sum, the advantage of LGM is its ability to estimate both the level and rate of change of the variables of interest while controlling for covariates.

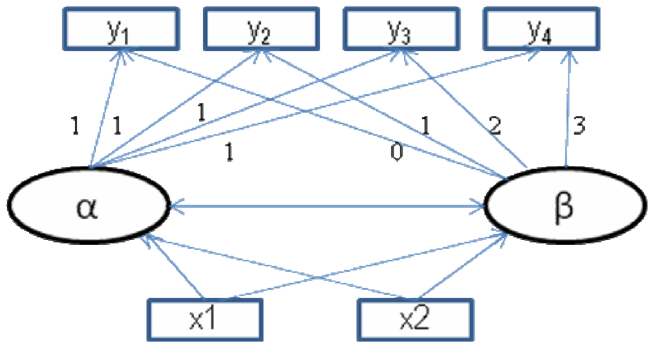

Figure 2: A Conditional Latent Growth Model with two Predictors

LGM Formulation 3: Multivariate Models

When there are multiple time-varying variables and each of them follows a latent growth process, a multivariate LGM needs to be introduced to examine the inter-play of these latent growth processes. Without loss of generality, suppose we are interested in two such variables, *Y* and *Z*, both of which are observed overall four time periods. Moreover, suppose *Y* is estimated with a conditional LGM with several time-invariant covariates *X*, as shown in the upper part of Figure 3; and *Z* is estimated with an unconditional LGM, as shown in the lower part of Figure 3.

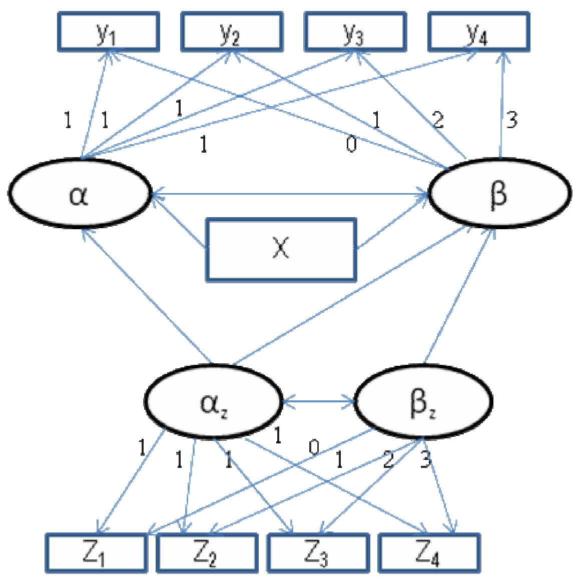

Figure 3: A Bivariate Latent Growth Model

A number of potential associations between the two latent variables can be examined, such as: "how does the initial level of Z affect the initial level and slope of Y" or "how are the slopes of the two variables associated"? Figure 3 represents the scenario where the intercept of *Y* is affected by the initial level of *Z*, and the slope of *Y* is affected by both the level and slope of *Z*. Many relationships among the level and slope of these two variables can be specified. Figure 3 show an example of specifying the relationships for a tri-variate LGM for three time-varying latent growth variables.

LGM Formulation 4: Nonlinear Latent Growth

The previous formulations focused on *linear* latent growth models, implying a constant rate of change across time. However, more complex trajectories, such as non-linear growth may emerge in practice. In LGM, non-linear trajectory is achieved through the coding of time in λ_t. Nevertheless, we need to know the specific non-linear form of the rate of change to correctly specify a non-linear LGM. Though there is only one way for a trajectory to be linear, there are infinite non-linear ways. Thus, not much can be said about the general non-linear form of LGM. Bollen and Curran (2006, Chapter 4) discussed quadratic, cubic, and exponential trajectories, and we refer readers for these specific treatments. In general, non-linear models require measurements at least four time periods. Herein we adopt a more general approach that can model the change process, regardless of the particular functional form. This general approach is called the Level/Shape (LS) strategy, which does not a priori assume a particular trajectory (Raykov and Marcoulides, 2008). In the LS modeling strategy, λ_t (referred to as the loadings on β_i) is normally fixed to be 0 for the first time occasion and 1 for the last time occasion, respectively. Unlike the coding in the slope-intercept approach introduced earlier, the λ_t's in the middle occasions (other than the first and last) are not fixed and were treated as free parameters to be estimated from the data.[6] This coding mechanism ensures that the Slope factor is interpreted as a change factor confined within 0 and 1. Freeing the loadings of the remaining occasions implies that the loadings reflect the part of the total change (100%) that occurs between the first and the last measurements (Duncan et al., 1999; Raykov and Marcoulides, 2008). Figure 4 shows the path diagram of this time coding scheme.

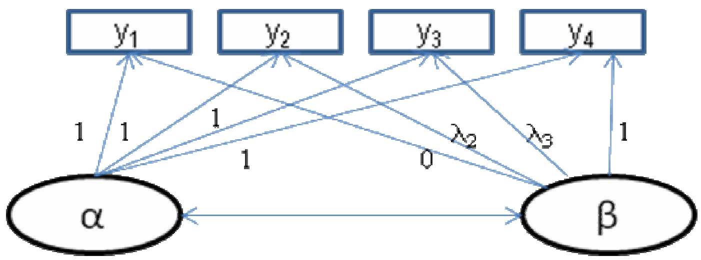

Figure 4: Level and Shape (LS) Time Coding Scheme

Fitness Assessment

The overall model fitness can be evaluated with a number of indices. These include the overall goodness-of-fit (GFI), adjusted goodness-of-fit (AGFI), overall χ^2, standardized mean squared residual (SRMR), comparative fit index (CFI), and normal fitness index (NFI). These indices are widely used (e.g., Raykov and Marcoulides, 2008, Bollen and Curran, 2006). Detailed criteria for evaluation of model fit based on these indices can be found in Byrne (1998), Hu and Bentler (1999), Marcoulides and Hershberger (1997), and Raykov and Marcoulides (2006; 2008). There is no general consensus on which criterion is better or which set of criteria should the researcher adopt. The most commonly used fit indices include AGFI>0.9, a CFI > 0.9, and SRMR<0.05.

