Running head: Piecewise Trajectories

Power of Latent Growth Curve Models to Detect Piecewise Linear Trajectories

Thierno M. O. Diallo*, Ph.D.

University of Sherbrooke, Sherbrooke, Canada &

Statistiques & M.N., Sherbrooke, Canada

Alexandre J.S. Morin *, Ph.D.

Centre for Positive Psychology and Education, University of Western Sydney, Sydney, Australia

* The authors (TMOD & AJSM) contributed equally to this article and their order was determined at

random: both should be considered first authors.

Corresponding author:
Alexandre J.S. Morin
Centre for Positive Psychology and Education,
University of Western Sydney, Bankstown Campus, Building 19
Locked Bag 1797, Penrith, NSW 2751, Australia
E-mail: A.Morin@uws.edu.au
Phone : +61 2 9772 6130
Fax : +61 2 9772 6193

Power of Latent Growth Curve Models to Detect Piecewise Linear Trajectories

Abstract

Latent curve models (LCMs) have been used extensively to analyse longitudinal data. However, little is known about the power of LCMs to detect nonlinear trends when they are present in the data. This simulation study was designed to investigate the Type I error rates, rates of nonconvergence and the power of LCMs to detect piecewise linear growth and mean differences in the slopes of the two joined longitudinal processes represented by the piecewise model. The impact of seven design factors was examined: number of time points, growth magnitude (slope mean), inter-individual variability, sample size, position of the turning point, and the correlation of the intercept and the second slope as well between the two slopes. The results show that previous results based on linear LCM can not be fully generalized to a nonlinear model defined by two linear slopes. Interestingly, design factors specific to the piecewise context (position of the turning point and correlation between the two growth factors) had some effects on the results but these effects remained minimal and much lower than the effects of other design factors. Similarly, observed rates of inadmissible solutions are comparable to those previously reported for linear LCMs. The major finding of this study is that a moderate sample size is ($n = 200$) need to detect piecewise linear trajectories, but that much larger samples ($n = 1500$) are required to achieve adequate statistical power to detect slope mean difference of small magnitude.

Keywords: power analysis, latent curve models, piecewise trajectory, mean difference, structural equation models, Monte Carlo, convergence, Type I error rates.

Latent Curve Models (LCM; Meredith & Tisak, 1990) are naturally suited to the analysis of longitudinal processes and increasingly used in applied research settings. LCM can be equivalently operationalized (e.g., Curran, 2003) within the multilevel (e.g., Bryk & Raudenbush, 1992) or the Structural Equation Modeling (SEM; e.g., Bollen & Curran, 2006) framework, although our focus is on the SEM approach. LCM can be extended to incorporate a variety of polynomial trends (i.e., quadratic, cubic, etc.) or even more complex nonlinear functions (e.g., exponential, logistic, etc.) of growth (e.g., Blozis, 2007; Browne & DuToit, 1991; Grimm, Ram, & Hamagami, 2011; Ram & Grimm, 2007). However, although polynomial functions allows developmental trends to evolve over time, these models usually estimate smoothed trajectories and assume that the direction of the developmental process will eventually change over time. For instance, an initially sharp increase will become less pronounced over time, before flattening out, and eventually starting to decrease in a quadratic inverted U-Shape function. These models thus fail to consider the possibility of sharp changes that are routinely expected with experimental/clinical studies, organizational changes, or important lifetime transitions (e.g., Eccles et al., 1993; Jones & Meredith, 2000; Shadish & Cook, 2009), as well as the fact that these developmental trends may not necessarily change direction over time (i.e., sharp growth may stop, and then becomes less pronounced, without showing a decrease).

Piecewise linear (PWL) LCM (Bollen & Curran, 2006; Flora, 2008) are a flexible alternative when specific transition points can be expected (life transitions, experiments, interventions, etc.). PWL-LCM are naturally suited to experimental or clinical studies were turning points can be specified as the beginning, or end, of the treatment. In PWL-LCM, nonlinearity is modeled by including two interrelated linear slopes reflecting the growth trajectory before and after the transition point. PWL models allow for the fact that both slopes might be differently related to predictors, outcomes, or time-varying covariates, while presenting either the same, or a different, average shape.

Despite the extensive use of LCMs, very few studies provide systematic estimates of the statistical power of the LCM to detect specific types of development (linear, quadratic, piecewise, etc.). Most of statistical studies of the power of LCM were concerned with the capacity of these models to detect between group differences (Duncan, Duncan, Strycker, & Li, 2002; Fan, 2003; Muthén & Curran, 1997), a covariance component among two linear rates of change (Hertzog,

Lindenberger, Ghisletta, & von Oertzen, 2006), individual differences in change (Hertzog, von Oertzen, Ghisletta, & Lindenberger, 2008), covariate intercepts interactions (Sun & Willson, 2009), or longitudinal mediation (Cheong, 2011), rather than the ability of these models to correctly estimate the shape of the trajectories, which represents an important pre-requisite to the use of these models. To our knowledge, only two studies did so. The first of those studies was not designed as a simulation study. Zhang and Wang (2009) developed a SAS macro to help applied researchers to conduct a priori power analyses for linear and nonlinear LCM functions (excluding PWL models) as a function of a limited set of conditions including sample size, growth magnitude and number of measurement points. To illustrate this macro, they report a short power analysis of linear and exponential LCM. Their results showed that for linear trajectories, power increased with sample size (50 to 1000), effect sizes (three different mean of the linear slope factor: .1, .2, .3), and number of measurement points (3 to 6). Regarding the exponential trajectory, fewer conditions were investigated and power was found to increase with sample size (100 to 1000). In the second study, Fan and Fan (2005) compared the capacity of linear LCM to detect linear growth as a function of the number of time points (4 conditions: 3 to 9), growth magnitude (6 conditions: .20 to .80) and sample size (10 conditions: 50 to 500). Their results confirm the role of most of these conditions on the power of LCM to detect linear growth, but surprisingly showed that the number of repeated measures had no effect. This result was unexpected given previous observations that this factor played a significant role in influencing LCM power to detect linear and exponential growth (Zhang & Wang, 2009) or the covariance between linear slopes (Hertzog et al., 2006). Their results also showed that, with 3 time points, LCM was associated with increased rates of nonconvergence. Clearly, these results deserve replication.

Most previous studies focused on linear trajectories, which is a very strict assumption to hold when modeling real life longitudinal data where nonlinear trajectories have frequently been observed (e.g., Grimm et al., 2011; Morin, Maïano, Nagengast, Marsh, Morizot, & Janosz, 2011; Ram & Grimm, 2007). In addition, none of these studies specifically considered PWL models. Thus, although we can often reasonably expect longitudinal processes to follow nonlinear trajectories (e.g., Grimm et al., 2011; Ram & Grimm, 2007), we currently have little information regarding the power of LCMs to detect nonlinear trends, especially PWL processes, when they are present in the data. Given that PWL

models are naturally suited to experimental or treatment studies where a priori determination of power is often of major importance, this lack of knowledge seems quite dramatic. The present simulation study examined the power of LCM in detecting nonlinear growth defined by two linear trends. In addition, we also address the issue of nonconvergence (i.e., converging on inadmissible solutions).

Piecewise Linear Trajectory Model

Let us assume a series of six equally spaced repeated measurements Y_{it}, for individual i at time t. In scalar terms, a PWL-LCM is represented as:

$$Y_{it} = I_i + \lambda_{1t}S_{1i} + \lambda_{2t}S_{2i} + \varepsilon_{it} \tag{1.1}$$

where I_i is the intercept of individual i growth trajectory, λ_{1t} is the time score for the first slope S_{1i}, λ_{2t} is the time score for the second slope S_{2i}, and ε_{it} is the time specific residual for individual i. Growth is represented by imposing constraints on the time scores (λ_{1t} or λ_{2t}) usually specified to reflect the passage of time. S_{1i} and S_{2i} reflect the level of change in Y that is observed for individual i between two consecutive time points (i.e., when λ_{1t} or λ_{2t} changes by one unit). S_{1i} reflects linear growth occurring before the turning point (or transition point, or knot) whereas S_{2i} reflects linear growth occurring after the turning point. Because each of these latent factors is a random variable, I_i, S_{1i}, and S_{2i} can be described by the average level observed in the sample (μ_I, μ_{S_1}, μ_{S_2}), plus a variance component reflecting inter-individual deviations from the growth factor means (ζ_{I_i}, $\zeta_{S_{1i}}$, $\zeta_{S_{2i}}$):

$$\begin{aligned} I_i &= \mu_I + \zeta_{I_i} \\ S_{1i} &= \mu_{S_1} + \zeta_{S_{1i}} \\ S_{2i} &= \mu_{S_2} + \zeta_{S_{2i}} \end{aligned} \tag{1.2}$$

This representation enables us to have two distinct linear "pieces", one before the turning point and one after the turning point, which should usually be determined based on theoretical or experimental considerations. Strictly speaking, there is no need for μ_{S_2} to be significantly different from μ_{S_1} or for ζ_{S_2} to be significantly different from ζ_{S_1} in order to properly specify a PWL-LCM– only a sufficient reason to expect individual trajectories to change before and after to turning point. Indeed, it may be far more interesting to verify whether the determinants or outcomes of S_{1i} differ, or not, from those of S_{2i}, or whether S_{1i} predicts S_{2i} in the context of complex mediation models (e.g., Cheong, MacKinnon,

& Khoo, 2003). However, changes located at the level of the average growth process or levels of intra-individual variability may also present a substantive interest in the context of experimental studies, and sometimes one may even be interested in locating an empirically determined turning point, for instance, in order to study latency in experimental effects. In this context, this turning point can be located with modification indexes showing where the fixed loadings on the slope factor need to be relaxed due to the presence of the turning point (e.g. Kwok, Luo, & West, 2010) or with a latent basis model with empirically estimated loadings imposing no specific shape on the trajectory (e.g. Grimm et al., 2011; Ram & Grimm, 2007). Further extensions are available in the context of Mixed Linear Models (Cudeck & Harring, 2007; Cudeck & Klebe, 2002), allowing for the estimation of PWL models where the turning point is empirically identified and allowed to differ between subjects.

