Is Timing Everything? Measurement Timing and the Ability to Accurately Model Longitudinal Data

by

Sebastian L.V. Sciarra

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of

Doctorate of Philosophy

in

Psychology

Guelph, Ontario, Canada

ABSTRACT

IS TIMING EVERYTHING? MEASUREMENT TIMING AND THE ABILITY TO ACCURATELY MODEL LONGITUDINAL DATA

Sebastian L.V. Sciarra
University of Guelph, 2022

Advisor(s):
David Stanley

The preface pretty much says it all. This is additional content. The preface pretty much says it all. This is additional content. The preface pretty much says it all. This is additional content. The preface pretty much says it all. This is additional content. The preface pretty much says it all. This is additional content.

DEDICATION

You can have a dedication here if you wish. You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.

ACKNOWLEDGEMENTS

I want to thank a few people.You can have a dedication here if you wish. You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.You can have a dedication here if you wish.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1 Introduction

"Neither the behavior of human beings nor the activities of organizations can be defined without reference to time, and temporal aspects are critical for understanding them" (Navarro et al., 2015, p. 136).

The topic of time has received a considerable amount of attention in organizational psychology over the past 20 years. Examples of well-received articles published around the beginning of the 21st century discuss how investigating time is important for understanding patterns of change and boundary conditions of theory (Zaheer et al., 1999), how longitudinal research is necessary for disentangling different types of causality (T. R. Mitchell & James, 2001), and explicate a pattern of organizational change (or institutionalization; Lawrence et al., 2001). Since then, articles have emphasized the need to address time in specific areas such as performance (Dalal et al., 2014; C. D. Fisher, 2008), teams (Roe et al., 2012), and goal setting (Fried & Slowik, 2004) and, more generally, throughout organizational research (Aguinis & Bakker, 2021; George & Jones, 2000; Kunisch et al., 2017; Navarro et al., 2015; Ployhart & Vandenberg, 2010; Roe, 2008; Shipp & Cole, 2015; Sonnentag, 2012; Vantilborgh et al., 2018).

The importance of time has also been recognized in organizational theory. In defining a theoretical contribution, Whetten (1989) discussed that time must be discussed in regard to setting boundary conditions (i.e., under what circumstances does the theory apply) and in specifying relations between variables over time (George & Jones, 2000; see also T. R. Mitchell & James, 2001). Even if a considerable number of organizational theories do not adhere to the definition of Whetten (1989), theoretical models in organizational psychology consist of path diagrams that delineate the causal underpinnings of a process. Given that temporal precedence is a necessary condition for establishing causality (Mill, 2011), time has a role, whether implicitly or explicitly, in organizational theory.

Despite the considerable emphasis that has been placed on investigating processes over time and its ubiquity in organizational theory, the prevalence of longitudinal research has historically remained low. One study examined the prevalence of longitudinal research from 1970–2006 across five organizational psychology journals and found that 4% of articles used longitudinal designs (Roe, 2014). Another survey of two applied psychology journals in 2005 found that approximamely 10% (10 of 105 studies) of studies used longitudinal designs (Roe, 2008). Similarly, two surveys of studies employing longitudinal designs with mediation analysis found that, across five journals, only about 10% (7 of 72 studies) did so in 2005 (Maxwell & Cole, 2007) and approximately 16% (15 of 92 studies) did so in 2006 (M. A. Mitchell & Maxwell, 2013).[1] Thus, the prevalence of longitudinal research has remained low.

In the six sections that follow, I will explain why longitudinal research is necessary and the factors that must be considered when conducting such research. In the first section, I will explain why conducting longitudinal research is essential for understanding the dynamics of psychological processes. In the second section, I will overview patterns of change that are likely to emerge over time. In the the third section, I will overview design and analytical issues involved in designing longitudinal studies. In the fourth section, I will explain how design and analytical issues encountered in conducting longitudinal

[1]Note that the definition of a longitudinal design in Maxwell & Cole (2007) and M. A. Mitchell & Maxwell (2013) required that measurements be taken over at least three time points so that measurements of the predictor, mediator, and outcome variables were separated over time.

research can be investigated. In the fifth section, I will provide a systematic review of the research that has investigated design and analytical issues involved in conducting longitudinal research. Finally, in the sixth section, I will briefly explain strategies for modelling nonlinear change. A summary of the three simulation experiments that I conducted in my dissertation will then be provided.

1.1 The Need to Conduct Longitudinal Research

Longitudinal research provides substantial advantages over cross-sectional research. Unfortunately, researchers commonly discuss the results of cross-sectional analyses as if they have been obtained with a longitudinal design. However, cross-sectional and longitudinal analyses often produce different results. Oneexample of the assumption that cross-sectional findings are equivalent to longitudinal findings comes from the large number of studies employing mediation analysis. Given that mediation is used to understand chains of causality in psychological processes (Baron & Kenny, 1986), it would thus make sense to pair mediation analysis with a longitudinal design because understanding causality, after all, requires temporal precedence. Unfortunately, the majority of studies that have used mediation analysis have done so using cross-sectional designs—with estimates of approximately 90% (Maxwell & Cole, 2007) and 84% (M. A. Mitchell & Maxwell, 2013)—and have often discussed the results as if they were longitudinal. Investigations into whether mediation results remain equivalent across cross-sectional and longitudinal designs have repeatedly concluded that using mediation analysis on cross-sectional data can return different, and sometimes completely opposite, results from using it on longitudinal data (Cole & Maxwell, 2003; Maxwell et al., 2011; Maxwell & Cole, 2007; M. A. Mitchell & Maxwell, 2013; O'Laughlin et al., 2018). Therefore, mediation analyses based on cross-sectional analyses may be misleading.

The non-equivalence of cross-sectional and longitudinal results that occurs with mediation analysis is, unfortunately, not due to a specific set of circumstances that only arise with mediation analysis, but a consequence of a broader systematic cause that affects the results of almost every analysis. The concept of ergodicity explains why cross-sectional and longitudinal analyses seldom yield similar results. To understand ergodicity, it is first important to realize that variance is central to many statistical analyses—correlation, regression, factor analysis, and mediation are some examples. Thus, if variance remains unchanged across cross-sectional and longitudinal data sets, then analyses of either data set would return the same results. Importantly, variance only remains equal across cross-sectional and longitudinal data sets if two conditions put forth by ergodic theory are satisfied (homogeneity and stationarity; Molenaar, 2004; Molenaar & Campbell, 2009). If these two conditions are met, then a process is said to be ergodic. Unfortunately, the two conditions required for ergodicity are highly unlikely to be satisfied and so cross-sectional findings will frequently deviate from longitudinal findings (see [Technical Appendix A][Technical Appendix A: Ergodicity and the Need to Conduct Longitudinal Research] for more information).

Given that cross-sectional and longitudinal analyses are, in general, unlikely to return equivalent findings, it is unsurprising that several investigations in organizational research—and psychology as a whole—have found these analyses to return different results. Beginning with an example from Curran & Bauer (2011), heart attacks are less likely to occur in people who exercise regularly (longitudinal finding), but more likely to happen when exercising (cross-sectional finding). Correlational studies find differences in correlation magnitudes between cross-sectional and longitudinal data sets J. Fisher et al. (2018).[2] Moving on to perhaps the most commonly employed analysis in organizational research of mediation, several articles have highlighted cross-sectional data can return different, and sometimes completely opposite, results to longitudinal data (Cole & Maxwell, 2003; Maxwell et al., 2011; Maxwell & Cole, 2007; O'Laughlin et al., 2018). Factor analysis is perhaps the most interesting example: The well-documented five-factor model of personality seldom arises when analyzing person-level data that was obtained by measuring personality on 90 consecutive days (Ellen L. Hamaker et al., 2005). Therefore, cross-sectional analyses are rarely equivalent to longitudinal analyses.

Fortunately, technological advancements have allowed researchers to more easily conduct longitudinal research in two ways. First, the use of the experience sampling method (Beal, 2015) in conjunction with modern information transmission technologies—whether through phone applications or short message services—allows data to sometimes be sampled over time with relative ease. Second, the development of analyses for longitudinal data (along with their integration in commonly used software) that enable person-level data to be modelled such as multilevel models (Raudenbush & Bryk, 2002), growth mixture models (Mo Wang & Bodner, 2007), and dynamic factor analysis (Ram et al., 2013) provide researchers with avenues to explore the temporal dynamics of psychological processes. With one recent survey estimating that 43.3% of mediation studies (26 of 60 studies) used a longitudinal design (O'Laughlin et al., 2018), it appears that the prevalence of longitudinal research has increased from the 9.5% (Roe, 2008) and 16.3%

[2]Note that J. Fisher et al. (2018) also found the variability of longitudinal correlations to be considerably larger than the variability of cross-sectional correlations.

(M. A. Mitchell & Maxwell, 2013) values estimated at the beginning of the 21st century. Although the frequency of longitudinal research appears to have increased over the past 20 years, several avenues exist where the quality of longitudinal research can be improved, and in my dissertation, I focus on investigating these avenues.

1.2 Understanding Patterns of Change That Emerge Over Time

Change can occur in many ways over time. One pattern of change commonly assumed to occur over time is that of linear change. When change follows a linear pattern, the rate of change over time remains constant. Unfortunately, a linear pattern places demanding restrictions on possible patterns of change. If change were to follow a linear pattern, then any pauses in change (or plateaus) or changes in direction would not occur and effects would simply grow over time. Unfortunately, effect sizes have been shown to diminish over time (for meta-analytic examples, see Cohen, 1993; Griffeth et al., 2000; Hom et al., 1992; Riketta, 2008; Steel et al., 1990; Steel & Ovalle, 1984). Moreover, many variables display cyclic patterns of change over time, with mood (Larsen & Kasimatis, 1990), daily stress (Bodenmann et al., 2010), and daily drinking behaviour (Huh et al., 2015) as some examples. Therefore, change over is unlikely to follow a linear pattern.

A more realistic pattern of change to occur over time is a nonlinear pattern (for a review, see Cudeck & Harring, 2007). Nonlinear change allows nonconstant rates of change such that change may occur more rapidly during certain periods of time, stop altogether, or reverse direction. When looking at patterns of change observed across psychology, examples appear in the declining rate of speech errors throughout child development (Burchinal & Appelbaum, 1991), forgetting rates in memory (Murre & Dros, 2015), development of habits over time (Fournier et al., 2017), and the formation of opinions over time (Xia et al., 2020). Given nonlinear change appears more likely than linear change, my dissertation will assume change over time to be nonlinear.