Alternative Models

An alternative to LGM for modeling longitudinal data is the class of model under the broad umbrella of *multi-level modeling* (Goldstein, 1995). Multi-level models, such as the popular hierarchical linear models (HLM), allow hierarchical partitioning of variance (Raudenbush and Bryk, 2002). This is an important ability when analyzing hierarchically nested data, such as firms are nested in an industry (Mithas et al. 2007) and departments nested within a division (Boh et al., 2007). MLM and LGM stem from different research traditions, and each has developed its own terminology and standard ways of framing research questions. However, equivalent models may be specified within each framework so that identical parameter estimates can be obtained (Bauer, 2003; Bauer and Curran, 2002). Raudenbush and Bryk (2002, p.187) note that the choice between LGM and HLM techniques reflects limitations in current software capabilities rather than limitations in modeling possibilities. For a detailed review and tutorial on these models, see Preacher et al. (2008) and Otondo et al. (2009). As concluded by Preacher et al. (2008, p. 78), the distinction between LGM and MLM may disappear altogether with advances in model specification and software.

---

[6] For example, in our study, λ_1=0 and λ_4=1, while the two loadings (λ_2 and λ_3) for the middle two years are estimated as additional parameters from the data.

Illustrative Example

Literature on IT Investments and Firm Performance

As an illustrative application, we use LGM to identify the relationship between IT investment and firm performance. IT investment plays an increasingly important role in today's business and government statistics suggest that over 50% of capital expenditure in US firms is invested in IT (BEA, 2009). The relationship between IT investment and firm performance has been analyzed extensively in the literature (Kohli and Devaraj, 2003; Melville et al., 2004). A variety of research methods has been used in the literature, ranging from correlation-based assessments, cross-sectional regressions, to panel data analyses (Kohli and Devaraj, 2003). One limitation of these studies is their focus on the influence of IT investment on the *level* of firm performance (Brynjolfsson and Hitt, 2003). This focus implicitly assumes a static relationship between IT investment and firm performance, as the hypotheses typically focus on how a certain level of increase in IT investment leads to a certain level of increase in firm performance in a certain period. Nonetheless, Brynjolfsson and Hitt (2003) argue that the relationship between IT investment and firm performance goes beyond a static relationship. IT investment not only affects the *level* of firm performance but also the *growth* of firm performance. They find that the benefit created by IT investments in the short term is about equal to the cost of the IT investments. Still, the impact of IT investments on firm performance "rises by a factor of two to eight as longer time periods are considered" (Brynjolfsson and Hitt, 2000, p. 33). The importance of IT investment to long-term firm performance growth is due to the nature of IT. IT is a type of generic technologies and the impact of such technologies depends on complementary firm investments, such as organizational change processes, human resource practices, and business strategies. As such, IT investments enable the future growth of firm performance.

The idea that technological and IT investments affect the growth of firm performance has long been recognized in the literature (e.g., Penrose and Pitelis 2009). However, it is often difficult to model such a dynamic relationship between IT investments and the growth in firm performance. One approach is a *time difference model* where the time difference of firm performance serves as the dependent variable and the time difference in IT investments serve as the independent variable (Brynjolfsson and Hitt, 2003). Such a model, however, imposes a restrictive relationship between the IT investment and the growth of firm performance. In contrast, we demonstrate below that LGM allows a more flexible modeling of this dynamic relationship between IT investments and firm performance.

Data and Variables

We obtained data from two sources: First, we obtain IT data from Harte-Hanks CI Technology database that provides detailed information on IT assets in over 500,000 business establishments in the United States and Canada. Harte-Hanks creates this database by conducting annual interviews with senior IT executives. The information in the database covers a wide range of IT assets, including PC, servers, storage, and IT labors. Earlier versions of this database were used in prior studies in IS research (e.g., Forman, 2005; Chen and Forman, 2006; Xue et al., 2010).

For this study, we obtained data on IT assets for Fortune 1000 firms from 2001-2004. We used the data to estimate each firm's IT capital stock for each of the 4 years. Second, we obtain firm financial data from the Standard & Poor (S&P) Compustat/CRSP databases. The S&P Compustat database archives detailed financial data extracted from firms' quarterly and annual financial reports and covers all US public firms. We used the S&P database to estimate each firms' ordinary capital stock and other variables related to firm performance. The CRSP database provides detailed daily stock price, volume and dividend information for publicly traded firms. We used the CRSP database to estimate the market value of each firm for the calculation of firm performance.

Dependent Variable

Firm performance is measured with Tobin's q. Tobin's q is a market-based measure, defined as the ratio of a firm's capital market value divided by replacement value of its assets. Tobin's q has been widely used in economic, business, and IS studies which show that market-based measures can reflect more information on firm performance than accounting measures. Also, market-based measures are forward-looking, allowing them to capture long-term impacts on firm performance. We calculated Tobin's q for each firm each year using data from S&P Compustat and CRSP databases. Our calculation followed Bharadwaj et al. (1999) and identified Tobin's q as follows:

Tobin's q = (market value of common stocks + liquidating value of preferred stocks + market value of debt) / book value of total assets.

We calculate market value of common stocks using the closing stock price on the last trading day of a calendar year (from CRSP database) times the total number of outstanding common shares reported in the Compustat database. Liquidating value of preferred stocks is obtained from the Compustat database. Market value of debt is calculated as the sum of long-term debt and current liabilities after subtracting liquid current assets (current assets – inventory).