More precisely, when estimated within the SEM framework, LCM are specified as restricted factor models, where the growth factors (I, S_1 and S_2) influence the repeated measures through fixed loadings reflecting the passage of time. In matrix form, for a PWL-LCM model with a turning point occurring in the third occasion, equation (1.1) is expressed as:

$$Y_i = \Lambda \eta_i + \varepsilon_i, \text{ with } \varepsilon_i \sim N(0, \Theta) \tag{1.3}$$

where

$$Y_i = \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \\ Y_{i5} \\ Y_{i6} \end{bmatrix}, \Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \eta_i = \begin{bmatrix} I_i \\ S_{1i} \\ S_{2i} \end{bmatrix}, \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \\ \varepsilon_{i4} \\ \varepsilon_{i5} \\ \varepsilon_{i6} \end{bmatrix}, \text{ and } \Theta = \begin{bmatrix} \theta_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_6 \end{bmatrix}$$

and where, Y_i is a 6x1 vector of repeated measures for individual i over the six repeated measurements, Λ is a 6x3 matrix of factor loadings, η_i is a 3x1 vector of latent variables, and ε_i is a 6x1 vector of time specific residuals following a normal distribution with mean 0 and a 6x6 covariance matrice (Θ) of ε_i with residual variances θ_1 to θ_6 (for consistency with previous studies, in this study this matrix is assume to be diagonal and homoscedastic). The factor loadings associated with I are in the first column {1, 1, 1, 1, 1, 1}. The loadings associated with S_1 are in the second column {0, 1, 2, 2, 2, 2} and reflect linear growth between the first three time points (after which the

equal loadings allow the remaining growth information to be absorbed by S_2), and the factor loadings associated with S_2 are in the third column {0, 0, 0, 1, 2, 3}, reflecting linear growth between the last three time points (before which the equal 0 loadings allow the preceding growth information to be absorbed by S_1). An alternative representation of the PWL-LCM model can also be estimated where the loadings on the S_1 are rather specified as in a linear LCM {0, 1, 2, 3, 4, 5}. In this added-rate model S_2 represents the difference in linear change between the second and first period and is particularly useful in some of the contexts in which one wishes to evaluate how people have been deflected from the trajectories they were on before as a function of intervention/treatment.

The mean structure of equation (1.3) can be expressed as

$$\eta_i = \mu_\eta + \zeta_i, \text{ with } \zeta_i \sim N(0, \Phi) \tag{1.4}$$

where μ_η is a 3x1 vector of means and ζ_i is a 3x1 vector of inter-individual variability around μ_η, that is assumed to follow a normal distribution with a mean of 0 and a 3x3 covariance matrice of individual deviations from the growth factor means Φ such that

$$\mu_\eta = \begin{bmatrix} \mu_I \\ \mu_{S_1} \\ \mu_{S_2} \end{bmatrix}, \zeta_i = \begin{bmatrix} \zeta_{I_i} \\ \zeta_{S_{1i}} \\ \zeta_{S_{2i}} \end{bmatrix}, \text{ and } \Phi = \begin{bmatrix} \phi_I & \phi_{IS_1} & \phi_{IS_2} \\ \phi_{IS_1} & \phi_{S_1} & \phi_{S_1 S_2} \\ \phi_{IS_2} & \phi_{S_1 S_2} & \phi_{S_2} \end{bmatrix}$$

where ϕ_I is the inter-individual variance of I, ϕ_{S_1} is the inter-individual variance of S_1, ϕ_{S_2} is the inter-individual variance of S_2, ϕ_{IS_1} is the covariance between I and S_1, ϕ_{IS_2} is the covariance between the I and S_2, and $\phi_{S_1 S_2}$ is the covariance between S_1 and S_2. Figure 1 graphically presents this PWL-LCM.

Purposes of the study

In this study, we investigate empirical power rates associated with the detection of PWL growth over time (defined as μ_{S_2}), as well as the difference between the two slopes (defined as $\mu_{S_1} - \mu_{S_2}$). More specifically, we investigate the effects of the number of time points, growth magnitude and inter-individual variability, sample size, the position of the turning points and the correlation between the second slope factor and other growth factors (the intercept and the first slope) on the statistical power to detect μ_{S_2} and $\mu_{S_1} - \mu_{S_2}$, Type I error rates and rates of nonconvergence.

Method

Statistical Model

The population models used in this study are PWL-LCM as previously defined. The data were generated under multivariate normality conditions. All observed variables were specified as continuous, and the piecewise linear growth was modeled with equally spaced time intervals. In the data generation process, we fixed the mean and variance of I and S_1 and the covariance between I and S_1, and varied the mean and variance of S_2. More precisely, μ_I was fixed to 1, μ_{S_1} was fixed to .16 (reflecting a medium growth for the first slope; e.g., Kwok, West, & Green, 2007; Kwok et al., 2010), ϕ_I was fixed to .2, ϕ_{S_1} was fixed to .1 (reflecting a medium level of inter-individual variability for the first slope; e.g., Kwok et al., 2007, 2010), and ϕ_{IS_1} was fixed to .05. Across conditions, the residual variances of the measures were specified as homoscedastic and fixed to .2.[1]

Data are generated under different conditions defined by the combination of growth magnitude of S_2, sample size, number of measurement occasions and position of the turning point. These conditions represent potentially critical factors that can affect the statistical power. Based on previous studies (Cheong, 2011; Fan & Fan, 2005; Hertzog et al., 2006; Sun & Willson, 2009), it is expected that the power to detect PWL growth will be enhanced as the sample size and the size of the effect (i.e., μ_{S_2} or $\mu_{S_1} - \mu_{S_2}$) increase. We also considered the effect of the variability of S_2 (ϕ_{S_2}) since the amount of inter-individual variability in growth has been previously found to play a role in influencing the size of biases induced by model misspecifications in LCM (Kwok et al., 2007; Voelkle, 2008). In this regard, we expect power to decrease as a function of the size of ϕ_{S_2} and to increase as a function of t, the number of repeated measurements (Cheong, 2011; Hertzog et al., 2006; but see Fan & Fan, 2005). Given the equivalence of SEM-based and multilevel-based LCM (e.g., Curran, 2003), repeated measurements constitute observations taken at the level 1 (within-person). It has also been previously argued that increased number of measurement points allows for a greater precision in the estimation of LCM (Cheong, 2011; Singer & Willett, 2003), and to result in lower rates of convergence problems (Fan & Fan, 2005). We also investigate the impact of the position of

the turning point. Indeed, the position of the turning point influences the number of time points available to estimate S_2 and thus combine with the total number of measurement occasions and sample size in influencing empirical power rates, Type I error rates, and rates of nonconvergence. Finally, another potentially important facet of PWL is the correlation between the growth factors, especially between the second slope and the intercept and first slope as these reflect the extent to which the trajectories estimated before and after the turning points exchange information. Given the total number of manipulated factors, we rely on a combination of a full factorial design for the main factors usually manipulated in power studies (mean and variance of S_2, sample size, and t) with a partial factorial design for the remaining factors (position of the turning points and correlations between the growth factors) which are each fully crossed with the main manipulated factors, but not with one another (e.g., Beauchaine & Beauchaine, 2002; Tofighi & Enders, 2007).

Full Factorial Design

   Mean and variance of S_2. In this simulation study we focus on the detection of μ_{S_2} and of $\mu_{S_1} - \mu_{S_2}$. These parameters (slopes, and slope difference) refer to clear and testable effects of direct relevance to developmental and experimental studies. Following Kwok et al. (2010), we simulated data with 4 different values of μ_{S_2} reflecting respectively a small, moderate, large and null change from μ_{S_2}: (a) .11 (i.e., $\mu_{S_1} - \mu_{S_2} = .16-.11 = .05$, a small difference); (b) 0 (i.e., $\mu_{S_1} - \mu_{S_2} = .16 - 0 = .16$, a moderate difference); (c) .55 (i.e., $\mu_{S_1} - \mu_{S_2} = .16-.55 = -.39$, a large difference); (d) .16 (i.e., $\mu_{S_1} - \mu_{S_2} = .16-.16 = 0$, a null difference). Figure 2 presents the first three average trajectories. We voluntarily limited this simulation to situations where PWL development may be less obvious, simply marked by a change of magnitude rather than a change in direction. However, we expect the results to fully generalize to the similar scenarios marked by changes in directions (e.g., where $\mu_{S_2} = -.11$ or -.55, or $\mu_{S_1} - \mu_{S_2} = -.05, -.16, -.39$). We also considered 3 levels of ϕ_{S_2}, to reflect low ($\phi_{S_2} = .05$, or SD = .22), moderate ($\phi_{S_2} = .16$, or SD = .4) or high ($\phi_{S_2} = .36$, or SD = .6) levels of variability.