1.3 Challenges Involved in Conducting Longitudinal Research

Conducting longitudinal research presents researchers with several challenges. Many challenges are those from cross-sectional research only amplified (for a review, see Bergman & Magnusson, 1990).[3] For example, greater efforts have to be made to to prevent missing data which can increase over (Dillman et al., 2014; Newman, 2008). Likewise, the adverse effects of well-documented biases such as demand characteristics (Orne, 1962) and social desirability (Nederhof, 1985) have to be countered at each time point. Outside challenges share with cross-sectional research, conducting longitudinal research also presents new challenges. Analyses of longitudinal data have to consider complications such as how to model error structures (Grimm & Widaman, 2010), check for measurement non-invariance over time (the extent to which a construct is measured with equivalent accuracy over time; Schoot et al., 2012), and how to center/process data to appropriately answer research questions (Enders & Tofighi, 2007; Wang & Maxwell, 2015).

Although researchers must contend with several issues in conducting longitudinal research, three issues are of particular interest in my dissertation. The first issue concerns how many measurements to use in a longitudinal design. The second issue concerns how to space the measurements. The third issue focuses on how much error is incurred if the time structuredness of the data is overlooked. The sections that follow will review each

[3]It should be noted that conducting a longitudinal study does alleviate some issues encountered in conducting cross-sectional research. For example, taking measurements over multiple time points likely reduces common method variance (Podsakoff et al., 2003; for an example, see Ostroff et al., 2002).

of these issues.

1.3.1 Number of Measurements

Researchers have to decide on the number of measurements to include in a longitudinal study. Although using more measurements increases the accuracy of results—as noted in the results of several studies (e.g., Coulombe et al., 2016; Finch, 2017; Fine et al., 2019; Timmons & Preacher, 2015)—taking additional measurements often comes at a cost that a researcher may be unable account for with a limited budget. One important point to mention is that a researcher designing a longitudinal study must take at least three measurements to obtain a reliable estimate of change and, perhaps more importantly, to allow a nonlinear pattern of change to be modelled (Ployhart & Vandenberg, 2010). In my dissertation, I hope to determine whether an optimal number of measurements exists when modelling a nonlinear pattern of change.

1.3.2 Spacing of Measurements

Additionally, a researcher must decide on the spacing of measurements in a longitudinal study. Although discussions of measurement spacing often recommend that researchers use theory and previous studies to implement measurement spacings that Dormann & Griffin (2015), organizational theories seldom delineate a period of time over which a process unfolds, and so the majority of longitudinal research uses intervals of convention and/or convenience to space measurements (Dormann & Ven, 2014; T. R. Mitchell & James, 2001). Unfortunately, using measurement spacing lengths that do not account for the temporal pattern of change of a psychological process can lead to inaccurate results (e.g., Chen et al., 2014). As an example, Cole & Maxwell (2009) provide show how correlation magnitudes are affected by the choice of measurement spacing intervals. In my dissertation, I hope to determine whether an optimal measurement spacing schedule exists when modelling a nonlinear pattern of change.

1.3.3 Time Structuredness

Last, and perhaps most pernicious, analyses of longitudinal data are likely to incur error from an assumption they make about data collection conditions. Many analyses assume that, across all collection points, participants provide their data at the same time. Unfortunately, such a high level of regularity in the response patterns of participants is unlikely: Participants are more likely to provide their data over some period of time after a data collection window has opened. As an example, consider a study that collects data from participants at the beginning of each month. If participants respond with perfect regularity, then they would all provide their data at the exact same time (e.g., noon on the second day of each month). If the participants respond with imperfect regularity, then they would provide their at different times after the beginning of each month. The regularity of responding observed across participants in a longitudinal study determines the time structuredness of the data and the sections that follow will provide overview of time structuredness.

1.3.3.1 Time-Structured Data

Many analyses assume that data are *time structured*: Participants provide data at the same time at each collection point. By assuming time-structured data, an analysis can incur error because it will map time intervals of inappropriate lengths onto the time intervals that occurred between participant's responses.[4] As an example of the conse-

[4] It should be noted that, although seldom implemented, analyses can be accessorized to handle time-unstructured data by using definition variables (Mehta & Neale, 2005; Mehta & West, 2000).

quences of incorrectly assuming data to be time structured, consider a study that assessed the effects of an intervention on the development of leadership by collecting leadership ratings at four time points each separated by four weeks (Day & Sin, 2011). The employed analysis assumed time-structured data; that is, each each participant provided ratings on the same day—more specifically, the exact same moment—each time these ratings were collected. Unfortunately, it is unlikely that the data collected from participants were time structured: At any given collection point, some participants may have provided leadership ratings at the beginning of the week, while others may only provide ratings two weeks after the survey opened. Importantly, ratings provided two weeks after the survey opened were likely influenced by changes in leadership that occurred over the two weeks. If an analysis incorrectly assumes time-structured data, then it assumes each participant has the same response rate and, therefore, will incorrectly attribute the amount of time that elapses between most participants' responses. For instance, if a participant only provides a leadership rating two weeks after having received a survey (and six weeks after providing their previous rating), then using an analysis that assumes time-structured data would incorrectly assume that each collection point of this participant is separated by four weeks (the interval used in the experiment) and would, consequently, model the observed change as if it had occurred over four weeks. Therefore, incorrectly assuming data to be time structured leads an analysis to overlook the unique response rates of participants across the collection points and, as a consequence, incur error (Coulombe et al., 2016; Mehta & Neale, 2005; Mehta & West, 2000).

1.3.3.2 Time-Unstructured Data

Conversely, some analyses assume that data are *time unstructured*: Participants provide data at different times at each collection point. Given the unlikelihood of one response pattern describing the response rates of all participants in a given study, the data obtained in a study are unlikely to be time structured. Instead, and because participants are likely to exhibit unique response patterns in their response rates, data are likely to be time unstructured. One way to conceptualize the distinction between time-structured and time-unstructured data is on a continuum. On one end of the continuum, participants all provide data with identical response patterns, thus giving time-structured data. When participants show unique response patterns, the resulting data are time unstructured, with the extent of time-unstructuredness depending on the length of the response windows. For example, if data are collected at the beginning of each month and participants only have one day to provide data at each time, then, assuming a unique response rate for each participant, the resulting data will have a low amount of time unstructuredness. Alternatively, if data are collected at the beginning of each month and participants have 30 days to provide data each time, then, assuming a unique response rate for each participant, the resulting data will have a high amount of time unstructuredness. Therefore, the continuum of time struturedness has time-structured data on one end and time-unstructured data with long response rates on another end. In my dissertation, I hope to determine how much error is incurred when time-unstructured data are assumed to be time structured.

1.3.4 Summary

In summary, researchers must contend with several issues when conducting longitudinal research. In addition to contending with issues encountered in conducting cross-sectional research, researchers must contend with new issues that arise from conducting longitudinal research. Three issues of particular importance in my dissertation are the number of measurements, the spacing of measurements, and incorrectly assuming data to be time structured. These issues will be serve as a basis for a systematic review of the simulation literature.

1.4 Using Simulations To Assess Modelling Accuracy

In the next section, I will present the results of the systematic review of the literature that has investigated the issues of measurement number, measurement spacing, and time structuredness. Before presenting the results of the systematic review, I will provide an overview of the Monte Carlo method used to investigate issues involved in conducting longitudinal research.

To understand how the effects of longitudinal issues on modelling accuracy can be investigated, the inferential method commonly employed in psychological research will first be reviewed with an emphasis on its shortcomings (see Figure 1.1). Consider an example where a researcher wants to estimate a population mean (μ) and understand how sampling error affects the accuracy of the estimate. Using the inferential method, the researcher samples data and then estimates the population mean (μ) by computing the mean of the sampled data. Because collected samples are almost always contaminated by a variety of methodological and/or statistical deficiencies (such as sampling error, measurement error, assumption violations, etc.), the estimation of the population parameter is likely to be imperfect. Unfortunately, to estimate the effect of sampling error on the accuracy of the population mean estimate (μ), the researcher would need to know the value of the population mean; without knowing the value of the population mean, it is impossible to know how much error was incurred in estimating the population mean and, as as a result, impossible to know the extent to which sampling error contributed to this error. Therefore, a study following the inferential approach can only provide estimates of population parameters.