Independent Variables

IT capital stock, IT infrastructure capital stock and IT labor capital stock

We estimated each firm's IT capital stock using the Harte-Hanks data. Harte-Hanks tracks the number of PCs, servers, and IT employees for each firm. We estimated a firm's IT capital stock in two parts: IT infrastructure capital stock and IT labor capital stock (Hitt and Brynjolfsson, 2000). IT infrastructure capital stock is estimated using IT PC costs and IT server costs. The yearly PC costs and server costs were collected from Gartner Dataquest Global PC Annual Forecast and IDC Worldwide Server Quarterly Tracker. The costs were then adjusted to real costs using the BEA price index for Computers and Peripheral Equipment. We then measured IT hardware capital stock by multiplying the number of PCs and servers with their respective real costs. A firm's IT labor costs were measured by multiplying the number of a firm's IT employees with the IT labor costs obtained from BLS occupational compensation data, deflated by the Index of Total Compensation Cost. IT labor capital stock was estimated as three times the IT labor costs (Hitt and Bryjolfsson, 1995). Since the IT capital stock varies significantly across firms, we normalized the data using the log value of IT capital stocks.

Control Variables

Firm Size

Studies have shown that firm size has a significant impact on firm performance and growth. Large firms have the advantage of economies of scale and scope compared to smaller firms. Still, there is a debate on whether the growth opportunities of large firms are limited (Hall, 1987). We measure firm size with the log of the number of employees.

Firm Type

Studies show that firms in the manufacturing and services differ significantly in growth potentials (Heshmati, 2003). We identified each firm's type based on its NAICS industry code, as reported in the S&P Computstat database. We considered all firms in NAICS two-digit industry codes between 11-33 as manufacturing firms and the remaining as service firms.

Market Share

The market share of a firm in the industry also significantly influences a firm's performance and growth. Firms with a dominant market share have stronger pricing power and generally a better performance. However, being in a dominant market position also indicates limited potential for future growth. We measured a firm's market share by its sales divided by industry total sales at the 4-digit NAICS level.

Advertising and R&D expenditures

Advertising and R&D are important aspects of a firm's business strategy. While advertising and R&D expenditures are typically expensed in the year they incurred, studies suggest that these expenditures create invisible capitals and have a long-term impact on firm performance and growth. To measure these invisible capitals, we used a 5-year rolling-average advertising expenditure and 5-year rolling-average R&D expenditure, respectively. We further standardized these measures by the firm's sales.

Debt to Equity ratio

The ability of firms to invest in new projects and opportunities to obtain long-term growth is determined by their ability to obtain external financing (Demiguc-Kunt and Maksimovic, 1998). The debt to equity ratio is thus an important factor in determining firm performance and growth. In this study, we calculated each firm's debt to equity ratio as the total liability divided by total equity, as reported in Compustat database.

Ordinary capital stock and Capital investment

Firm growth is driven by its existing level of capital stock and the amount of capital investment. A firm with a large existing capital base that continues to spend a large amount in capital investments will have better performance and higher growth in the future. We used two measures for a firm's capital stock and capital investment. We measured a firm's existing capital stock with the log of its net property, plants, and equipment, as reported in the Compustat database. We measured a firm's current capital investment level by its total invested capital divided by total assets.

As our main goal of this study is to provide a parsimonious illustration of GLM, we treated all control variables as time-invariant by taking the average across the study's four years (2001-2004).

LGM Analyses

It is possible to use many commercial packages to estimate LGM, such as Mplus, AMOS, and SAS (see Preacher et al. 2008, p.5; Bollen and Curran 2006, p.12-14; Singer and Willett 2003, p. 280-302). In this study, we implemented the various LGM models using SAS Proc CALIS (Tan et al. 2010) and the sample code is available upon request.

Unconditional LGM Model

We start by illustrating three simplest unconditional, linear LGM models as plotted in Figure 1 for the three variables of interest – Tobin's Q, IT infrastructure, and IT labor. The three models (Table 1) shed light on the individual growth pattern of the three variables. All three slope factors are significant, indicating there were significant changes for each of the three variables over time. We also report the covariance between the level and the slope factors. All three covariances are negative and significant, implying the high the initial level is, the slower the potential for growth becomes. Taken together, the unconditional LGM models for each of the three focal variables suggest that all three variables have an increasing trajectory during the focal 2001-2004 period, albeit the potential for growth attenuates if the initial levels in each variable in the first period (2001) are high.

Table 1. Unconditional Models for the Three Focal Variables

	Tobin's Q	IT Infrastructure	IT Labor
Level	-0.075***	2.789***	4.596**
Slope	0.054***	0.062**	0.341*
Covariance (Level vs Slope)	**-0.031****	**-0.22****	**-0.440*****
AGFI	0.905	0.919	0.921
RSMR	0.048	0.011	0.007
NFI	0.931	0.991	0.992
CFI	0.978	0.988	0.992

Note: the Level and Slope values are the means as shown in equation 2.

Conditional LGM Model

We then ran a conditional model for Tobin's Q (Figure 2) only, with all eight control variables (as explained above). Compared with the unconditional model, the fit is improved for all the three measures, as shown in Table 2.

The main added benefit under the conditional LGM model is its ability to examine which cross-sectional (control) variables influence the two primary factors (Level and Slope). For example, Table 2 shows that the growth of firm performance is not influenced by firm size. The result supports Gibrat's Law, and it is consistent with prior empirical studies on the relationship between firm size and growth (Hall 1987). Our analysis also reveals that market share and capital investments are positively associated with the initial level of firm performance, but they have no significant impact on the slope of firm performance. However, advertising expenditures have a positive effect on both the initial level and the longitudinal growth trajectory of firm performance. This combination of the results suggest that advertising expenditures create invisible capital that lays the foundation for future growth in firm performance but tangible capital (as represented by capital investments) has little influence on the future growth.