   Sample size. We simulated data based on 11 different sample sizes in line with values

considered in previous studies (Ferron, Dailey, & Yi, 2002; Kwok et al., 2010): 30, 50, 100, 150, 200, 300, 500, 1000, 1500, 2000, and 3000. The first two values were chosen in order to evaluate the power and Type I error rates in samples smaller that what is usually seen in applied LCM research, especially in combination with increasing number of measurement points so as to reflect a reality that is more common in the context of time series analyses with few participants, but multiple waves (e.g. Hamaker, Dolan, & Molenaar, 2005). Even though the empirical power curves flattened after 300 cases for μ_{S_2}, the last four values were more relevant to the detection of $\mu_{S_1} - \mu_{S_2}$.

Number of measurement occasions. We considered 3 different conditions regarding the number of measurement points: 6, 8 and 10. We chose 6 as the minimum number of measurement points in order to be able to estimate all turning points conditions within each time point conditions while keeping all models fully identified (Bollen & Curran, 2006). Finally, the last two conditions were selected to reflect a moderate and high number of measurement occasions.

Partial Factorial Design

Turning point. In this study we generated data with the transition point occurring at different time points: 2, 3 and 4. Each of these three conditions was investigated in combination with all of the levels on the previous factors, for a total of 1188 (4 values of μ_{S_2} X 3 levels of ϕ_{S_2} X 11 sample sizes X 3 conditions for the number of measurement points X 3 positions for the turning point) conditions. The first transition point enabled us to study situation where transition occurs early in the longitudinal process (e.g., Kwok et al., 2010), which can be a realistic condition for treatment studies with multiple baseline measurements (i.e. initial evaluation and intake). This condition represents a suboptimal LCM model where the first slope is estimated based on only two measurement points and is thus only identified here because the errors are specified as homoscedastic. However, clinical studies seldom include more than two baseline measures (and often a single one, precluding PWL-LCM analyses altogether) so we deemed it of practical significance to verify the properties of this condition. We note however that "suboptimal" does not mean that this model is problematic in any way as the homoscedasticity of the residuals is an assumption commonly made in LCM studies (Chou et al., 1998; Cheong, 2011; Fan & Fan, 2005; Sun & Willson, 2009). The last two conditions reflect a

situation where the transition occurs later in the longitudinal process. In all conditions where the

effect of the position of the turning point was manipulated, both ϕ_{IS_2} and $\phi_{S_1 S_2}$ were fixed to 0.

Correlation of the second slope growth factor with the intercept and the first slope. To

investigate the impact of ϕ_{IS_2} and $\phi_{S_1 S_2}$ on the power, Type I error rates and rates of nonconvergence,

we generated data with 4 different values for correlation between I and S_2 [corr(I,S_2)] as well for the

correlation between S_1 and the S_2 [corr(S_1,S_2)]: 0, .1, .25 and .5 to reflect no, low, moderate and high

correlation. Each of these conditions was investigated in combination with all of the levels on the

main factors, for a total of 6336 [4 values of μ_{S_2} X 3 levels of ϕ_{S_2} X 11 sample sizes X 3 conditions

for the number of measurement points X 4 values of corr(I,S_2) X 4 values of corr(S_1,S_2)] conditions.

In all conditions where the effect of the correlations among the growth factors was investigated, the

turning point was located at the third measurement occasion.

Data generation and Analysis.

In other to ensure that stability of the Monte Carlo simulation has been reached, 10,000

replications were generated for each design cell. First, we generated data for each cell and recorded

the percentage of the nonconverging samples before discarding those samples. Second, new samples

were generated until 10,000 converging samples were obtained for each cell. All simulations were

conducted using the Mplus 6.11 statistical package (Muthén & Muthén, 2011) and the true models

were always estimated. The outcome variables were the empirical power to detect μ_{S_2} and $\mu_{S_1} - \mu_{S_2}^2$

when they are in fact different from 0 in the population model, the empirical Type I error rates when

μ_{S_2} or $\mu_{S_1} - \mu_{S_2}$ are equal to 0 in the population model, and the proportion of inadmissible solutions

(e.g., nonpositive definite variance-covariance matrix). Because of the strong ceiling effects of power

data, parametric and nonparametric analyses were conducted and results were compared. The

Kruskal-Wallis (K-W) test agreed with the parametric one-way ANOVAs most of the time. These

results are reported in the Table. Extensive results tables are available in the online supplements

accompanying this manuscript. This simulation was made possible through the possibility to access

Compute Canada high performance computing facilities (https://computecanada.ca/).

Results

Type I error rates for μ_{S_2} and for $\mu_{S_1} - \mu_{S_2}$.

Type I error rates were significantly related to the sample size, the number of measurement points, and the variability of S_2. The effects of the other conditions [position of the turning point, and the size of corr(I, S_2) and corr(S_1, S_2)] were not statistically significant predictors of Type I error rates associated with μ_{S_2}, but the size of corr(S_1, S_2) did influence Type I errors rates associated with $\mu_{S_1} - \mu_{S_2}$. Type I error rates remained low across conditions, ranging between 3% and 6% with very few cells reaching 7%. The average Type I error rates were all close to the nominal value of .05 across all conditions. For instance, for μ_{S_2}, the average errors rates did not differ across conditions and were .046, .047 and .045 (.051, .050, and .051 for $\mu_{S_1} - \mu_{S_2}$) respectively for a turning point located at the second, third, and fourth measurement points. The number of measurement points slightly increased Type I error rates to .043, .049 and .048 (.047, .052, and .052 for $\mu_{S_1} - \mu_{S_2}$) respectively for models including 6, 8 and 10 measurement points. Significantly larger Type I error rates were associated with the smallest sample sizes (n =30, for μ_{S_2} M = .057, for $\mu_{S_1} - \mu_{S_2}$ M = .064). The average errors rates are not influenced by the correlation between I and S_2 and are equal to .048, .047, .047 and .047 (.051, .05, .05 and .049 for $\mu_{S_1} - \mu_{S_2}$) respectively when corr(I, S_2) equal 0, .1, .25, and .5 across conditions. Conversely, the Type I error rate associated with $\mu_{S_1} - \mu_{S_2}$ [.051, .050, .050, and .049 for $\mu_{S_1} - \mu_{S_2}$ when corr(S_1, S_2) equal 0, .1, .25, and .5], but not μ_{S_2} (.047, .049, .047 and .046), shows a significant, yet negligible, association with the size of corr(S_1, S_2). Finally, the average Type I error rate was .046 when S_2 showed a small variation (SD = .22), .046 when S_2 showed a moderate variation (SD = .4) and .048 when S_2 showed a large variation (SD = .6) (.049, .051, and .05 for $\mu_{S_1} - \mu_{S_2}$).

Power to detect μ_{S_2}

Empirical power rates were significantly influenced by level and variability of S_2 and sample size. However, the number of measurement points, the position of the turning point and the size of the

correlation between I and S_2 as well between S_1 and S_2 did not affect the power to detect μ_{S_2}.

Level and variability of S₂. Consistent with statistical theory, power rates were affected by μ_{S_2}, and the SD of S_2. Power increased as μ_{S_2} increased but decreased as the SD of S_2 increased. The average power was of .845 when μ_{S_2} was .11 and 1 when μ_{S_2} was .55. Similarly, the average power was .977 when S_2 showed a small variation (SD = .22), .926 when S_2 showed a medium variation (SD = .4) and .865 when S_2 showed a large variation (SD = .6) (all differences significant).

Sample size. Power increased as a function of sample size, varying from .692 (averaged across conditions) for $n = 30$ to 1 for $n = 3000$, and reached an acceptable level of .80 at $n = 100$.

Summary. Figure 3 presents these results for the conditions where the μ_{S_2} was .11, with power curves presented as a function of sample size, SD of S_2, and the number of time points. Since power was not related to the position of the turning point or corr(I, S_2) and corr(S_1, S_2), power curves were plotted for the main fully factorial condition where the turning point was located at the third measurement occasion, and corr(I, S_2) and corr(S_1, S_2) were fixed to zero. Power curves were not plotted when μ_{S_2} was .55 as power rates remained systematically close to 1 in this condition. These curves show that sample sizes as low as 30 are sufficient to detect a large μ_{S_2}. However, larger sample sizes are required to detect smaller μ_{S_2}. To ensure a satisfactory power rate higher than .80 for the detection of a small μ_{S_2} showing substantial inter-individual variability, our results suggest that PWL studies should be based on samples of at least $n = 200$.

Power to detect $\mu_{S_1} - \mu_{S_2}$.