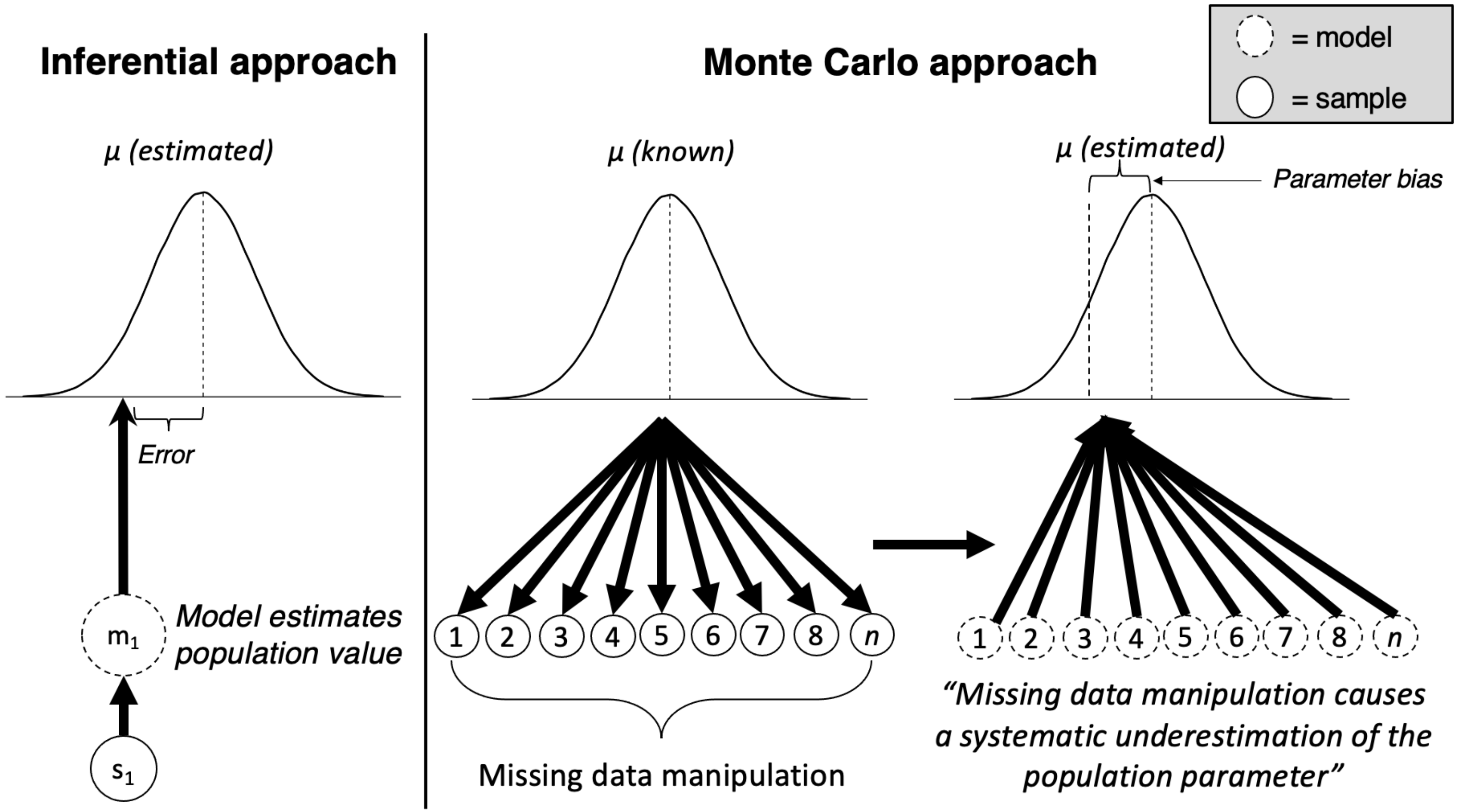
The Monte Carlo method has a different goal. Whereas inferential methods focus on estimating parameters from sample data, the Monte Carlo method is used to understand the factors that influence the accuracy of the inferential approach. Figure 1.1 shows that the Monte Carlo method works in the opposite direction of the inferential approach: Instead of collecting a sample, the Monte Carlo method begins by assigning a value to at least one parameter to define a population. Many sample data sets are then generated from the defined population, with some methodological deficiency built in to each data set. In the current example, each data set is generated to have a specific amount of missing data. Each generated sample is then analyzed and the population estimates of each statistical model are averaged and compared to the pre-determined parameter value.[5] The difference between the average of the estimates and the known population value constitutes bias in parameter estimation (i.e., parameter bias). In the current example, the missing data manipulation causes a systematic underestimation, on average, of the population parameter. By randomly generating data, the Monte Carlo method can determine how a variety of methodological and statistical factors affect the accuracy of a

[5]A statistical deficiency can also be introduced in the analysis of each generated data set.

model (for a review, see Robert & Casella, 2010).

Figure 1.1

Depiction of Monte Carlo Method



Note. Comparison of inferential approach with the Monte Carlo approach. The inferential approach begins with a collected sample and then estimates the population parameter using an appropriate statistical model. The difference between the estimated and population value can be conceptualized as error. Because the population value is generally unknown in the inferential approach, it cannot estimate how much error is introduced by any given methodological or statistical deficiency. To estimate how much error is introduced by any given methodological or statistical deficiency, the Monte Carlo method needs to be used, which constitutes four steps. The Monte Carlo method first defines a population by setting parameter values. Second, many samples are generated from the pre-defined population, with some methodological deficiency built in to each data set (in this case, each sample has a specific amount of missing data). Third, each generated sample is then analyzed and the population estimates of each statistical model are averaged and compared to the pre-determined parameter value. Fourth, the difference between the estimate average and the known population value defines the extent to which the missing data manipulation affected parameter estimation (the difference between the population and average estimated population value is the parameter bias).

Monte Carlo simulations have been used to evaluate a variety of methodological and statistical deficiencies. Beginning with the simple bivariate correlation, Monte Carlo simulations have shown that realistic values of sample size and measurement accuracy produce considerable variability in estimated correlation values (Stanley & Spence, 2014). Monte Carlo simulations have also provided valuable insights into more complicated statistical analyses. In investigating more complex statistical analyses, simulations have shown that mediation analyses are biased to produce results of complete mediation because the statistical power to detect direct effects falls well below the statistical power to detect indirect effects (Kenny & Judd, 2014). Finally, as an example of the utility of Monte Carlo simulations for evaluating growth mixture models, Monte Carlo simulations have shown that class enumeration accuracy (the ability to identify the correct number of response groups) decreases with nonnormal data (Bauer, 2003). Given the ability of the Monte Carlo method to evaluate statistical methods, the experiments in my dissertation used it to evaluate the effects of measurement number, measurement spacing, and time structuredness on modelling accuracy.[6]

1.5 Systematic Review of Simulation Literature

To understand the extent to which issues involved in conducting longitudinal research had been investigated, I conducted a systematic review of the simulation literature. The sections that follow will first present the method I followed in systematically reviewing the literature and then summarize the findings of the review.

[6]My simulation experiments also investigated the effects of sample size and nature of change on modelling accuracy.

1.5.1 Systematic Review Methodology

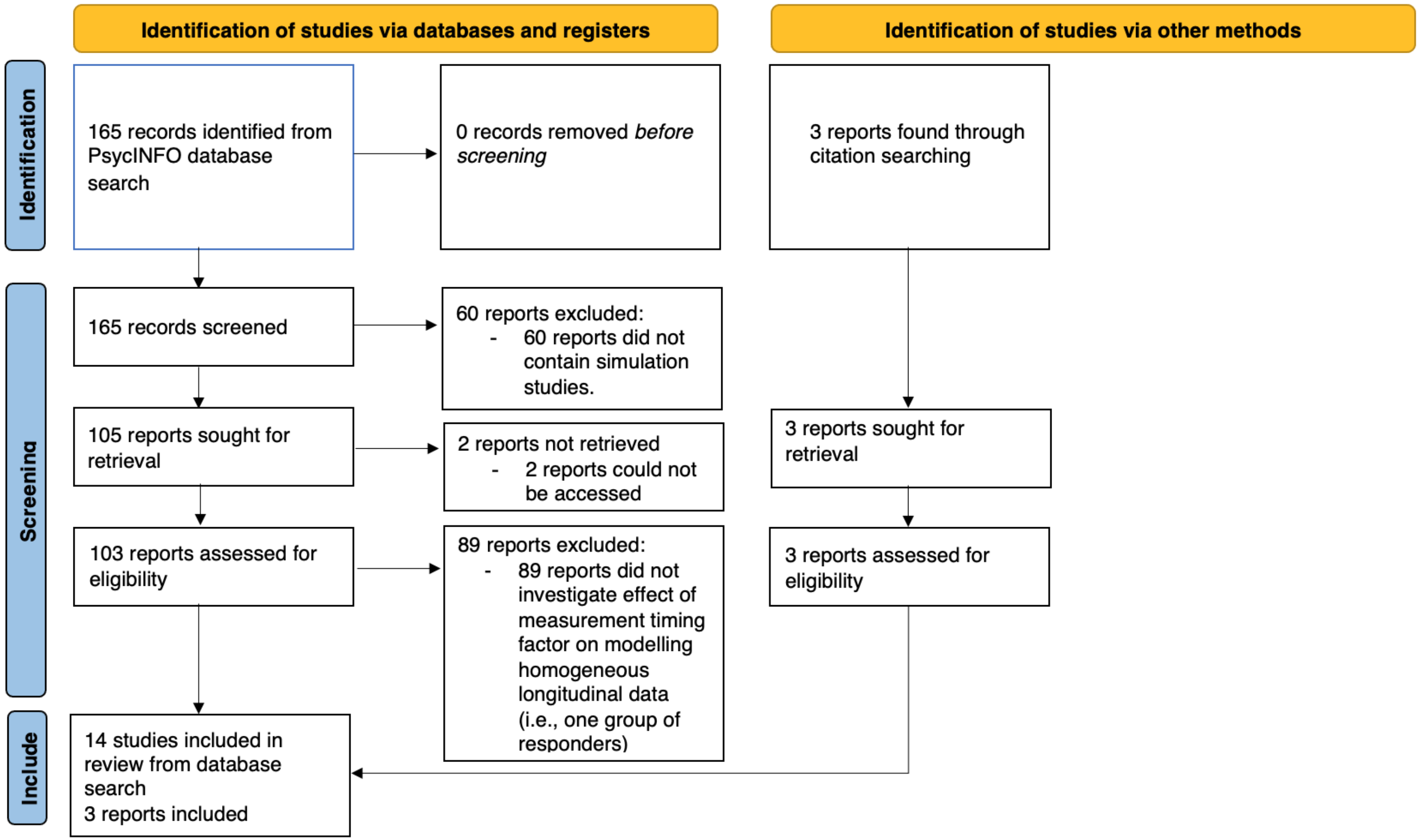
I identified the following keywords through citation searching and independent reading: "growth curve," "time-structured analysis," "time structure," "temporal design," "individual measurement occasions," "measurement intervals," "methods of timing," "longitudinal data analysis," "individually-varying time points," "measurement timing," "latent difference score models," "parameter bias," and "measurement spacing." I entered these keywords entered into the PsycINFO database (on July 23, 2021) and any paper that contained any one of these key words and the word "simulation" in any field was considered a viable paper (see Figure 1.2 for a PRISMA diagram illustrating the filtering of the reports). The search returned 165 reports, which I screened by reading the abstracts. Initial screening led to the removal of 60 reports because they did not contain any simulation experiments. Of the remaining 105 papers, I removed 2 more popers because they could not accessed (Stockdale, 2007; Tiberio, 2008). Of the remaining 103 identified simulation studies, I deemed a paper as relevant if it investigated the effects of any design and/or analysis factor relating to conducting longitudinal research (i.e., number of measurements, spacing of measurements, and/or time structuredness) and did so using the Monte Carlo simulation method. Of the remaining 103 studies, I removed 89 studies being removed because they did not meet the inclusion criteria, leaving fourteen studies to be included the review, with. I also found an additional 3 studies through citation searching, giving a total of 17 studies.