These findings help identify which cross-sectional (control) variables can influence the level and shape of a dependent variable, thus guiding researchers as to how to influence not only the initial level (similar to prior research which primarily focus on the effect of mean level) but also the trajectory of a dependent variable over time (thus extending prior research).

Table 2. Conditional Model for Tobin's Q		
	Level	**Slope**
Intercept	-0.341	0.178***
Covariance (Level vs Slope)	-0.027	
Firm Size	0.002	-0.006
Firm Type	0.065	0.014
Market Share	0.443***	0.0122
Capital Investment	0.521***	-0.016
Advertising Expenditure	0.554 ***	0.067***
R&D Expenditure	0.294 **	-0.207**
Debt to Equity Ratio	-0.001	-0.002
Ordinary Capital Stock	-0.012	-0.012
AGFI	0.961	
RSMR	0.044	
NFI	0.977	
CFI	0.974	

Non-Linear Multivariate LGM Models[7]

Considering that firm performance (Tobin's Q), IT infrastructure, and IT labor all can exhibit different growth trajectories during the 2001-2004 period, we propose two tri-variate LGM models to represent the longitudinal relationships among the three principal variables using the general LS coding (Figures 5 and 6).

Two structural models could emerge to model the relationships among these three variables. The first structural model would assume a *mediated* relationship where the effect of IT infrastructure is mediated by IT labor. This suggests that the IT infrastructure is used by IT labor to facilitate firm performance (Pavlou and El Sawy, 2006). The second structural model assumes that the relationship between IT infrastructure and IT labor can be bi-directional. This is based on the logic that IT labor can both be enhanced by the quality of the IT infrastructure and also that it enhances the IT infrastructure. We depict both models in Figure 5 and Figure 6, respectively. For simplicity, we do not detail the control variables in these two structural models. Besides, we simplified the Level-Shape (LS) coding.[8]

Mediated LGM Model

It should be clear that each of the observed variables (e.g., Tobin's Q for each of the years 2001-2004) loads on two latent variables (α and β). These path coefficients are directly shown in Figure 5.

The results show that the initial Level of IT infrastructure positively but marginally (at $p<0.1$ level) affects the Slope of Tobin's Q, consistent with the IS literature that expects a positive relationship between IT infrastructure and firm performance growth over time. However, while the initial Level of IT labor positively affects the Level of performance (Tobin's Q), it negatively affects the Slope of Tobin's Q. This is an interesting finding that suggests that IT labor may not help accelerate firm performance over time, potentially due to the non-scalable nature of IT labor and the law of diminishing marginal returns. Moreover, the initial Level of IT infrastructure is positively associated ($\beta=0.512$, $p<.01$) with the Level of IT labor, but it is negatively associated with the Slope of IT labor ($\beta=-0.082$, $p<0.1$). While the IS literature expects that the level of IT infrastructure to positively influence IT labor over time, this relationship is not particularly strong in a longitudinal sense. This result suggests that, while IT

---

[7] Although the general trend in Tobin's Q between 2001and 2004 was increasing, it was not monotonic since there was a dip between 2001 and 2002. This suggests that the growth trajectory in firm performance is not simply linear.

[8] In the mediated model, the estimated loadings for Tobin's q for the middle two years are $\lambda_2=-0.759$ and $\lambda_3=0.514$, implying a decline of performance in year 2002 (compared to 2001) and then an increase in 2003. This demonstrates the ability of LS to capture a general non-linear trajectory. Also, for IT infrastructure, $\lambda_2=0.410$ and $\lambda_3=0.938$; for IT labor, $\lambda_2=0.322$ and $\lambda_3=0.716$.

infrastructure complements IT labor in the short term, it potentially has a substitutive effect on IT labor in the long run. Moreover, in terms of the relationship among the growth of the two IT variables, the Slope of IT infrastructure positively impacts the Slope of IT labor (β=0.368, p<.01). These results are largely consistent with the IS literature.

Finally, there is a negative relationship between the Level and Slope within each variable (β=-0.029 for Tobin's Q, β=-0.614 for IT infrastructure, and β=-1.110 for IT labor), as also shown for the unconditional model in Table 1. These results suggest that the higher the initial level in each variable, the lower the prospects for growth over time.

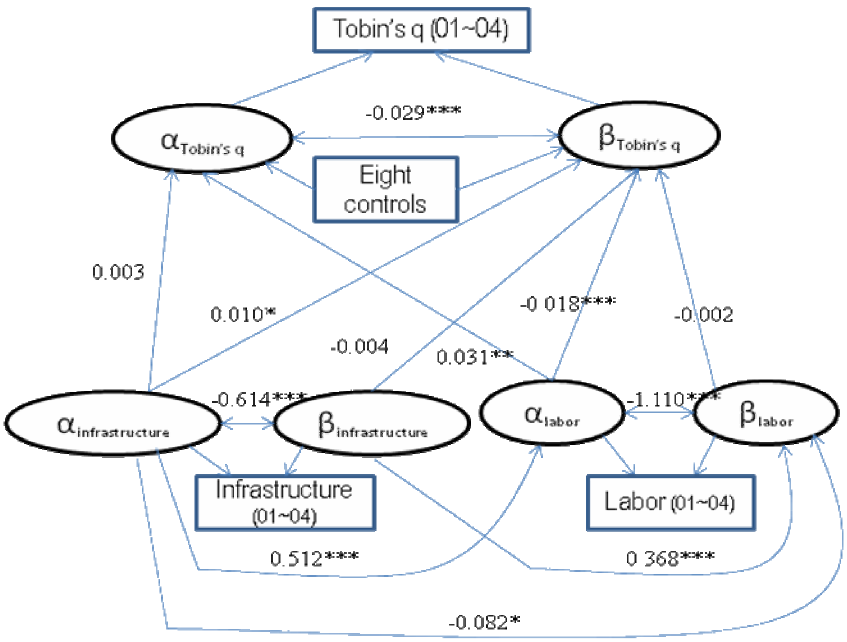

Figure 5: Mediated Trivariate LGM

(GFI=0.94, AGFI=0.90, SRMR=0.02, CFI=0.98, NFI=0.98)

Bi-Directional LGM Model

Figure 6 presents the results on the bi-directional case. The results suggest that initial Level of IT infrastructure positively (p<0.01 level) affects the growth of Tobin's Q, similar to the unidirectional case. The Level of IT labor positively affects the Level of Tobin's Q, but it negatively affects the Slope of Tobin's Q. Again these results are consistent with the results reported in Figure 5.