Empirical power to detect $\mu_{S_1} - \mu_{S_2}$ was statistically related to the size of the difference, sample size, variability of S_2, position of the turning point, number of measurement points, and the correlation between S_1 and S_2, but not to the correlation between the I and S_2.

Slope mean difference and variability of S₂. Power increased as the size of $\mu_{S_1} - \mu_{S_2}$ increased. While the average power was satisfactory across conditions with a moderate (.16) and large (-.39) mean difference, it was very difficult to detect small mean differences (.05). More specifically,

the average power was of .472 with a small mean difference (.05), .861 with a medium mean

difference (.16), and .992 with a large mean difference (-.39). The power to detect $\mu_{S_1} - \mu_{S_2}$ was also

negatively affected by the variability of S_2 (all differences significant): .809 for a small variation (SD

= .22), .782 for a moderate variation (SD = .4) and .734 for a large variation (SD = .6).

Sample size. Power increased as a function of sample size, varying from .462 (averaged

across conditions) for $n = 30$ to .995 for $n = 3000$, and reached an acceptable level of .80 at $n = 500$.

Turning point. Power significantly increased as a function of the number of measurement

points before the turning point, with average power of .700, .780, and .767 for turning points

occurring at the second, third and fourth measurement occasions. Power rates for turning points

occurring at the third and fourth occasions did not significantly differ from one another and were

significantly higher than those associated with a turning point occurring at the second occasion.

Number of measurement points. Power was positively and significantly affected by the

number of measurement points. Furthermore, while the results show that average power remained

slightly under the acceptable level of .8 across conditions for all time points, a statistically significant

difference was noted between 6 measurement points (.757) versus 8 (.778) or 10 (.790).

Correlation of the two slopes. Power was affected by the correlation between S_1 and S_2. The

average power was .753 with no correlation between S_1 and S_2, .769 with a small correlation of .1,

.783 with a moderate correlation of .25, and .811 with a large correlation of .5. A statistically

significant difference was observed between the high correlation and the other conditions, as well as

between the no and moderate conditions. No other differences proved significant.

Summary. Since power rates were only slightly affected by the size of the correlation

between S_1 and S_2, power curves are reported in Figures 4 to 6 for the main fully factorial condition

where the turning point was located at the third measurement occasion and corr (S_1, S_2) was fixed to

zero. There were no significant differences between this situation and the situation where the turning

point was located at the fourth occasion, and only a slight decrease in power was observed when the

turning point was located at the second occasion. To illustrate this (small) difference, power curves

for the suboptimal situation where the turning point occurred at the second measurement point are

reported at the end of the online supplements. In summary, these curves show that, when the turning point is located at least at the third measurement occasion, sample sizes of $n = 1500$, but ideally $n = 2000$ are needed to detect small slope mean differences, at least $n = 200$ to detect moderate slope mean differences, and as low as $n = 30\text{-}50$ to detect large slope mean differences. The comparative figures when only two measurement occasions are available before the turning point were of $n = 2000$ but ideally $n = 3000$ to detect a small difference, $n = 200$ but ideally $n = 300$ for a moderate difference, and $n = 50$ for a large difference. These highly discrepant results as a function of the size of the slope mean differences clearly shows that researchers would do well to design their studies with clear a priori expectations regarding the size of the slope mean differences they wish to detect.

Convergence.

The percentage of nonconverging samples was statistically related to sample size, number of measurement points, position of the turning point, and size of the correlation between I and S_2 and between S_1 and S_2. For the variation of S_2 the nonparametric and parametric tests did not agree, but the results show this effect to be relatively small. The size of $\mu_{S_1} - \mu_{S_2}$ did not affect nonconvergence.

Turning point. The rates of nonconvergence were negatively affected by the position of the turning point, being higher when the turning point was located earlier, and decreasing dramatically as the turning point position increased (i.e., 12.67%, 4.4%, 1.55% for turning points located at the second, third and fourth measurement points). This result suggests that rates of nonconvergence are impacted by the number of measurement occasions available before the turning point, but not after the turning point, and only when the number of measurement points available to estimate the first piece is not fully optimal. Indeed, when one compares rates of nonconvergence with only six measurement points, the rates of nonconvergence remain considerably higher when the turning point is at the second occasion (14.5%) than at the third (5.4%) or fourth occasion (2.6%).

Sample size. Rates of nonconvergence decreased as sample size increased, varying between 27.47% and 5% for n < 100, but systematically less than 1% for sample sizes > 200.

Number of measurement points Rates of nonconvergence were significantly decreased when the number of measurement points increased (5.74%, 4.17% and 4.14% for 6, 8 and 10

occasions, averaged across conditions). A significant difference was observed between 6 and all other measurement occasions but not between 8 and 10 measurement occasions.

Correlations. Rates of nonconvergence were slightly related to the size of the correlation between I and S_2. The rates of nonconvergence, averaged across conditions, were 5.36%, 4.44%, 4.1%, and 4.33% for no, small (.1), moderate (.25), and large (.5) correlations between I and S_2. The conditions of small, moderate and high correlations were not significantly different from one another whereas the condition of no correlation was different from all other conditions. Rates of nonconvergence also varied as function of the correlation between S_1 and S_2. More precisely, rates of nonconvergence were respectively 4.68% (no correlation), 3.68% (small correlations), 3.94% (moderate correlation), and 6.44% (large correlation). The small and moderate correlations conditions did significantly differ from one another but all other conditions were different one another.

Variability of S_2. According to the parametric test, rates of nonconvergence were slightly but significantly related to the variability of S_2. Average rates of nonconvergence, were 5.21% when S_2 had a small level of variation, 4.51% when S_2 had a moderate level of variation and 4.34% when S_2 had a large level of variation. The conditions of moderate and large variation were not significantly different from one another but significantly lower than the condition with small variation.

Summary. The rates of nonconvergence were less than 5 % across conditions. Mainly, these rates were higher for smaller sample sizes and when the turning point was closer to the beginning. This suggests that, in order to avoid higher than acceptable risks of nonconvergence, researchers should attempt to collect at least three measures before the expected turning point. In these cases, sample sizes as low as $n = 100$ result in acceptable rates of nonconvergence. Thus, if the choice is between collecting more measures before, or after, the expected turning point, the results suggest that it is better to invest before reaching the turning point, at least in terms of convergence.

Discussion

This study aimed to provide more precise guidelines in terms of sample size requirements for research using PWL models in order to estimate the magnitude of two joint longitudinal processes and the magnitude of the growth differences between these two processes. To this end, we examined empirical power rates as a function of the number of repeated measurements, sample size, size of the

second slope, levels of inter-individual variability in the second slope, position of the turning point, and correlations between the growth factors of the PWL trajectory. We similarly investigated the rates of nonconvergence and Type I error as a function of these same design factors.

Our results generally supported our hypotheses regarding the impact of some of the design factors on Type I error rates, power and nonconvergence. The empirical Type I error rates were all very close to the nominal value of 5% and fluctuated normally around this value. Type I error rates related to the second slope as well the mean differences between the slopes were only significantly related to sample size and to the number of measurement points. In particular, Type I error rates decreased as a function of the number of measurement occasions and sample size. Overall, the results show that the PWL-LCM seldom ends up falsely detecting PWL growth or significant growth differences when none are present in the data, at least based on the conditions simulated in this study.

For power, consistent with statistical theory, the results showed that power estimates were larger when the size of the effect (μ_{S_2} or $\mu_{S_1} - \mu_{S_2}$) was larger, sample sizes larger, and with smaller levels of inter-individual variability on the second slope factor. However, the correlation between the intercept and the second slope were not related to the power to detect the second slope or the slope mean difference. Regarding the number of measurement points, the position of the turning point and the correlation between the first and second slope, they were all significantly related to the power to detect the slope mean difference but not to the power to detect the second slope. Thus, having a turning point located at the second measurement point resulted in slight decreases in the power to detect differences between two slopes whereas increases in the correlation between the slopes resulted in increases in the power to detect the slope mean difference. Our results showed that relatively low sample sizes were required to achieve a satisfactory .80 level in power when the objective was only to detect whether the second slope was significant. This could be a reasonable objective of studies using PWL simply to model distinct trajectories (including different sets of predictors and outcomes) before and after some transition point. More precisely, our results show that $n = 30\text{-}50$ is sufficient to detect a large S_2 with power levels over .80 whereas $n = 200$ is required to detect a smaller slope. However, researchers may not only be interested in modeling distinct, yet interrelated, developmental processes.

Rather, they may also want to test whether the level of growth significantly change before and after the turning point. Our results show that larger sample sizes are required to achieve satisfactory power in this context and that researchers would do well to a priori define the size of the effect they want to detect. More precisely, sample sizes of $n = 1500\text{-}2000$ are needed to detect a small slope mean difference with a power levels of .80, at least $n = 200$ to detect a moderate difference, and only $n = 30\text{-}50$ to detect a large difference, unless only two measurement occasions are available before the turning point, in which case the required sample sizes rise to $n = 2000\text{-}3000$, $n = 200\text{-}300$, and $n = 50$.