The findings of my systematic review are summarized in Tables 1.1–1.2. Tables 1.1–1.2 differ in one way: Table 1.1 indicates how many studies investigated each effect, whereas Table 1.2 provides the reference of each study and detailed information about each study's method. Otherwise, all other details of Tables 1.1–1.2 are identical. The first column lists the longitudinal design factor (alongside with sample size) and the corresponding two- and three-way interactions. The second and third columns list whether each effect has been investigated with linear and nonlinear patterns of change, respectively. Shaded cells indicate effects that have not been investigated, with cells shaded in light blue indicating effects that have not been investigated with linear patterns of change and cells shaded in dark blue indicating effects that have not been investigated with nonlinear patterns of change.[7]

[7] Table 1.2 lists the effects that each study (identified by my systematic review) investigated and notes the following methodological details (using superscript letters and symbols): the type of model used in each paper, assumption and/or manipulation of complex error structures (heterogeneous variances and/or correlated residuals), manipulation of missing data, and/or pseudo-time structuredness manipulation. Across all 17 simulation studies, 5 studies (29%) assumed complex error structures (Gasimova et al., 2014; Liu & Perera, 2021; Y. Liu et al., 2015; Miller & Ferrer, 2017; Murphy et al., 2011), 1 study (6%) manipulated missing data (Fine et al., 2019), and 2 studies (12%) contained a pseudo-time structuredness manipulation (Fine et al., 2019; Fine & Grimm, 2020) . Importantly, the pseudo-time structuredness manipulation used in Fine et al. (2019) and Fine & Grimm (2020) differed from the manipulation of time structuredness used in the current experiments (and from previous simulation experiments of Coulombe et al., 2016; Miller & Ferrer, 2017) in that it randomly generated longitudinal data such that a given person could provide all their data before another person provided any data.

Figure 1.2

PRISMA Diagram Showing Study Filtering Strategy



Note. PRISMA diagram for systematic review of simulation research that investigates measurement timing effects.

1.5.2 Systematic Review Results

Although the previous research appeared to sufficiently fill some cells of Table 1.1, two patterns suggest that arguably the most important cells (or effects) have not been investigated. First, it appears that simulation research has invested more effort in investigating the effects of longitudinal design factors with linear patterns than with nonlinear patterns of change. In counting the number of effects that remain unaddressed with linear and nonlinear patterns of change, a total of five cells (or effects) have not been

Table 1.1

Number of Simulation Studies That Have Investigated Longitudinal Issues with Linear and Nonlinear Change Patterns (n = 17)

Effect	Linear pattern	Nonlinear pattern
Main effects		
Number of measurements (NM)	11 studies	6 studies
Spacing of measurements (SM)	1 study	1 study
Time structuredness (TS)	2 studies	1 study
Sample size (S)	11 studies	7 studies
Two-way interactions		
NM x SM	1 study	1 study
NM x TS	1 study	**Cell 1 (Exp. 3)**
NM x S	9 studies	5 studies
SM x TS	**Cell 2**	**Cell 3**
SM x S	**Cell 4**	**Cell 5 (Exp. 2)**
TS x S	1 study	2 studies
Three-way interactions		
NM x SM x TS	**Cell 6**	**Cell 7**
NM x SM x S	**Cell 8**	**Cell 9 (Exp. 2)**
NM x TS x S	1 study	**Cell 10 (Exp. 3)**
SM x TS x S	**Cell 11**	**Cell 12**

Table 1.1

Number of Simulation Studies That Have Investigated Longitudinal Issues with Linear and Nonlinear Change Patterns (n = 17) (continued)

Effect	Linear pattern	Nonlinear pattern

Note. Cells are only numbered for effects that have not been investigated. Cells shaded in light blue indicate effects that have not been investigated with linear patterns of change and cells shaded in dark blue indicate effects that have not been investigated with nonlinear patterns of change.

Table 1.2

Summary of Simulation Studies That Have Investigated Longitudinal Issues with Linear and Nonlinear Change Patterns (n = 17)

Effect	Linear pattern	Nonlinear pattern
Main effects		
Number of measurements (NM)	Timmons & Preacher (2015)[a]; Murphy et al. (2011)[℧b]; Gasimova et al. (2014)[c℧]; Wu et al. (2014)[a]; Coulombe (2016)[a];Ye (2016)[a]; Finch (2017)[a]; O'Rourke et al. (2021)[d]; Newsom & Smith (2020)[a]; Coulombe et al. (2016)[a]	Timmons & Preacher (2015)[a]; Finch (2017)[a]; Fine et al. (2019)[e∘▽]; Fine & Grimm (2020)[e,f▽];J. Liu et al. (2019)[g]; Liu & Perera (2021)[h℧]; Y. Liu et al. (2015)[g℧]
Spacing of measurements (SM)	Timmons & Preacher (2015)[a]	Timmons & Preacher (2015)[a]
Time structuredness (TS)	Aydin et al. (2014)[a]; Coulombe et al. (2016)[a]	Miller & Ferrer (2017)[a℧]; Y. Liu et al. (2015)[g℧]
Sample size (S)	Murphy et al. (2011)[b℧]; Gasimova et al. (2014)[c℧]; Wu et al. (2014)[a]; Coulombe (2016)[a];Ye (2016)[a]; Finch (2017)[a]; O'Rourke et al. (2021)[d]; Newsom & Smith (2020)[a]; Coulombe et al. (2016)[a];Aydin et al. (2014)[a]; Coulombe et al. (2016)[a]	Finch (2017)[a]; Fine et al. (2019)[e∘▽]; Fine & Grimm (2020)[e,f▽];J. Liu et al. (2019)[g]; Liu & Perera (2021)[h℧]; Y. Liu et al. (2015)[g℧]; Miller & Ferrer (2017)[a℧]
Two-way interactions		
NM x SM	Timmons & Preacher (2015)[a]	Timmons & Preacher (2015)[a]
NM x TS	Coulombe et al. (2016)[a]	**Cell 1**

Table 1.2

Summary of Simulation Studies That Have Investigated Longitudinal Issues with Linear and Nonlinear Change Patterns (n = 17) (continued)

Effect	Linear pattern	Nonlinear pattern
NM x S	Murphy et al. (2011)[b℧]; Gasimova et al. (2014)[c℧]; Wu et al. (2014)[a]; Coulombe (2016)[a];Ye (2016)[a]; Finch (2017)[a]; O'Rourke et al. (2021)[d]; Newsom & Smith (2020)[a]; Coulombe et al. (2016)[a]	Finch (2017)[a]; Fine et al. (2019)[e∘▽]; Fine & Grimm (2020)[e,f▽];J. Liu et al. (2019)[g]; Liu & Perera (2021)[h℧]
SM x TS	**Cell 2**	**Cell 3**
SM x S	**Cell 4**	**Cell 5**
TS x S	Aydin et al. (2014)[a]	Y. Liu et al. (2015)[g℧]; Miller & Ferrer (2017)[a℧]
Three-way interactions		
NM x SM x TS	**Cell 6**	**Cell 7**
NM x SM x S	**Cell 8**	**Cell 9**
NM x TS x S	Coulombe et al. (2016)[a]	**Cell 10**
SM x TS x S	**Cell 11**	**Cell 12**

Note. Cells are only numbered for effects that have not been investigated. Cells shaded in light and dark blue indicate effects that have not, respectively, been investigated with linear and nonlinear patterns of change.

[a] Latent growth curve model. [b] Second-order latent growth curve model. [c] Hierarchical Bayesian model. [d] Bivariate latent change score model. [e] Functional mixed-effects model. [f] Nonlinear mixed-effects model. [g] Bilinear spline model. [g] Parallel bilinear spline model.

[∘] Manipulated missing data. [℧] Assumed complex error structure (heterogeneous variances and/or correlated residuals). [▽] Contained pseudo-time structuredness manipulation.

1.5.3 Next Steps

Given that longitudinal research is needed to understand the temporal dynamics of psychological processes, it is necessary to understand how longitudinal design and analysis factors interact with each other (and with sample size) in affecting the accuracy with which nonlinear patterns of change are modelled. With no study to my knowledge having conducted a comprehensive investigation of how longitudinal design and analysis factors affect the modelling of nonlinear change patterns, my simulation experiments are designed to address this gap in the literature. Specifically, my simulation experiments investigate how measurement number, measurement spacing, and time structuredness affect the accuracy with which a nonlinear change pattern is modelled (see Cells 1, 5, 9, and 10 of Table 1.1).

1.6 Methods of Modelling Nonlinear Patterns of Change Over Time

Because my simulation experiments assumed change over time to be nonlinear, it is important to provide an overview of how nonlinear change is modelled. In this section, I will provide a brief review on how nonlinear change can be modelled and will contrast the commonly employed polynomial approach with the lesser known nonlinear function approach that I use in my simulations.[8][9]

[8] It should be noted that nonlinear change can be modelled in a variety of ways, with latent change score models (e.g., O'Rourke et al., 2021) and spline models (e.g., Fine & Grimm, 2020) offering some examples.

[9] The definition of a nonlinear function is mathematical in nature. Specifically, a nonlinear function contains at least one parameter that exists in the corresponding partial derivative. For example, in the logistic function $\theta + \frac{\alpha-\theta}{1+exp(\frac{\beta-t}{\gamma})}$ is nonlinear because β exists in $\frac{\partial y}{\partial \beta}$ (in addition to γ existing in its corresponding partial derivative). The n^{th} order polynomial function of $y = a + bx + cx^2 + ... + nx^n$ is linear because the partial derivatives with respect to the parameters (i.e., $1, x^2, ..., x^n$) do not contain the associated parameter.