The initial Level of IT infrastructure is positively associated (β=0.843) with the Level of IT labor, but it is negatively associated (β=-0.362) with the Slope of IT labor. Besides, the Slope of the IT infrastructure is positively associated with the Slope of IT labor (β=0.565). Finally, there is a negative impact between the Level and Slope within each of the three variables (β=-0.029 for Tobin's Q, β=-0.617 for IT infrastructure, and β=-1.296 for IT labor).

In sum, these results are largely consistent with those in Figure 5 for the directional case, implying that there is no major difference in the results regardless of whether the relationship between IT infrastructure and IT labor is modeled as directional (IT infrastructure influencing IT labor) or bidirectional.

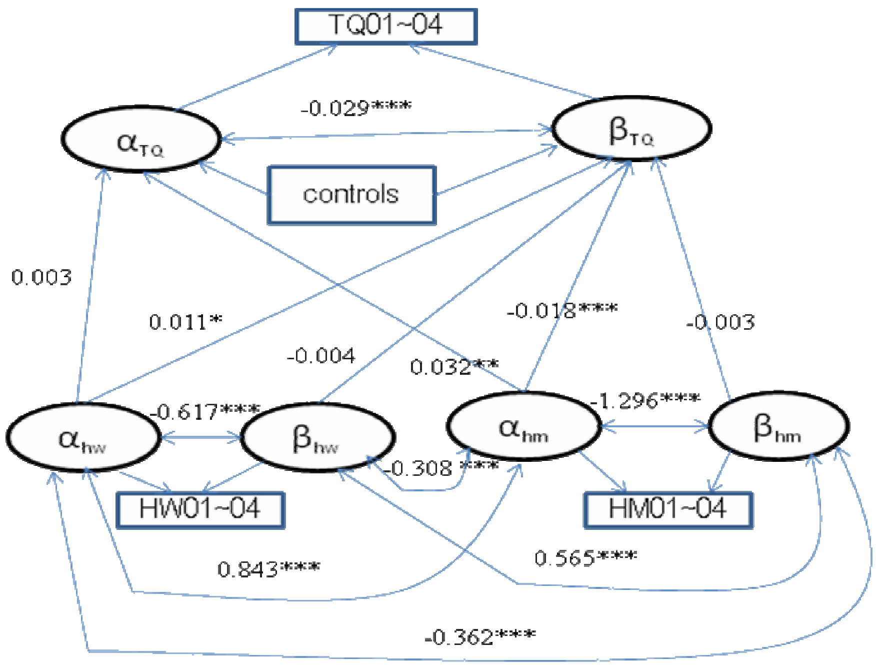

Figure 6. Bi-directional Trivariate LGM (GFI=0.94, AGFI=0.90, SRMR=0.02, CFI=0.98, NFI=0.98)

Comparative Results: LGM analysis versus Conventional Approaches

We compare LGM with two common approaches used in IS research, (a) Structural Equation Modeling (SEM) and (b) panel data econometric analysis using the Generalized Linear Model (GLM).

Cross-Sectional Structural Equation Modeling (SEM) Estimation

As our analysis of the longitudinal studies in MISQ and ISR indicates, during the last 10 years, the majority (65%) of empirical studies use SEM (e.g., Davis and Hufnage, 2007; Kim et al., 2009). To conduct the SEM analysis, we pooled the data across the four years and estimated the following path diagram (Figure 7). The SEM analysis is implemented in SAS Proc CALIS, which adopts LISREL estimation. Similar to the LGM models in Figures 5 and 6, we specified a simple mediated structure model where the effect of IT Infrastructure on firm performance is mediated by IT labor (Figure 7). The SEM results show that IT infrastructure has both a direct and an indirect effect on firm performance, while the direct impact of IT labor on firm performance is negative. For clarity, we omitted the path coefficients of the control variables from Figure 7.

Figure 7 (relative to Figure 5) shows how incorporating the time dimension into SEM adds new insights. Notably, the cross-sectional SEM model in Figure 7 does not permit to make inferences about longitudinal patterns. First, while the SEM results suggest that IT labor has a negative effect on firm performance (β =-0.095, $p<.01$) (Figure 7), the corresponding LGM results clarify that the effect of the Level of IT labor is positive and significant on the initial Level of firm performance but negative and significant on the Slope of IT infrastructure over time. Moreover, while the IT infrastructure has a significant positive effect on IT labor (β =0.39, $p<.001$) in a cross-sectional SEM model, the LGM results suggest that the relationship is positive in terms of the initial levels of IT infrastructure and IT labor but negative between the initial Level of IT infrastructure and the Slope of IT labor over time. Finally, while the SEM results show a positive effect of IT infrastructure on firm performance, the LGM results specify that the only significant relationship is that between the initial Level of IT infrastructure and the Slope of firm performance.

In sum, while the SEM results specify cross-sectional relationships among the proposed variables, the LGM models can identify specific relationships about both the initial levels and slope of these variables over time. This reckons with Mitchell and James (2001)'s concern that the simpler hypotheses built on cross-sectional analyses can turn out to be (possibly mistakenly) strong and are harder to falsify.