Although the position of the turning point only had a small effect on power, this factor had an a major influence on rates of nonconvergence, which were highly elevated in the condition where only two measurement points are available before the turning point. Thus, in the specific condition where only two measurement points are available before the turning point, sample sizes of $200 < n < 300$ seem to be a more reasonable guideline for researchers who aim to avoid an unreasonable risk of nonconvergence. This result also suggest that experimental or clinical studies would do well to conduct clear preliminary analyses of the relative costs of having to increase the sample size, or the number of measurement occasions before the beginning of treatment.

To our knowledge, this is the first LCM study to consider the variability of the slope in a power study. Our results clearly show the importance of this design factor in the determination of empirical power rates and the importance of incorporating this design factor in future simulations studies. Interestingly, this study also shows that some results based on linear LCM (e.g. Fan & Fan, 2005) cannot fully generalize to nonlinear models defined by two linear slopes. Indeed, our results showed that characteristics of the full PWL model, such as the number of measurement points used to estimate the first slope factor do not have a substantial impact on the power to detect the second slope factor but influence more clearly the detection of the slope mean difference. However, we cannot expect our results to generalize to other forms of nonlinear relations. We also observed, as reported by Fan and Fan (2005) for linear LCM, that PWL estimation was associated with similarly high rates of nonconvergence. Thus, adding an additional slope to capture the nonlinear component of change did not result in an increased probability of encountering improper solutions in the estimation process. Whether this conclusion will hold for other nonlinear functional forms (exponential, logistic, etc.) that

do not involve adding an additional growth factor to the model remains to be seen in future studies.

This study represents the first systematic investigation of the performance of PWL models based on simulated data. In order to be as comprehensive as possible, we strived to incorporate the broadest range of design conditions shown to be important to power determination within previous LCM studies, as well as multiple conditions likely to be relevant to the specific PWL context. This resulted in a very comprehensive study based on a total of 7524 design conditions, each based on 10,000 properly converged samples. However, some conditions obviously had to be put aside to maintain this study within reasonable boundaries but should be more systematically investigated in the context of future studies which will hopefully be able to pick up a reduced range of design conditions and levels within conditions based on the results from the current study. These directions for future research include, multilevel PWL models were the turning point is an estimated parameter or the systematic comparisons of linear, latent basis, and PWL models for data simulated to correspond to either a linear or a PWL growth process. Similarly, the current study relied on complete data that met multivariate normality and homoscedasticity assumptions. However, real data tend to include missing data that may follow different mechanisms (e.g., missing at random), to be nonnormal and to be heteroscedastic. It is not known to what extent the result presented in this study will hold for these various conditions but these clearly represents areas where future research is required. Our expectation is that larger sample sizes may be required in these conditions.

Endnotes

[1] Power is known to be influenced by the magnitude of the effect to be detected so that the magnitude of μ_{S_2} and $\mu_{S_1} - \mu_{S_2}$ reflects one of the critical element to consider in this study. However, another indicator of the magnitude of the effect associated with the full LCM model (less relevant here to the detection of a specific parameter) is the R^2 of the repeated measures. R^2 values are a function of the time score, the variances and covariances of the growth factors and the variances of the residuals. More precisely, R^2 values of the PWL model can be calculated using the following formula:

$$R^2 (y_t) = \frac{\phi_I + \lambda^2_{1t}\phi_{S_1} + \lambda^2_{2t}\phi_{S_2} + 2\lambda_{1t}\phi_{IS_1} + 2\lambda_{2t}\phi_{IS_2} + 2\lambda_{1t}\lambda_{2t}\phi_{S_1 S_2}}{\phi_I + \lambda^2_{1t}\phi_{S_1} + \lambda^2_{2t}\phi_{S_2} + 2\lambda_{1t}\phi_{IS_1} + 2\lambda_{2t}\phi_{IS_2} + 2\lambda_{1t}\lambda_{2t}\phi_{S_1 S_2} + \theta_t},$$

where y_t is the outcome at time t, λ_{1t} is the time score for S_1, λ_{2t} is the time score for S_2, ϕ_I is the inter-

individual variance of I, ϕ_{S_1} is the inter-individual variance of S_1, ϕ_{S_2} is the inter-individual variance of S_2, ϕ_{IS_1} is the covariance between I and S_1, ϕ_{IS_2} is the covariance between the I and S_2, $\phi_{S_1 S_2}$ is the covariance between S_1 and S_2, and θ_t is the time specific residual variance. Thus, R² is essentially composed by two terms: $\phi_I + \lambda^2_{1t}\phi_{S_1} + \lambda^2_{2t}\phi_{S_2} + 2\lambda_{1t}\phi_{IS_1} + 2\lambda_{2t}\phi_{IS_2} + 2\lambda_{1t}\lambda_{2t}\phi_{S_1 S_2}$ and θ_t. The first term is strictly increasing over time given that all growth factors variances and covariances are greater than zero in this study. The second term is constant due to the homoscedastic assumption ($\theta_t = \theta$) of this study. As a result the R² strictly increases over time. Based on this formula, it is clear that the specific R² values change across design conditions under the influence of multiple design conditions, making it complex to report specific values associated with each of the time points within each of the 7524 design conditions. Overall, R² values range from .5 for the first time point to .99 for the tenth time point. This increase over time is reasonable as more time point usually provide more precision in the estimation of the underlying trajectories. However, the higher values observed at later time points might have resulted in slightly inflated power estimates.

[2] In order to test the statistical significance of $\mu_{S_1} - \mu_{S_2}$, we relied on a Wald test of statistical significance. Alternatively, Likelihood ratio tests (LRT), which are well suited for situations where the normality assumption is met and multiple parameters are estimated, could also have been used. The Wald test is routinely provided by most software making it simpler to use than the comparison of models based on LRTs. For this reason, the Wald test is more frequently used in practice. However, the squared version of the Wald test and the Likelihood ratio test are asymptotically equivalent and follow a chi-square distribution of one degree of freedom (Bollen, 1989; DasGupta, 2008). These two test should thus give similar results in most situations. In fact, small differences may potentially be expected when the asymptotic equivalence between the Wald and the Likelihood ratio test no longer hold, such as in small sample sizes. The specific conditions where this equivalence breaks down should clearly be more systematically investigated in the context of future studies.

References

Beauchaine, T. & Beauchaine, R. (2002). A comparison of maximum covariance and k-means cluster analysis in classifying cases into known taxon groups. *Psychological Methods, 7*, 245-261.

Blozis, S. A. (2007). On fitting nonlinear latent curve models to multiple variables measured longitudinally. *Structural Equation Modeling*, *14*, 179-201.

Bollen, K. A. (1989). *Structural Equations with Latent variables.* New York, NY: Wiley.

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

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.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear model: Applications and data analysis methods.* Newbury Park, CA: Sage.

Cheong, J. (2011). Accuracy of estimates and statistical power for testing mediation in latent growth curve modeling. *Structural Equation Modeling*, *18*, 195-211.

Cheong, J., MacKinnon, D. P., & Khoo, S. T. (2003). Investigation of meditational processes using parallel process latent growth curve modeling. *Structural Equation Modeling*, *10*, 238-262.

Chou, C.-P, Bentler, P. M., & Pentz, M. A. (1998). Comparisons of two statistical approaches to study growth curves: The multilevel model and the latent curve analysis. *Structural Equation Modeling*, *5*, 247-266.

Cudeck, R. & Harring, J.R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology* 58, 615-637.

Cudeck, R. & Klebe, K.J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods* 7, 41-63.

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, *38*, 529-569.

DasGupta, A. (2008). *Asymptotic theory of statistics and probability.* New York, NY: Springer.

Duncan, T. E., Duncan, S.C., Strycker, L. A., & Li, F. (2002). A latent variable framework for power estimation within intervention contexts. *Journal of Psychopathology & Behavioral Assessment, 24*, 1-12.

Eccles, J.S., Midgley, C., Wigfield, A., Buchanan, C.M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, *48*, 90-101.

Fan, X. (2003). Power of latent growth modeling for detecting group differences in linear growth trajectory parameters. *Structural Equation Modeling, 10*, 380-400.

Fan, X., & Fan, X. (2005). Power of latent growth modeling for detecting linear growth: Number of measurements and comparison with other analytical approaches. *The Journal of Experimental*

Education, 73, 121-139.

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

Flora, D. B. (2008): Specifying Piecewise Latent Trajectory Models for Longitudinal Data. *Structural Equation Modeling*, *15*, 513-533.

Grimm, K. J., Ram, N., & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development, 82,* 1357-1371.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rational and application of multivariate stationary time series analyses. *Multivariate Behavioral Research*, *40*, 207-233.

Hertzog, C., Lindenberger, U., Ghisletta, P., & Oertzen, T. v. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods, 11,* 244-252.

Hertzog, C., Oertzen, T. v., Ghisletta, P., & Lindenberger, U. (2008). Evaluating the power of latent growth curve models to detect individual differences in change. *Structural Equation Modeling*, *15*, 541-563.

Jones, C. J., & Meredith, W. (2000). Developmental paths of psychological health from early adolescence to later adulthood. *Psychology and Aging*, *15*, 351-360.

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.