Consider an example where an organization introduces a new incentive system with the goal of increasing the motivation of its employees. To assess the effectiveness of the incentive system, employees provide motivation ratings every month days over a period of 360 days. Over the 360-day period, the motivation levels of the employees increase following an s-shaped pattern of change over time. One analyst decides to model the observed change using a **polynomial function** shown below in Equation 1.1:

$$y = a + bx + cx^2 + dx^3. \tag{1.1}$$

A second analyst decides to model the observed change using a **logistic function** shown below in Equation 1.2:

$$y = \theta + \frac{\alpha - \theta}{1 + e^{\frac{\beta - time}{\gamma}}} \tag{1.2}$$

Figure 1.3A shows the response pattern predicted by the polynomial function of Equation 1.1 with the estimated values of each parameter (a, b, c, and d) and Figure 1.3B shows the response pattern predicted by the logistic function (Equation 1.2) along with the values estimated for each parameter (θ, α, β, and γ). Although the logistic and polynomial functions predict nearly identical response patterns, the parameters of the logistic function have the following meaningful interpretations (see Figure 1.4):

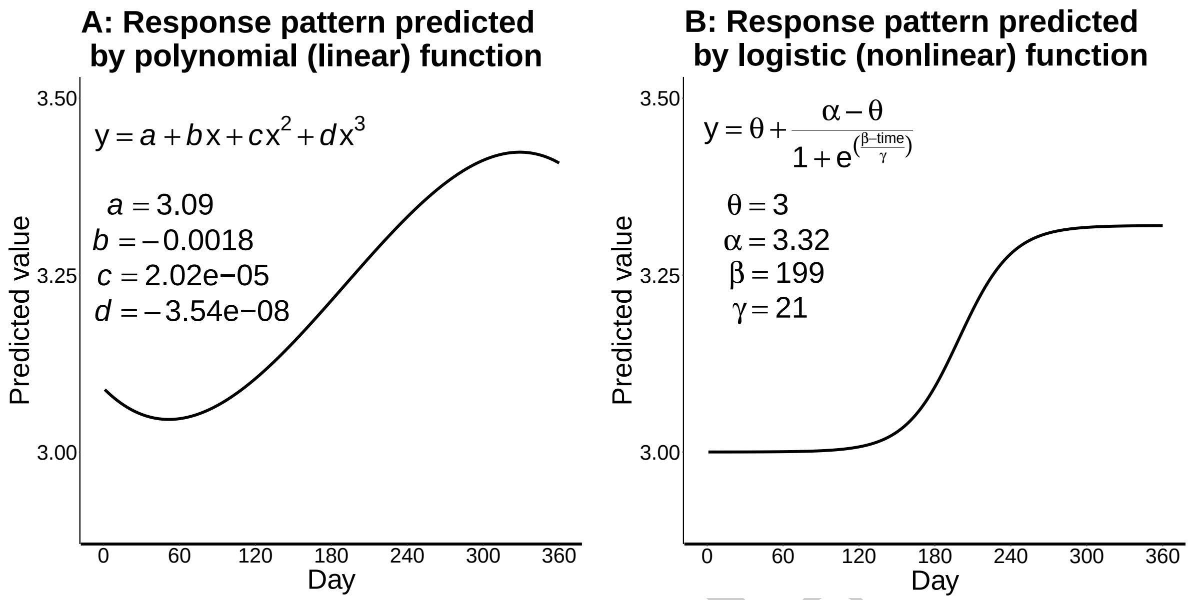
- θ specifies the value at the first plateau (i.e., the starting value) and so is called the *baseline* parameter (see Figure 1.4A).
- α specifies the value at the second plateau (i.e., the ending value) and so is called the the *maximal elevation* parameter (see Figure 1.4B).

- β specifies the number of days required to reach the half the difference between the first and second plateau (i.e., the midway point) and so is called the *days-to-halfway-elevation* parameter (see Figure 1.4C).
- γ specifies the number of days needed to move from the midway point to approximately 73% of the difference between the starting and ending values (i.e., satiation point) nd so is called the *halfway-triquarter delta* parameter (see Figure 1.4D).

Applying the parameter meanings of the logistic function to the parameter values estimated by using the logistic function (Equation 1.2), the predicted response pattern begins at a value of 3 (baseline) and reaches a value of 3.32 (maximal elevation) by the end of the 360-day period. The midway point of the curve is reached after 199 days (days-to-halfway elevation) and the satiation point is reached 21days later (halfway-triquarter delta; or 220 days after the beginning of the incentive system is introduced). When looking at the polynomial function, aside from the '*a*' parameter indicating the starting value, it is impossible to meaningfully interpret the values of any of the other parameter values. Therefore, using a nonlinear function such as the logistic function provides a meaningful way to interpret nonlinear change.

Figure 1.3

Response Patterns Predicted by Polynomial (Equation 1.1) and Logistic (Equation 1.2) Functions



Note. Panel A: Response pattern predicted by the polynomial function of Equation (1.1). Panel B: Response pattern predicted by the logistic function of Equation (1.2).

1.7 Overview of Simulation Experiments

To investigate the effects of longitudinal design and analysis factors on modelling accuracy, I conducted three Monte Carlo experiments. Before summarizing the simulation experiments, one point needs to be mentioned regarding the maximum number of independent variables used in each experiment. No simulation experiment manipulated more than three variables because of the difficulty associated with interpreting interactions between four or more variables. Even among academics, the ability to correctly interpret interactions sharply declines when the number of independent variables increases from three to four (Halford et al., 2005). Therefore, none of my simulation experiments manipulated more than three variables so that results could be readily interpreted.

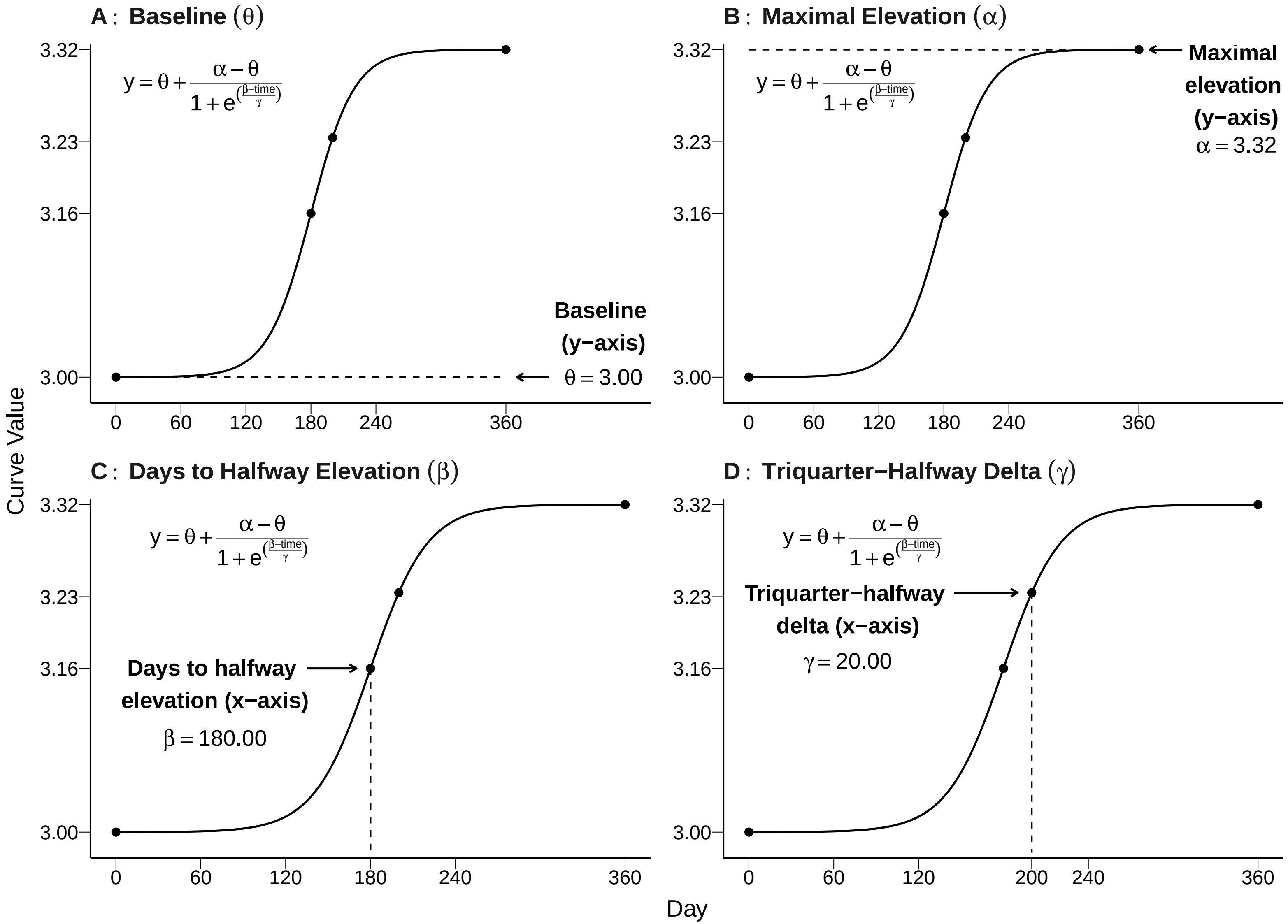
To summarize the three simulation experiments, the independent variables of each simulation experiment are listed below:

- Experiment 1: number of measurements, spacing of measurements, and nature of change.
- Experiment 2: number of measurements, spacing of measurements, and sample size.
- Experiment 3: number of measurements, sample size, and time structuredness.

The sections that follow will present each of the simulation experiments and their corresponding results.

Figure 1.4

Description Each Parameters Logistic Function (Equation 1.2) Functions



Note. Panel A: The baseline parameter (θ) sets the starting value of the of curve, which in the current example has a value of 3.00 (θ = 3.00). Panel B: The maximal elevation parameter (α) sets the ending value of the curve, which in the current example has a value of 3.32 (α = 3.32). Panel C: The days-to-halfway elevation parameter (β) sets the number of days needed to reach 50% of the difference between the baseline and maximal elevation. In the current example, the baseline-maximal elevation difference is 0.32 ($\alpha - \theta$ = 3.32 - 3.00 = 0.32), and so the days-to-halfway elevation parameter defines the number of days needed to reach a value of 3.16. Given that the days-to-halfway elevation parameter is set to 180 in the current example ($\beta = 180$), then 180 days are neededto go from a value of 3.00 to a value of 3.16. Panel D: The halfway-triquarter delta parameter (γ) sets the number of days needed to go from halfway elevation to approximately 73% of the baseline-maximal elevation difference of 0.32 ($\alpha - \theta$ = 3.32 - 3.00 = 0.32). Given that 73% of the baseline-maximal elevation difference is 0.23 and the halfway-triquarter delta is set to 40 days ($\gamma = 40$), then 40 days are needed to go from the halfway point of 3.16 to the triquarter point of approximately 3.23).