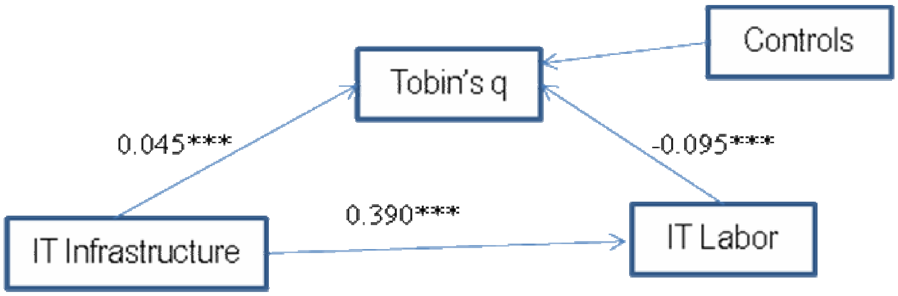

Figure 7. The SEM Estimation using Pooled 2001-2004 Data

(* 0.1, ** 0.05, ***0.01, GFI=0.994, AGFI=0.928, SRMR=0.003, CFI=0.995, NFI=0.997)

Generalized Linear Model (GLM)

During our analysis of the papers published in MISQ and ISR during the last ten years, about 30 longitudinal studies use panel data analysis approaches, such as the fixed time effects in GLM. Thus, we compared GLM with LGM. Table 3 shows the GLM results that treat Year as fixed effects (Dewan and Ren, 2007; Smith and Telang, 2009). The GLM results show that IT infrastructure does not have significant impact on firm performance, but IT labor has a positive and significant role in firm performance. All year fixed effects are significant (versus default Year 2004).

These results suggest that IT labor plays a major role in firm performance. However, our results from LGM models offer an in-depth understanding of the relationship. The results reveal that while the initial level of IT infrastructure has no effect on the initial *level* of firm performance, it has a significant effect on the *growth* of firm performance. On the other hand, while IT labor is positively associated with the initial *level* of firm performance, it has a negative impact on the *growth* of firm performance. The comparison shows that similar to SEM, by focusing on static relationships, the GLM results in Table 3 do not offer any specific information on the longitudinal effects of IT infrastructure and IT labor on firm performance. The inclusion of fixed time effects merely controls for variations of firm performance over time, but it does not reveal the longitudinal relationship between variables offered by LGM.

Table 3: GLM results with year as fixed effects

Parameter	Estimate	STD	t-value	p-value
Intercept	-0.130	0.070	-1.86	0.063
IT Infrastructure	-0.001	0.007	-0.22	0.829
IT Labor	0.021	0.006	3.34	0.001
Year 2001	-0.063	0.024	-2.57	0.010
Year 2002	-0.202	0.024	-8.41	<.0001
Year 2003	-0.056	0.024	-2.36	0.018
Year 2004	0.000	.	.	.
Firm Size	0.085	0.018	4.78	<.0001
Firm Type	-0.001	0.008	-0.1	0.924
Market Share	0.430	0.056	7.69	<.0001
Advertising Expenditure	5.592	0.370	15.11	<.0001
R&D Expenditure	2.610	0.151	17.28	<.0001
Debt to Equity Ratio	0.000	0.001	-0.26	0.794
Capital Investment	0.500	0.053	9.37	<.0001
Ordinary Capital Stock	-0.032	0.007	-4.67	<.0001
N	4820			
R-square	0.139			
F	60.440			
P	<0.001			

Discussion

Key Contributions

Given the relative dearth of research on theorizing longitudinal models and proposing longitudinal hypotheses in the IS literature, the study's first contribution is to introduce the basic tenets of LGM and demonstrate with an illustrative example how to use LGM to empirically explore longitudinal relationships among IT variables, such as IT infrastructure and IT labor, and their effects on firm performance over time. Another important contribution to the IS literature is to provide practical guidelines on how to implement various LGM approaches to model various types of longitudinal relationships among IT variables and their effects on firm performance.

Comparative Advantages and Disadvantages of LGM

The first advantage of LGM models is to explicitly model change over time. While the current practice of SEM is to focus on cross-sectional analysis, LGM models integrate a longitudinal time dimension to SEM models. Besides, while the GLM panel data analysis acknowledges that there are time effects across the four years in our sample, unlike LGM, it does not explicitly model changes in relationships over time to specify longitudinal effects.

Second, LGM can also model longitudinal variation within individual variables. Unlike SEM or GLM, LGM breaks down the growth trajectory of a single variable into its initial Level and Slope, thus adding to our understanding of how an individual variable changes over time in terms of its initial value in a period and its growth trajectory.

Third, multi-variate LGM models add to our understanding on how change of one (longitudinal) variable affects change in other variables, an important aspect not captured by SEM or GLM methods and notably missing from IS research. The ability to predict how the longitudinal change in one variable affects the longitudinal change in another opens many new avenues for IS research by exploring complex relationships among IS variables over time.

Finally, while SEM and GLM models focus on linear and directional relationships, LGM can readily model both non-linear and also bi-directional relationships, thus capturing more complex relationships among variables over time that are not necessarily linear or directional.

Optimal Number of Time Periods for Latent Growth Models

While our study used four time periods for the simplicity of illustration and for showing that longitudinal relationships can be adequately modeled with a relatively small number of time periods,[9] there is no general consensus in the LGM literature on the optimal number of time periods. This is largely determined by the model identification, i.e., whether unique values of the model parameters can be identified given the model's structure (e.g. the structural specification and coding of time λ_t in LGM etc.) and data (Bollen and Curran 2006, p.21). In the simplest intercept-slope coding (Figure 1), LGM requires a minimum of *three* time periods. However, for non-linear models, such as Bollen and Curran (2006), at least *four* time periods are needed (p. 91). The proposed Level-Shape (LS) coding used in this study can accommodate both linear and also non-linear models with a minimum requirement of *three* time periods.