Kwok, O., Luo, W., & West, S. G. (2010). Using modification indexes to detect turning points in longitudinal data: A Monte Carlo study. *Structural Equation Modeling*, *17*, 216-240.

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

Morin, A. J. S., Maïano, C., Nagengast, B., Marsh, H. W., Morizot, J., & Janosz, M. (2011). Growth mixture modeling of adolescents trajectories of anxiety across adolescence: The impact of untested invariance assumptions on substantive interpretations. *Structural Equation Modeling, 18,* 613-648.

Muthén, B. O., & Curran, P. (1997). General growth modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2*, 371-402.

Muthén, L. K., & Muthén, B. O. (2011). Mplus user's guide (6th ed). Los Angeles: Muthén & Muthén.

Ram, N., & Grimm, K. J. (2007). Using simple and complex growth models to articulate developmental change: Matching theory to method. *International Journal of Behavioral Development*, *31*, 303-316.

Shadish, W.R., & Cook, T.D. (2009). The renaissance of field experimentation in evaluating interventions. *Annual Review of Psychology*, *60*, 607-629.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Sun, R., & Willson, V. L. (2009). Evaluating intercept-slope interactions in latent Growth modeling. *Structural Equation Modeling*, *16*, 226-244.

Tofighi, D., & Enders, C. K. (2007). Identifying the correct number of classes in growth mixture models. In G. R. Hancock (Ed.), *Mixture models in latent variable research* (pp. 317-341). Greenwich, CT: Information Age.

Voelkle, M. C. (2008). Reconsidering the use of Autoregressive Latent Trajectory (ALT) Models. *Multivariate Behavioral Research*, *43*, 564-591.

Zhang, Z., & Wang, L. (2009). Statistical power analysis for growth curve models using SAS. *Behavior Research Methods*, *41*, 1083-1094.

Table. Main results regarding the effects of the design factors on the Type I error rates, empirical power rates and rates of nonvergence.

| Design factor	Parametric ANOVA F (df)	p	Non-Parametric Kruskal-Wallis χ^2(df)	p
Type I error rate (μ_{S2})				
Slope Variance	$F(2, 1878) = 3.114$	.045	$\chi^2(2) = 7.202$	.027
Sample Size	$F(10, 1870) = 15.681$	.000	$\chi^2(10) = 594.931$	.000
Number of measurement points	$F(2, 1878) = 21.988$	.000	$\chi^2(2) = 153.327$	.000
Position of the turning point	$F(2, 1878) = .796$	.451	$\chi^2(2) = 5.296$	.071
Size of Corr(I, S_2)	$F(3, 1877) = .589$	.622	$\chi^2(3) = .351$	.950
Size of Corr(S_1, S_2)	$F(3, 1877) = 1.72$	.161	$\chi^2(3) = 4.91$	.179
Type 1 error rate (μ_{S1} - μ_{S2})				
Slope Variance	$F(2, 1878) = 3.946$	.019	$\chi^2(2) = 10.149$	.006
Sample Size	$F(10, 1870) = 83.575$	.000	$\chi^2(10) = 481.561$	.000
Number of measurement points	$F(2, 1878) = 90.239$	.000	$\chi^2(2) = 170.053$	.000
Position of the turning point	$F(2, 1878) = .407$	.666	$\chi^2(2) = .505$	.777
Size of Corr(I, S_2)	$F(3, 1877) = 2.481$	.059	$\chi^2(3) = 7.105$	.069
Size of Corr(S_1, S_2)	$F(3, 1877) = 2.903$	.034	$\chi^2(3) = 8.248$	.041
Power to detect μ_{S2}				
Slope Mean	$F(1, 3758) = 725.69$	.000	$\chi^2(1) = 1034.249$	.000
Slope Variance	$F(2, 3757) = 113.788$	.000	$\chi^2(2) = 147.455$	.000
Sample Size	$F(10, 3749) = 151.893$	.000	$\chi^2(10) = 957.255$	.000
Number of measurement points	$F(2, 3757) = 1.599$	.202	$\chi^2(2) = 4.298$	.117
Position of the turning point	$F(2, 3757) = .114$	.892	$\chi^2(2) = .268$	.875
Size of Corr(I, S_2)	$F(3, 3756) = .005$	1	$\chi^2(3) = .119$	.989
Size of Corr(S_1, S_2)	$F(3, 3756) = .001$	1	$\chi^2(3) = .109$	.991
Power to detect μ_{S1} - μ_{S2}				
Slope Mean	$F(2, 5640) = 2356.822$	.000	$\chi^2(2) = 2943.542$	.000
Slope Variance	$F(2, 5640) = 25.878$	.000	$\chi^2(2) = 80.528$	.000
Sample Size	$F(10, 5632) = 221.481$	.000	$\chi^2(10) = 1680.686$	.000
Number of measurement points	$F(2, 5640) = 4.972$	.007	$\chi^2(2) = 15.333$	.000
Position of the turning point	$F(2, 5640) = 8.396$	.000	$\chi^2(2) = 20.311$	.000
Size of Corr(I, S_2)	$F(3, 5639) = 1.024$	.381	$\chi^2(3) = 4.63$	.201
Size of Corr(S_1, S_2)	$F(3, 5639) = 8.457$	.000	$\chi^2(3) = 39.684$	.000
Rates of nonconvergence				
Slope Mean	$F(3, 7520) = 0$	1	$\chi^2(2) = 0$	1
Slope Variance	$F(2, 7521) = 6.081$	.002	$\chi^2(2) = 5.582$	.061
Sample Size	$F(10, 7513) = 3083.358$	.000	$\chi^2(10) = 6461.699$	.000
Number of measurement points	$F(2, 7521) = 23.902$	.000	$\chi^2(2) = 15.335$	.000
Position of the turning point	$F(2, 7521) = 175.893$	.000	$\chi^2(2) = 223.87$	.000
Size of Corr(I, S_2)	$F(3, 7520) = 7.937$	.000	$\chi^2(3) = 8.588$	.035
Size of Corr(S_1, S_2)	$F(3, 7520) = 28.187$	.000	$\chi^2(3) = 69.258$	.000
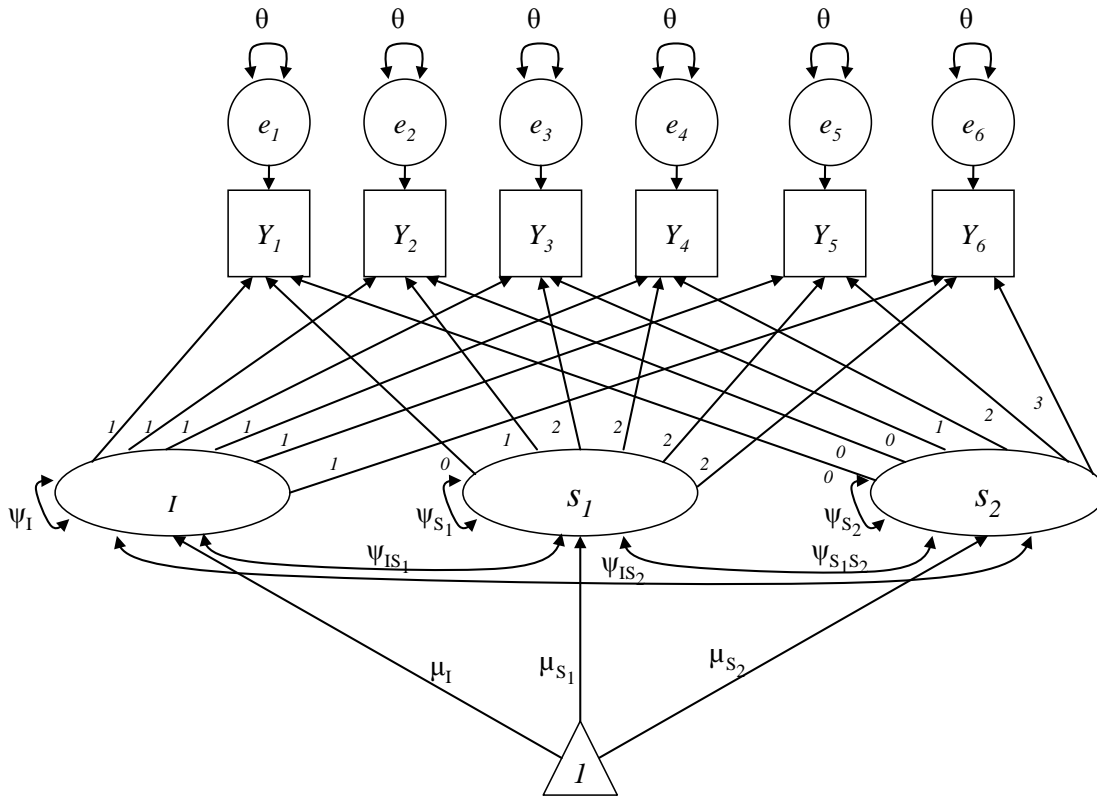
Figure 1. Graphical Presentation of a Piecewise LCM with a Turning Point at 3 measurement
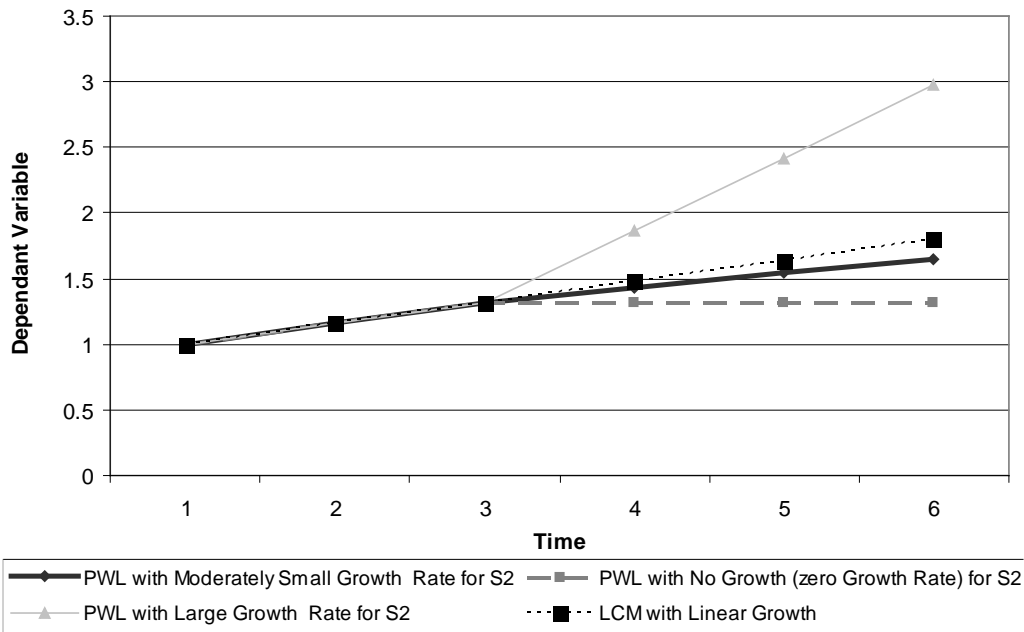
occasions.