2 Experiment 1

In Experiment 1, I investigated the number of measurements needed to obtain accurate modelling (i.e., unbiased and precise parameter estimation) under different spacing schedules and natures of change. Before presenting the results of Experiment 1, I will present my design and analysis goals. For the design, I conducted a 4(measurement spacing:equal, time-interval increasing, time-interval decreasing, middle-and-extreme x 4(number of measurements: 5, 7, 9, 11) x 3(nature of change: population value for the fixed-effect days-to-halfway elevation parameter [β_{fixed}] of 80, 180, or 280) study. For the analysis, I was interested in answering two questions. First, I was interested in whether spacing measurements near periods of change leads to higher modelling accuracy. To answer the first question, I determined whether modelling accuracy under each spacing schedule increased when measurements were taken closer to periods of change.

Second, I was interested in how to space measurements when the nature of change is unknown. When the nature of change is unknown, this translates to a situation where a researcher has little to no knowledge of how change unfolds over time, and so any nature of change is a viable candidate for the true change. Therefore, to determine how to space measurements when the nature of change is unknown, I averaged the modelling accuracy of each spacing schedule across all possible nature-of-change curves and considered the spacing schedule with the highest modelling accuracy to be the best one.

2.1 Methods

2.1.1 Variables Used in Simulation Experiment

2.1.1.1 Independent Variables

To build on current research, Experiment 1 used independent variable manipulations from a select number of previous studies. In looking at the summary of the simulation literature in Table 1.2, the study by Coulombe et al. (2016) was the only one to investigate three longitudinal issues of interest to my dissertation, and so represented the most comprehensive investigation. Because I was also interested in investigating measurement spacing, manipulations were inspired from the only other simulation study to manipulate measurement spacing (the study by Timmons & Preacher, 2015). The sections that follow will discuss each of the variables manipulated in Experiment 1.

2.1.1.1.1 Spacing of Measurements

The only simulation study identified by my systematic review that manipulated measurement spacing was Timmons & Preacher (2015). Measurement spacing in Timmons & Preacher (2015) was manipulated in the following four ways:

1) **Equal spacing**: measurements were divided by intervals of equivalent lengths.
2) **Time-interval increasing spacing**: intervals that divided measurements increased in length over time.
3) **Time-interval decreasing spacing**: intervals that divided measurements decreased in length over time.
4) **Middle-and-extreme spacing**: measurements were clustered near the beginning, middle, and end of the data collection period.

To maintain consistency with the established literature, I manipulated measurement spacing in the same way as Timmons & Preacher (2015) presented above. Importantly, because Timmons & Preacher (2015) did not create their measurement spacing schedules with any systematicity, I developed a novel and replicable procedure for generating measurement schedules for each of the four measurement spacing conditions, which is described in Appendix A. I also automated the generation of measurement schedules by creating a set of functions in R (RStudio Team, 2020).

Table 2.1 lists the measurement days that were used for all measurement spacing-measurement number cells. The first column lists the type of measurement spacing (i.e., equal, time-interval increasing, time-interval decreasing, or middle-and-extreme); the second column lists the number of measurements (5, 7, 9, or 11); the third column lists the measurement days that correspond to each measurement number-measurement spacing condition; and the fourth column lists the interval lengths that characterize each set of measurements. Note that the interval lengths are equal for the equal spacing, increase over time for the time-interval increasing spacing, and decrease over time for the time-interval decreasing spacing, For cells with middle-and-extreme spacing, the measurement days and and interval lengths corresponding to the middle of the measurement window have been emboldened.

Table 2.1

Measurement Days Used for All Measurement Number-Measurement Spacing Conditions

Spacing Schedule	Number of Measurements	Measurement Days	Interval Lengths
Equal	5	0, 90, 180, 270, 360	90, 90, 90, 90
	7	0, 60, 120, 180, 240, 300, 360	60, 60, 60, 60, 60, 60
	9	0, 45, 90, 135, 180, 225, 270, 315, 360	45, 45, 45, 45, 45, 45, 45, 45
	11	0, 36, 72, 108, 144, 180, 216, 252, 288, 324, 360	36, 36, 36, 36, 36, 36, 36, 36, 36, 36
Time-interval increasing	5	0, 30, 100, 210, 360	30, 70, 110, 150
	7	0, 30, 72, 126, 192, 270, 360	30, 42, 54, 66, 78, 90
	9	0, 30, 64.29, 102.86, 145.71, 192.86, 244.29, 300, 360	30, 34.29, 38.57, 42.86, 47.14, 51.43, 55.71, 60
	11	0, 30, 61.33, 94, 128, 163.33, 200, 238, 277.33, 318, 360	30, 31.33, 32.67, 34, 35.33, 36.67, 38, 39.33, 40.67, 42
Time-interval decreasing	5	0, 150, 260, 330, 360	150, 110, 70, 30
	7	0, 90, 168, 234, 288, 330, 360	90, 78, 66, 54, 42, 30
	9	0, 60, 115.71, 167.14, 214.29, 257.14, 295.71, 330, 360	60, 55.71, 51.43, 47.14, 42.86, 38.57, 34.29, 30
	11	0, 42, 82.67, 122, 160, 196.67, 232, 266, 298.67, 330, 360	42, 40.67, 39.33, 38, 36.67, 35.33, 34, 32.67, 31.33, 30
Middle-and-extreme	5	1, **150, 180, 210**, 360	150, **30, 30**, 150

Table 2.1

Measurement Days Used for All Measurement Number-Measurement Spacing Conditions (continued)

Spacing Schedule	Number of Measurements	Measurement Days	Interval Lengths
	7	1, 30, **150, 180, 210**, 330, 360	30, 120, **30, 30**, 120, 30
	9	1, 30, 60, **150, 180, 210**, 300, 330, 360	30, 30, 90, **30, 30**, 90, 30, 30
	11	1, 30, 60, **120, 150, 180, 210, 240,** 300, 330, 360	30, 30, 60, **30, 30, 30, 30**, 60, 30, 30

Note. For middle-and-extreme spacing levels, the measurement days and and interval lengths corresponding to the middle of measurement windows have been emboldened.

2.1.1.1.2 Number of Measurements

The smallest measurement number value in (**coulomble2016?**)(i.e., three) could not be used in Experiment 1 (or any other simulation experiment that manipulated measurement number in my dissertation) because doing so would have created non-identified models. The model used in my simulations estimated 9 parameters ($p = 9$; 4 fixed-effects + 4 random-effects + 1 error) and so the minimum number of measurements (or observed variables) required for model identification (and to allow model comparison) was 4.[10]. Although a measurement number of three could not be used in my manipulation of measurement number, the next highest measurement number values in Coulombe et al. (2016) of 5, 7, and 9 were used. Importantly, a larger value of 11 was added to test for a possible effect of a high measurement number. Therefore, my simulation experiments used the following values in manipulating the number of measurements: 5, 7, 9, and 11.

2.1.1.1.3 Population Values Set for The Fixed-Effect Days-to-Halfway Elevation Parameter β_{fixed} (Nature of Change)

The nature of change was manipulated by setting the days-to-halfway elevation parameter (β_{fixed}) to a value of either 80, 180, or 280 days (see Figure 1.4A). Note that no other study manipulated nature of change using logistic curves and so its manipulation in Experiment 1 is, to the best of my knowledge, unique (in this literature). Nature of change was manipulated to simulate situations where uncertainty exists in the nature of change.

[10] Degrees of freedom is calculated by multiplying the number of observed variables (p) by $p + 1$ and dividing it by 2 (($p[\{p + 1\}]/2$; see Loehlin & Beaujean, 2017)

2.1.1.2 Constants

Because sample size was not manipulated in Experiment 1, I set it to have a constant value across all cells. I decided to set the sample size value to the average sample size used in organizational research ($n = 225$; Bosco et al., 2015). Another variable set to a constant value across the cells was time structuredness (data were assumed to be time structured). That is, data were generated such that, at each time point, all data were obtained at the exact same time.

2.1.1.3 Dependent Variables

2.1.1.3.1 Convergence Success Rate

The proportion of iterations in a cell where models converged defined the *convergence success rate*.[11] Equation (4.5) below shows the calculation used to compute the convergence success rate:

$$\text{Convergence success rate} = \frac{\text{Number of models that successfully converged in a cell}}{n}, \tag{2.1}$$

where n represents the total number of models run in a cell.

2.1.1.3.2 Bias

Bias was calculated to evaluate the accuracy with which each logistic function parameter was estimated in each experimental cell. As shown below in Equation (4.6), *bias* was obtained by summing the differences between the population value set for a parameter and the value estimated for the parameter by each i converged model and then

[11]Specifically, convergence was obtained if the convergence code returned by OpenMx was 0.

dividing the sum by the number of N converged models.

$$\text{Bias} = \frac{\sum_i^N (\text{Population value for parameter} - \text{Average estimated value}_i)}{N} \tag{2.2}$$

Bias was calculated for the fixed- and random-effect parameters of the baseline (θ_{fixed}, θ_{random}), maximal elevation (α_{fixed}, α_{random}), days-to-halfway elevation (β_{fixed}, β_{random}), and the halfway-triquarter delta parameters (γ_{fixed}, γ_{random}) and the error parameter (ϵ).

2.1.1.3.3 Precision

In addition to computing bias, precision was calculated to evaluate the variability with which each parameter was estimated. Importantly, metrics used to evaluate precision in previous studies could not be used for two reasons. First, some metrics assume estimates are normally distributed (e.g., mean-squared error and empirical standard error). Because some parameters in my simulations had skewed distributions, using a metric that assumed a normal distribution would likely yield inaccurate results. Second, although some simulation studies have used confidence intervals to evaluate precision, there is no confidence interval calculation (to my knowledge) for structure latent growth curves. Therefore, I defined *precision* as the range of values covered by the middle 95% of values estimated for a logistic parameter, which could be interpreted as a range of plausible population estimates.