Suggestions for Future Work

Despite the demonstration of the basic tenets of LGM and offering guidelines to use LGM models in IS research, there is still a need to better explore and understand the value of LGM for IS research. In particular, LGM enables both theoretical and methodological advancements in IS research. As we mentioned earlier, LGM allow researchers to propose longitudinal hypotheses. However, there has been little theoretical guidance on how to develop such hypotheses. Hence, there is a need to expand prior IS theories to explicitly identify longitudinal relationships and develop longitudinal hypothesis. Further, future research could offer practical guidance on developing longitudinal hypotheses, including theorizing longitudinal changes within single variables, how the change in one variable affects the change in another variable, and how a set of variables can influence each other in non-linear and bi-directional ways. Future research could also explore how existing longitudinal studies (either regression, SEM, or GLM

[9] Since our LGM application is only for illustration and not for theory development and testing purposes, it is less critical to have more than four time periods. Nonetheless, future research could include additional time periods.

models) could be informed by LGM, and how existing models could be extended by the ability to model longitudinal effects. From a methodological perspective, the use of LGM in the domain of IS research requires methodology advancement to treat data issues such as unbalanced data, missing data, and measure errors. While techniques for treating these data issues (such as imputation and measurement error correction) are well-known in SEM and GLM models, little is known whether the same technique can be applied to LGM. Prior IS research also shows that relationships among IS variables are often complicated and non-linear. Future research is needed to extend LGM to model such complex relationships.

The study's ultimate purpose is to bring longitudinal aspects back to IS research and use appropriate methods that readily extend existing approaches, such as SEM and GLM, to model relationships among IS variables over time. We hope to entice future IS research to more carefully examine the longitudinal aspects of IS phenomena.

References

Adams, J. D., and A. B. Jaffe, "Bounding the Effects of R&D," *RAND Journal of Economics* (27:4), 1996, pp. 700-721.

Bauer, D.J. (2003). Estimating multilevel linear models as structural equation models. Journal of Educational and Behavioral Statistics, 28, 134-167.

Bharadwaj, A. S., Bharadwaj, S. G., and Konsynski, B. R. "Information Technology Effects on Firm Performance as Measured by Tobin's q," *Management Science* (45:6), 1999, pp. 1008-1024

Bollen, K., Curran, P. *Latent curve models: a structural equation perspective*. Wiley & Sons, Hoboken, NJ, 2006.

Boh, W., Slaughter, S., spinosa, S. 2007. Learning from Experience in Software Development: A Multilevel Analysis. Management Science, Vol. 53, No. 8, August 2007, pp. 1315-1331

Bresnahan T. F., E. Brynjolfsson, L. M. Hitt. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. Quarterly Journal of Economics (117:1), 2002, pp. 339-375.

Brynjolfsson, E. "The productivity paradox of information technology," Communications of the ACM (35:12), 1993, pp. 66-67.

Brynjolfsson, E., and Hitt, L. "Paradox Lost? Firm-level Evidence on the Returns to Information Systems Spending," *Management Science* (42:4), 1996, pp. 541-558

Chen, P., C. Forman. 2006. "Can vendors influence switching costs and compatibility in an environment with open standards?" *MIS Quarterly* 30 541-562

Curran, P., Willoughby, T. 2003. Implications of latent trajectory models for the study of developmental psychopathology. Dev Psychopathol. 2003 Summer;15(3):581-612.

Curran, P., Bauer, D., Willoughby, T. Testing Main Effects and Interactions in Latent Curve Analysis. Psychological Methods . Vol 9(2), Jun 2004, 220-237

Daft, R.L. & Lengel, R.H. (1986). Organizational information requirements, media richness and structural design. Management Science 32(5), 554-571.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

Davis, C. and E. Hufnagel. 2007. Through the Eyes of Experts: a Socio-Cognitive Perspective on the Automation of Fingerprint Work. *MIS Quarterly* 31(3): 1-23.

Devaraj, S., and Kohli, R. "Performance Impacts of Information Technology: Is Actual Usage the Missing Link?" *Management Science* (49:3), 2003, pp. 273-289

Dewan, S. and Ren, F., 2007. "Risk and Return of Information Technology Initiatives: Evidence from Electronic Commerce Announcements," *Information Systems Research,* December 2007, Vol. 18, No. 4, pp. 370-394

Duncan, T.E. & Duncan, S.C. (1995). Modeling the Processes of Development via Latent Variable Growth Curve Methodology. *Structural Equation Modeling*, 2(3), 187-213.

Duncan, T. E., Duncan, S. C., Strycker. 2006.. An introduction to latent variable growth curve mdoeling. Mahwah, NJ: Lawrence Erlbaum Associates.

Fichman, R., Kemerer, C. "The illusory diffusion of innovation: an examination of assimilation gaps," *Information Systems Research* (10:3), 1999, pp. 255-275.

Findlay, R., 1978, "Relative Backwardness, Direct Foreign Investment, and the Transfer of Technology: A Simple Dynamic Model," *The Quarterly Journal of Economics* 92(1), 1-16.

Forman, C. 2005. "The corporate digital divide: Determinants of internet adoption," *Management Science* 51(4), 641-654

George, J. M., and Jones, G. R. “The role of time in theory and theory building,” *Journal of Management* (26), 2000, pp. 657-684.

Goldstein, H. *Multilevel statistical models*. New York, Wiley, 1995.

Gottfried, A.E., Marcoulides, G.A., Gottfried, A.W., Oliver, P.H., Guerin, D.W. “Modeling the developmental decline in academic intrinsic math motivation: Childhood through adolescence,” *International Journal of Behavioral Development* (Special Issue on Longitudinal Modeling of Developmental Processes (31:4), 2007, pp. 317-327.