Figure 2. Illustration of the data generation conditions for the zero, moderately small, and large

growth rates for the second slope of the PWL-LCM with a Turning Point at 3 measurement occasions.
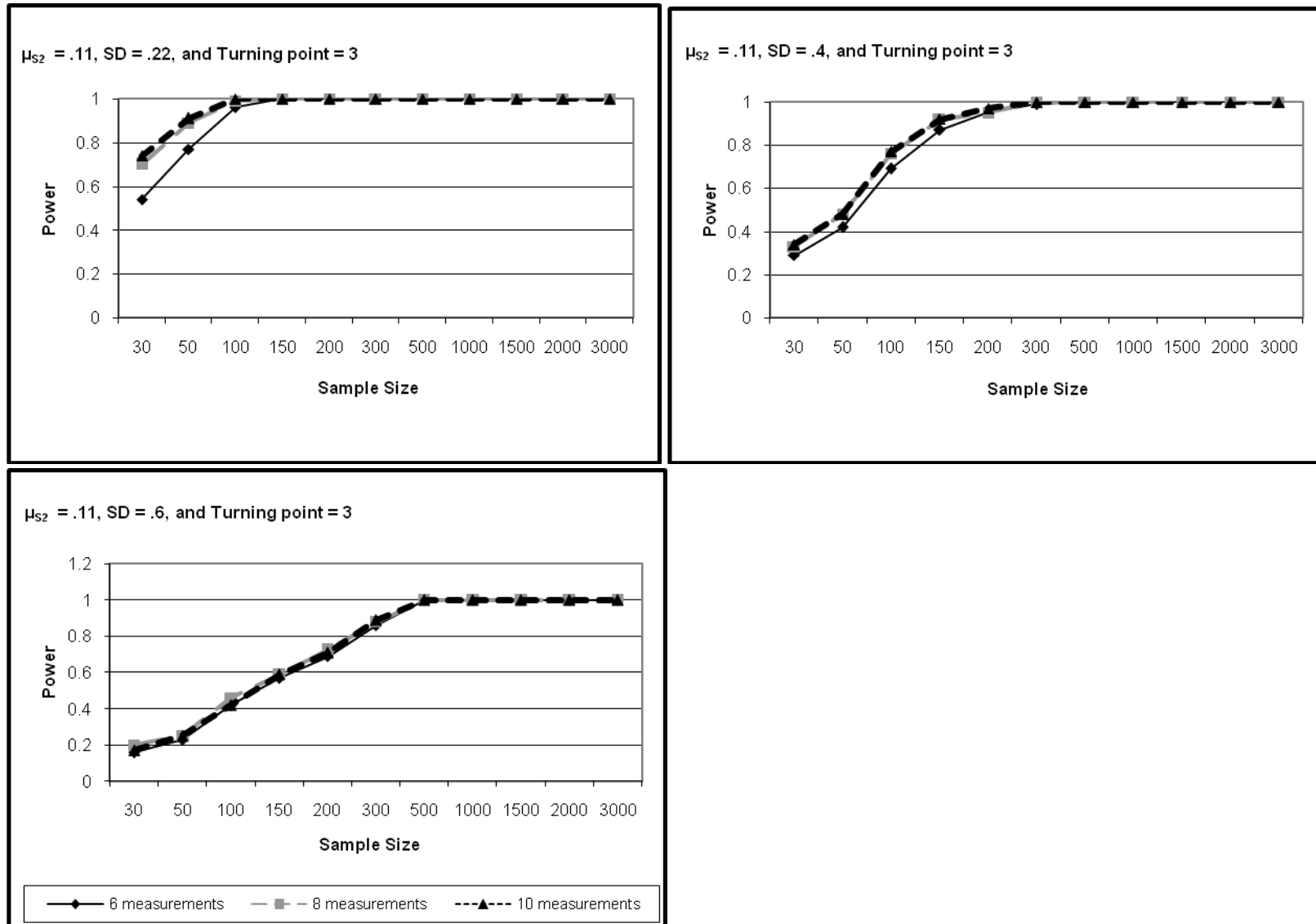
Figure 3. Empirical Power Curves to Detect a S_2 Mean of .11, for a Turning Point at Time 3 with Corr(I,S_1) = 0 and Corr(S_1,S_2) = 0.
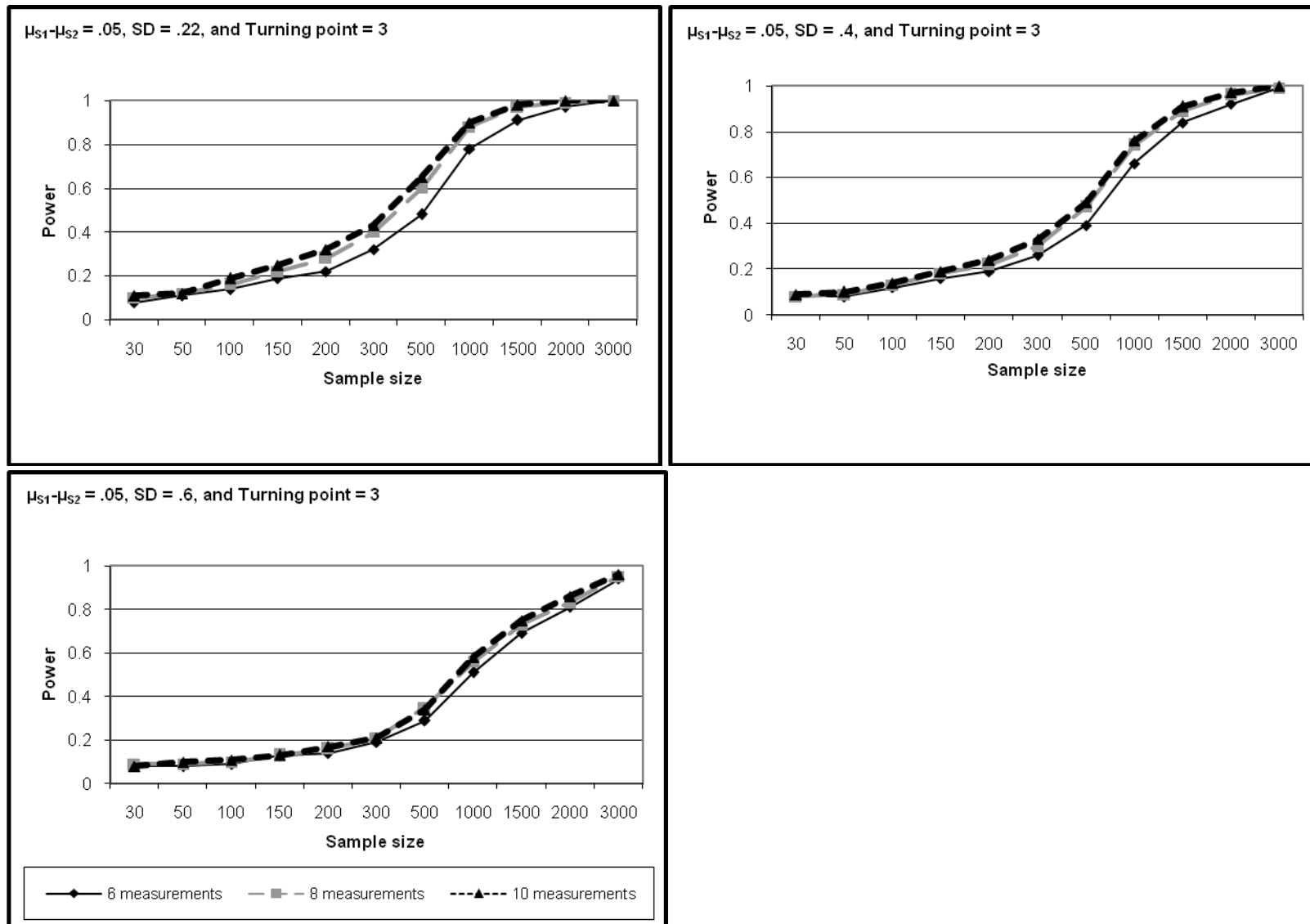