2.1.2 Overview of Data Generation

2.1.2.1 Data Generation

2.1.2.1.1 Function Used to Generate Each Data Set

Data for each simulation experiment were generated using R (RStudio Team, 2020). To generate the data, the *multilevel logistic function* shown below in Equation (2.3) was used:

$$y_{ij} = \theta_j + \frac{\alpha_j - \theta_j}{1 + e^{\frac{\beta_j - time_i}{\gamma_j}}} + \epsilon_{ij}, \tag{2.3}$$

where θ represents the baseline parameter, α represents the maximal elevation parameter, β represents the days-to-halfway elevation parameter, and γ represents triquarter-halfway delta parameter. Note that, values for θ, α, β, and γ were generated for each j person across all i time points, with an error value being randomly generated at each i time point(ϵ_{ij}; see Figure 1.4 for a review of each parameter). In other words, unique response patterns were generated for each person in each of the 1000 data sets generated per cell.

The logistic growth function (Equation 2.3 was used because it is a common pattern of organizational change (or institutionalization; Lawrence et al., 2001). Institutionalization curves follow an s-shaped pattern of the logistic growth function, and so their rates of change can be represented by the days-to-halfway elevation and triquarter-halfway delta parameters (β, γ, respectively), and the success of the change can be defined by the magnitude of the difference between baseline and maximal elevation parameters (α - θ, respectively).

2.1.2.1.2 Population Values Used for Function Parameters

Table 2.2 lists the parameter values that will be used for the population parameters. Given that the decisions for setting the values for the baseline, maximal elevation, and residual variance parameters were informed by past research, the discussion that follows highlights how these decisions were made. The difference between the baseline and maximal elevation parameters (θ and α, respectively) corresponded to the effect size most commonly observed in organizational research (i.e., the 50$^{\text{th}}$ percentile effect size value; Bosco et al., 2015). Because the meta-analysis of Bosco et al. (2015) computed effect sizes as correlations, the the 50$^{\text{th}}$ percentile effect size value of $r = .16$ was computed to a standardized effect size using the following conversion function shown in Equation 2.4 (Borenstein et al., 2009, Chapter 7):

$$d = \frac{2r}{\sqrt{1 - r^2}}, \tag{2.4}$$

where r is the correlation effect size. Using Equation 2.4, a correlation value of $r = .16$ becomes a standardized effect size value of $d = 0.32$. For the value of the residual variance parameter, its value in Coulombe et al. (2016) was set to the value used for the value of the intercept variance parameter. In the current context, the intercept of the logistic function (Equation 2.3) is the baseline parameter.[12] Given that the value for the variability of the baseline parameter was 0.05 (albeit in standard deviation units), the value used for the residual variance parameter was 0.05 ($\epsilon = 0.05$). Because justification

[12]The definition of an intercept parameter is the value of a curve when no time has elapsed, and this is precisely the definition of the baseline parameter (θ). Therefore, the variance of the intercept parameter carries the same meaning as the variance of the baseline parameter (θ_{random}).

for the other parameters could not be found in any of the simulation studies identified in my systematic review, values set for the other parameters was largely arbitrary.

To facilitate interpretation of the results, data were generated to resemble the commonly used Likert (range of 1–5) by using a standard deviation of 1.00 and change was assumed to occur over a period of 360 days. The decision to generate data in the context of a 360-day period was made because many organizational processes are often governed by annual events (e.g., performance reviews, annual returns, regulations, etc.). Importantly, because Coulombe et al. (2016) set covariances between parameters to zero, all the simulation experiments used zero-value covariances.

Table 2.2
Values Used for Multilevel Logistic Function Parameters

Parameter Means	Value
Baseline, θ	3.00
Maximal elevation, α	3.32
Days-to-halfway elevation, β	180.00
Triquarter-halfway delta, γ	20.00
Variability and Covariability Parameters (in Standard Deviations)	
Baseline standard deviation, ψ_{θ}	0.05
Maximal elevation standard deviation, ψ_{α}	0.05
Days-to-halfway elevation standard deviation, ψ_{β}	10.00
Triquarter-halfway delta standard deviation, ψ_{γ}	4.00
Baseline-maximal elevation covariability, $\psi_{\theta\alpha}$	0.00
Baseline-days-to-halfway elevation covariability, $\psi_{\theta\beta}$	0.00
Baseline-triquarter-halfway delta covariability, $\psi_{\theta\gamma}$	0.00
Maximal elevation-days-to-halfway elevation covariability, $\psi_{\alpha\beta}$	0.00
Maximal elevation-triquarter-halfway delta covariability, $\psi_{\alpha\gamma}$	0.00
Days-to-halfway elevation-triquarter-halfway delta covariability, $\psi_{\beta\gamma}$	0.00

Residual standard deviation, ψ_ϵ	0.05

Note. The difference between α and θ corresponds to the 50$^{\text{th}}$ percentile Cohen's d value of 0.32 in organizational psychology (Bosco et al., 2015).

2.1.3 Modelling of Each Generated Data Set

Previously, I described how data were generated. Here, I describe how the generated data were modelled.

Each data set generated by the multilevel logistic function (Equation 2.3) was analyzed using a modified latent growth curve model known as a structure latent growth curve model (K. J. Preacher & Hancock, 2015). Importantly, the model fit to each generated data set estimated nine parameters: A fixed-effect parameter for each of the four logistic function parameters, a random-effect parameter for each of the four logistic function parameters, and an error parameter. As with a multilevel model, a fixed-effect parameter has a constant value across all individuals, whereas a random-effect parameter represents the variability of values across all modelled people.[13] To fit the logistic function to a given data set (Equation 2.3), a linear approximation of the logistic function was needed so that it could fit within the linear nature of structural equation modelling framework.[14] To construct a linear approximation of the logistic function, a first-order Taylor series was constructed for the logistic function. For a detailed explanation of how the logistic function was fit into the structural equation modelling framework, see

[13]Estimating a random-effect for a parameter allows person- or data-point-specific values to be computed for the parameter.

[14]The logistic function (Equation 2.3) is a nonlinear function and so cannot be directly inserted into the structural equation modelling framework because this framework only allows linear computations of matrix-matrix, matrix-vector, and vector-vector operations. Unfortunately, the algebraic operations permitted in a linear framework cannot directly reproduce the operations in the logistic function (Equation 2.3) and so a linear approximation of the logistic function must be constructed so that the logistic function can be inserted into the structural equation modelling framework.

Technical Appendix B.

2.1.4 Analysis of Data Modelling Output and Accompanying Visualizations

To analyse and visualize modelling performancel, I calculated values for convergence success rate, bias, and precision in each experimental cell (see dependent variables). The sections that follow provide details on how I analysed each dependent variable and constructed plots to visualize bias and precision.

2.1.4.1 Analysis of Convergence Success Rate

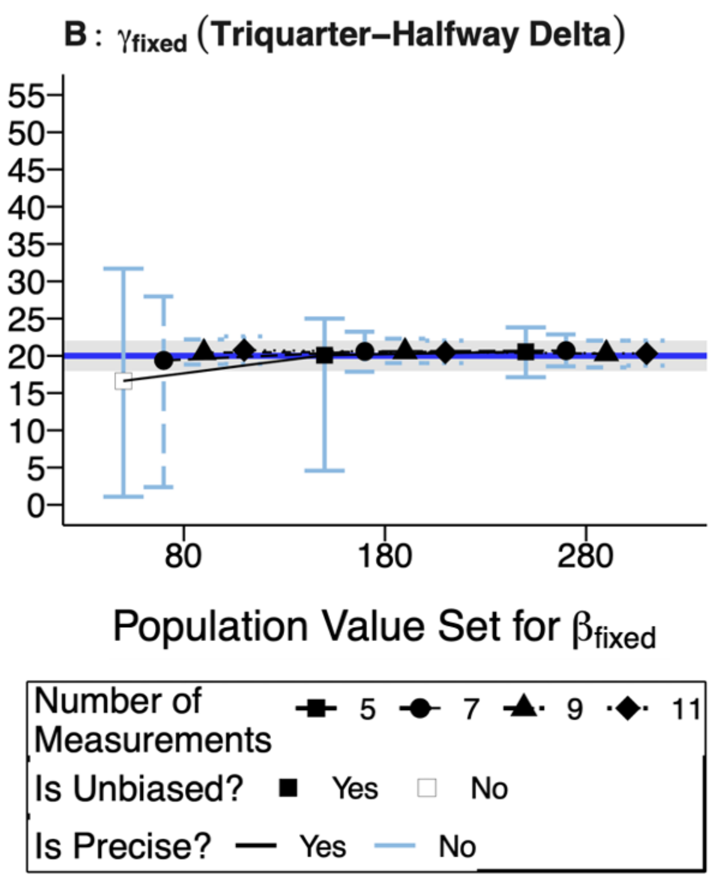
For the analysis of convergence success rate, the mean convergence success rate was computed for each cell in each experiment (see section on convergence success rate). Because convergence rates exhibited little variability across cells due to the nearly unanimous high rates (almost all cells across all experiments had convergence success rates above 90%), examining the effects of any independent variable on these rates would have provided little information. Therefore, I only reported the average convergence success rate for each cell (see Appendix B).

2.1.4.2 Analysis and Visualization of Bias

In accordance with several simulation studies, an estimate with a bias value within a ±10% margin of error of the parameter's population value was deemed unbiased (Muthén, 1997). To visualize bias, I constructed parameter estimation plots. Figure 2.1 shows a parameter estimation plot for the fixed-effect halfway-triquarter parameter (γ_{fixed}) for each measurement number and nature of change. The dots (squares, circles, triangles, diamonds) indicate the average estimated value (see bias). The horizontal blue line indicates the population value ($\gamma_{fixed} = 4.00$) and the gray band indicates the acceptable margin of error of ±10% of the parameter's population value. Dots that lie within the gray margin of error are filled and dots that lie outside of the margin remain unfilled. In the current example, the average value estimated for the fixed-effect halfway-triquarter delta parameter (γ_{fixed}) is only biased (i.e., lies outside the margin of error) with five measurements and a nature of change with an early halfway-elevation point (i.e., β_{fixed} = 80). Therefore, estimates are unbiased in almost all cells.