Gottfried, A.E., G.A. Marcoulides, A.W. Gottfried, and P.H. Oliver. 2009. “A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation.” Journal of Educational Psychology, 101(3): 729-739.

Greene, W. 2007. *Econometrics Analysis*. Prentice Hall.

Hall, B. H. “The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector,” *Journal of Industrial Economics* (35:4), 1987, pp. 583-606.

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.

Jackman , S. *Bayesian Analysis for the Social Sciences*. Wiley Publications, New York, NY, 2009.

Kher, H., M. A. Serva, S. Davidson, E. Monk. Leveraging latent growth models to better understand MIS theory: A primer. *Proceedings of the Special Interest Group on Management Information Systems' 47th Annual Conference on Computer Personnel Research*, Limerick, Ireland, 2009.

Kim, H.-W. and Kankanhalli, A. 2009. Investigating user resistance to information systems implementation: a

status quo bias perspective. MIS Quarterly, 33(3): 567-582

Kohli, R., S. Devaraj. “Measuring information technology payoff: A meta-analysis of structural variables in firm-level empirical research,” *Information Systems Research* (14:2), 2003, pp. 127-145.

Levinthal, D. and J. March . 1993. The Myopia of Learning. *Strategic Management Journal*, 14: 95-112.

Lewicki, R. J., Tomlinson, E. C., and Gillespie, N. “Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions,” *Journal of Management* (32), 2006, pp. 991-1022.

Massey, A., Montoya-Weiss, M. 2006. Unraveling the Temporal Fabric of Knowledge Conversion: A Model of Media Selection and Use. MIS Quarterly, 30 (1).

March*, J.G. (*1991*).* Exploration and Exploitation *in* Organizational Learning*' , Organizational Science* 1: 71-87.

Marcoulides, G. A., Hershberger, S. L. *Multivariate statistical methods*: A first course. Mahwah, NJ: Lawrence Erlbaum Associates, 1997.

Marcoulides, G., Hershberger, S. *Multivariate statistical methods: A first course*. Lawrence Erlbaum Associates Publishers, 1997.

Melville, N., K. Kraemer, V. Gurbaxani. “Information technology and organizational performance: An integrative model of IT business value,” *MIS Quarterly* (28:2), 2004, pp. 283-322

Melville, N., Kraemer, K., and Gurbaxani, V. “Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value,” *MIS Quarterly* (28:2), 2004, pp. 283-322

Mitchell, T. and James, L. “Building better theory: time and the specification of when things happen,” *Academy of Management Review* (26:4), 2001, pp. 530-547.

Mithas, S., Ramasubbu, N., Krishnan, M.S. and Fornell, C. "Designing Websites for Customer Loyalty across Business Domains: A Multilevel Analysis," Journal of Management Information Systems (23:3) 2006, pp. 97-127

Muthén, B. (1993). Goodness of fit with categorical and other non-normal variables. In K. A. Bollen, & J. S. Long (Eds.), Testing Structural Equation Models (pp. 205-243). Newbury Park, CA: Sage.

Otondo, R. F., Barnett, T., Kellermanns, F. W., Pearson, A. W., Pearson, R. A. Assessing Information Technology Use over Time with Growth Modeling and Hierarchical Linear Modeling: A Tutorial. *Communications of the Association for Information Systems*, 2009, 25, 1-37.

Pavlou, P., Gefen, D. 2004. Building Effective Online Marketplaces with Institution-Based Trust. *Information Systems Research.* 15 (1) , pp. 37-59.

Penrose, E., Pitelis, C. *The theory of the growth of the firm, 4th edition*. Oxford University Press, New York, 2009.

Pitariu, A., Ployhart, R. "Explaining change: theorizing and testing dynamic mediated longitudinal relationships," *Journal of Management* (36:2), 2010, pp. 405-429.

Preacher, K., Wichman, A., Maccallum, R., Briggs, N. *Latent Growth Curve Modeling*. Sage Publications, 2008.

Qureshi, I., Y. Wang, D. Compeau, D. Meister. Capturing the dynamics of adoption through Latent Curve Modeling, *DIGIT 2008 Proceedings.* 2008.

Raudenbush, W. and Bryk, S. *Hierarchical linear models: applications and data analysis methods*, Thousand Oaks, CA: Sage Publications, 2002.

Raykov, T., Marcoulides, G.A. *An introduction to applied multivariate analysis*, Lawrence Erlbaum Associates, Inc. Publishers, 2008.

Smith, Michael, Rahul Telang. 2009. Competing with Free: The Impact of Movie Broadcasts on DVD Sales and Internet Piracy. *Management Information Systems Quarterly*, **33**(2) 312-338.

Straub, D. "The effect of culture on IT diffusion: e-mail and fax in Japan and the U.S." *Information Systems Research* (5:1), 1994, pp. 23-47.

Tan, T.K., T. Kang, and D. Hogan. 2010. "Building latent growth models using PROC CALIS: A structural equation modeling approach." SAS Global Forum 2010, Paper 273-2010.

Tushman, Michael, and D. Nadler. "Information Processing as an Integrating Concept in Organization Design." *The Academy of Management Review* 3 (1978): 613-624.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, (46:2), 186-204.

Whiteman, S. and Mroczek, D. "A brief introduction to growth curve models," *The Irish Journal of Psychology* (28:1), 2007, pp. 77-85.

Xue, L, G. Ray, B. Gu. "Environmental uncertainty and IT infrastructure governance: A moral hazard perspective," *Information Systems Research*, 2010, forthcoming

Zhu, K. and Weyant, J. "Strategic decisions of new technology adoption under asymmetric information," *Decision Sciences* (34:4), 2003, pp. 643-675.