Figure 4. Empirical Power Curves to Detect a Mean Difference of .05, for a Turning Point at Time 3 with $Corr(I,S_1) = 0$ and $Corr(S_1,S_2) = 0$.
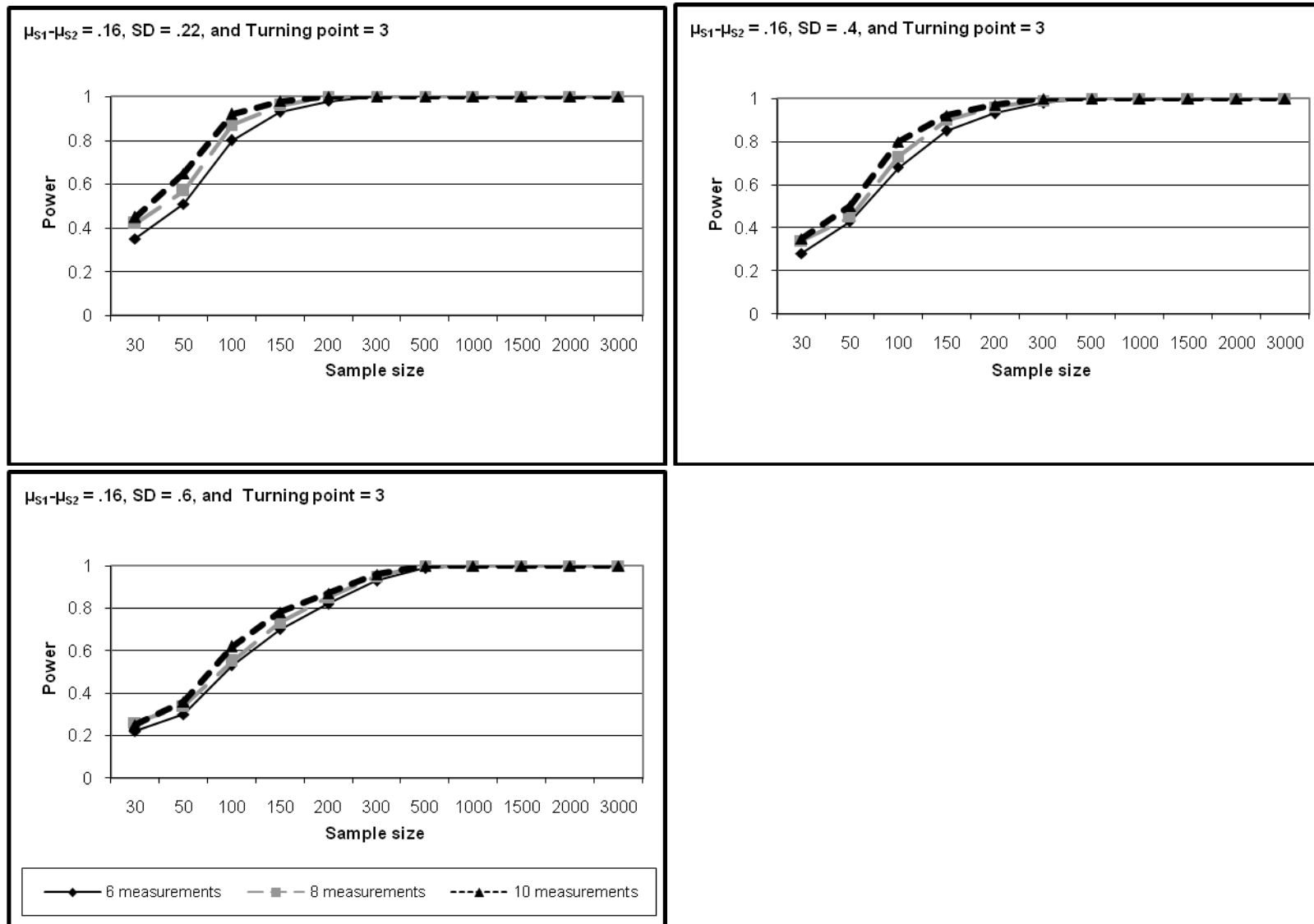
Figure 5. Empirical Power Curves to Detect a Mean Difference of .16, for a Turning Point at Time 3 with $Corr(I,S_1) = 0$ and $Corr(S_1,S_2) = 0$.
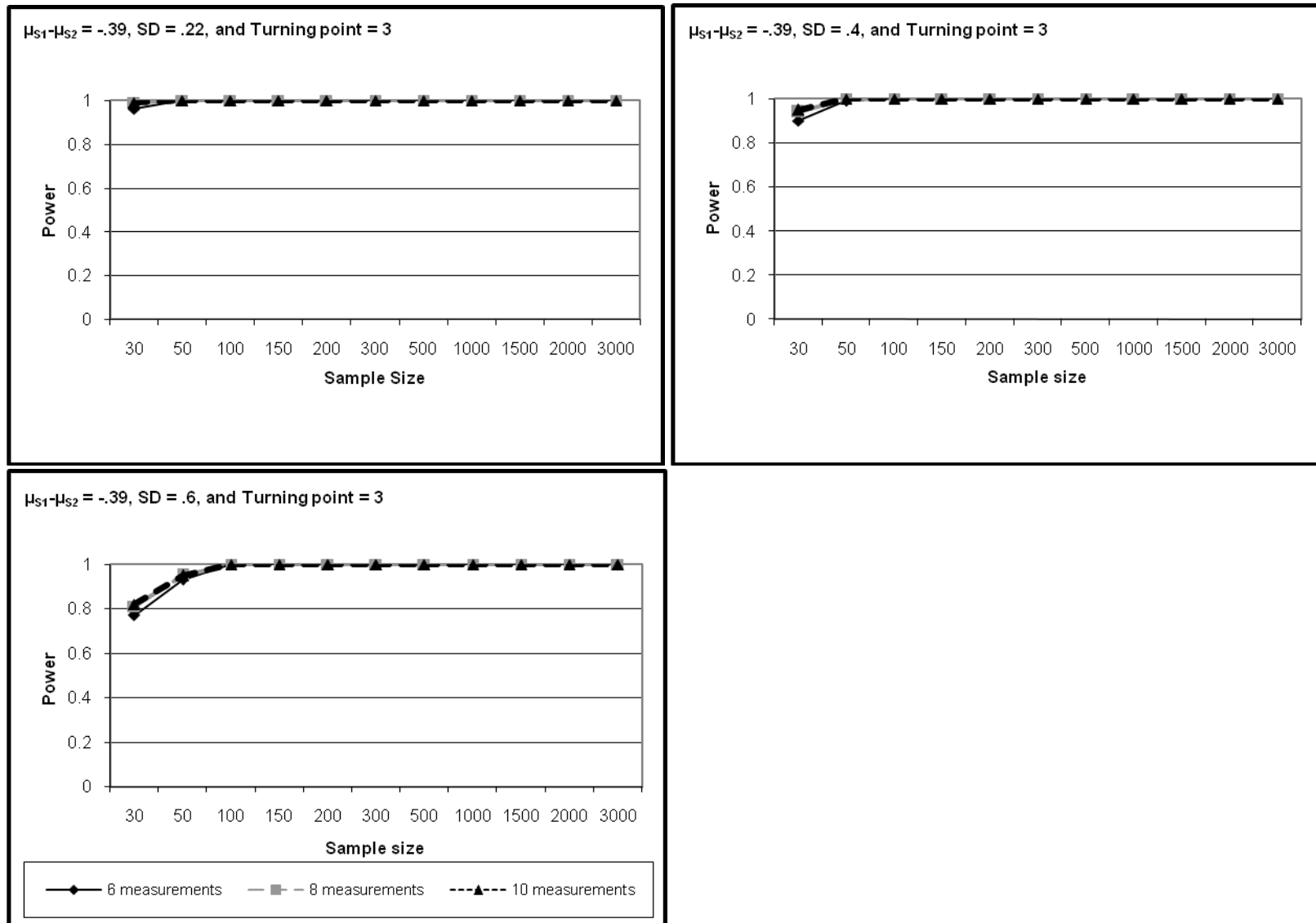
Figure 6. Empirical Power Curves to Detect a Mean Difference of -.39, for a Turning Point at Time 3 with Corr(I,S$_1$) = 0 and Corr(S$_1$,S$_2$) = 0.