Figure 2.1

Parameter Estimation Plot for the Fixed-Effect Days-to-Halfway Elevation Parameter (γ_{fixed})



Note. Dots (squares, circles, triangles, diamonds) indicate the average estimated value and error bars show the range of values covered by the middle 95% of the estimated values (see Precision). The horizontal blue line indicates the population value (γ_{fixed} = 4.00) and the gray band indicates the acceptable margin of error (i.e., ±10% of the population value) for bias. Dots that lie outside of the margin of error are unfilled and are considered biased estimates. Dots that lie inside the margin of error are filled and considered unbiased estimates. Error bars whose upper and/or lower whisker lengths exceed 10% of the parameter's population value are light blue and indicate parameter estimation that is not precise. Error bars whose upper and/or lower whisker lengths do not excced 10% of the parameter's population value are black and indicate parameter estimation that is precise.

2.1.4.3 Analysis and Visualization of Precision

As discussed previously, precision was defined as the range of values covered by the middle 95% of estimated values for a given parameter (see precision). The cutoff value used to estimate precision directly followed from the cutoff value used for bias. Given that bias values within a ±10% of a parameter's population value were deemed acceptable, an acceptable value for precision should not allow any bias values above the ±10% cutoff. That is, the range covered by the middle 95% of estimated values should not allow a bias value outside the the ±10% cutoff. If the range of values covered by the middle 95% of estimate values is conceptualized as an error bar centered on the population value, then an acceptable value for precision implies that neither the lower nor upper whiskers have a length greater than 10% of the parameter's population value. In summary, I deemed precision acceptable if no estimate within the range of values covered by the middle 95% of estimated values had a bias value greater than 10% of the population value, which also means that neither the lower nor upper whiskers of the error bar have a length greater than 10% of the population value.

Like bias, I also depicted precision in parameter estimation plots using error bar. Each error bar in the parameter estimation plot of Figure 2.1 indicates the range of values covered by the middle 95% of estimated values in the given cell for the fixed-effect halfway-triquarter delta parameter (γ_{fixed}). Importantly, if estimation is not precise, then at least one of the lower and/or upper whisker lengths exceeds 10% of the parameter's population value. When estimation is not precise, the error bar is light blue. When estimation is precise (i.e., neither of the lower or upper whisker lengths exceed 10% of the parameter's population value), the corresponding error bar is black. In the current example, all error bars are light blue and so precision is low in all cells.

2.1.4.3.1 Effect Size Computation for Precision

One last statistic I calculated was an effect size value to estimate the variance in parameter estimates for each independent variable. Among the several effect size metrics—at a broad level, effect size metrics can represent standardized differences or variance-accounted-for measures that are corrected or uncorrected for sampling error—the corrected variance-accounted-for effect size metric of partial ω^2 was chosen because of three desirable properties. First, partial ω^2 provides a less biased estimate of effect size than other variance-accounted-for measures (Okada, 2013). Second, partial ω^2 is more robust to assumption violations of normality and homogeneity of variance (Yigit & Mendes, 2018). Given that parameter estimates were often non-normally distributed across cells, effect size values computed with using partial ω^2 should be relatively less biased than other variance-accounted-for effect size metrics (e.g., η^2). Third, being partial effect size, partial ω^2 provides an effect size estimate that is not diluted by the inclusion of unaccountable variance in the denominator. To compute partial ω^2 value for each experimental effect, Equation 2.5 shown below was used:

$$\text{partial}\omega^2 = \frac{\sigma^2_{effect}}{\sigma^2_{effect} + MSE} \tag{2.5}$$

where σ^2_{effect} represents the variance accounted by an effect and MSE is the mean squared error. Importantly, σ^2_{effect} values were corrected values obtained by using the following formula in Equation 2.6 for a two-way factorial design with fixed variables (Howell, 2009):

$$\sigma^2_{effect} = \frac{(a-1)(MS_{effect} - MS_{error})}{nab}, \tag{2.6}$$

where a is the number of levels in the effect, b is the number of levels in the second effect, and n is the cell size. The variance accounted by the interaction was computed using the following formula in Equation 2.7:

$$\sigma^2_{AxB} = \frac{(a-1)(b-1)(MS_{AxB} - MS_{error})}{nab}. \tag{2.7}$$

To compute partial ω^2 values for effects, a Brown-Forsythe test was computed and the appropriate sum-of-squares terms were used to compute partial ω^2 values. A Brown-Forsythe test was used because to protect against the biasing effects of skewed distributions(Brown & Forsythe, 1974), which were were observed in the parameter estimate distributions in the current simulation experiments. To compute the Brown-Forsythe test, median absolute deviations in each cell were computed by calculating the absolute difference between each i estimate and the median estimated value in the given experimental cell as shown in Equation 2.8 below:

$$\text{Median absolute deviation}_i = |\text{Parameter estimate}_i - \text{Median parameter estimate}_{cell}|. \tag{2.8}$$

An ANOVA was then computed on the median absolute deviation values (using the independent variables of the experiment as predictors), with the terms in Equation 2.5 extracted from the ANOVA output to compute partial ω^2 values.

2.2 Results and Discussion

In the sections that follow, I organize the results by presenting them for each spacing schedule (equal, time-interval increasing, time-interval decreasing, middle-and-extreme). The results are presented for each spacing schedule because answering my research questions first requires knowledge of these results. To answer my first question of whether modelling accuracy increases from spacing measurements during periods of change, I need to determine whether modelling accuracy increases by placing measurements near periods of change for each spacing schedule. To answer my second question of how to space measurements when the nature of change is unknown, modelling accuracy across all manipulated nature-of-change values must be calculated for each spacing schedule. The spacing schedule that obtains the lowest modelling accuracy across all nature-of-change values can then be determined as the best schedule to use.

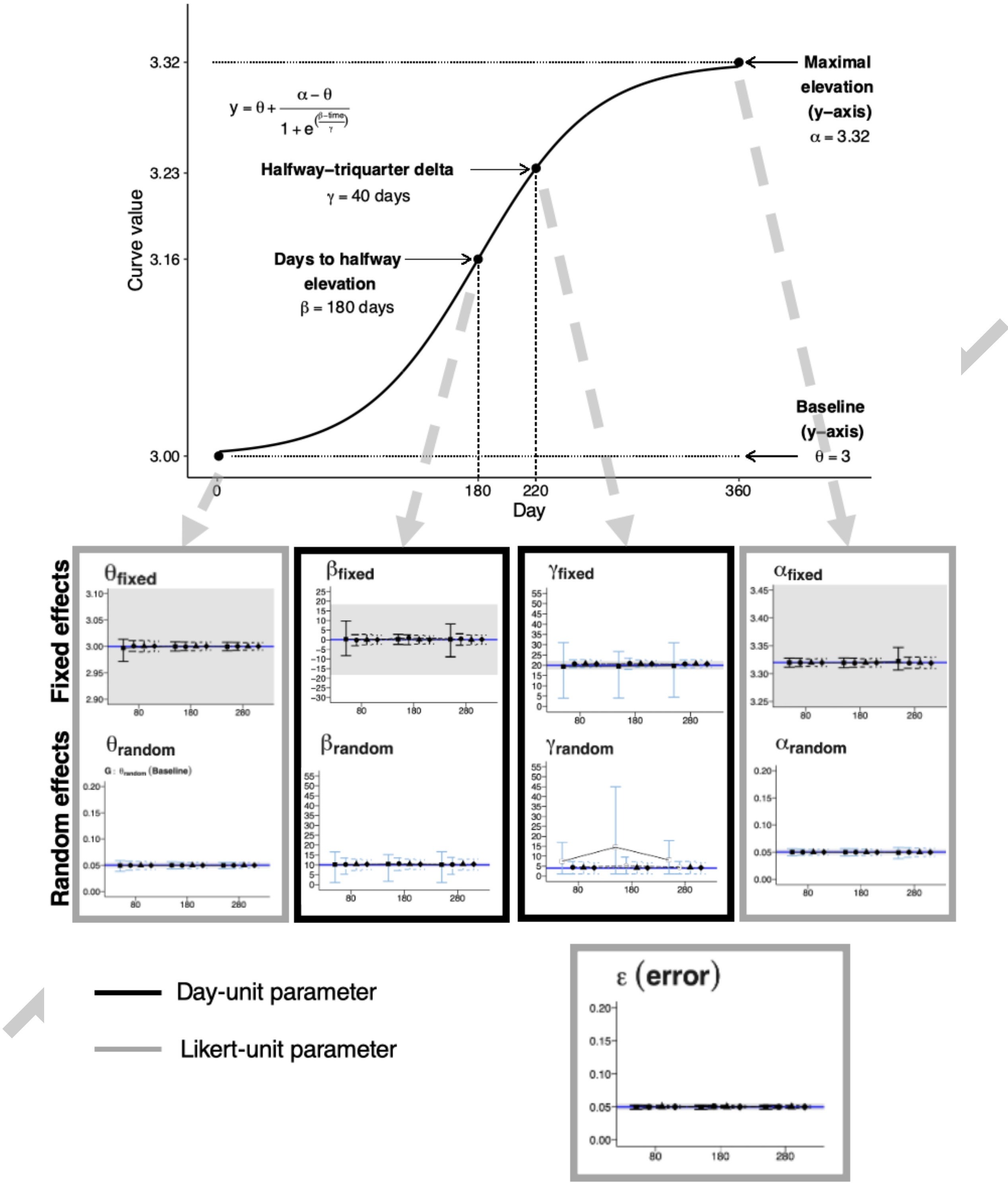
For each spacing schedule, I will first present a concise summary table of the results and then provide a detailed report for each column of the summary table. Because the lengths of the detailed reports are considerable, I provide concise summaries before the detailed reports to establish a framework to interpret the detailed reports. The detailed report of each spacing schedule presents the results of each day-unit's parameter estimation plot, modelling accuracy under each nature-of-change value, and then provides a qualitative summary. After providing the results for each spacing schedule, I then use the results to answer my research questions.

2.2.1 Framework for Interpreting Results

Because Experiment 1 (like all other experiments) had many cells (i.e., 48 cells in Experiment 1), the number of dependent variables to track in the results section can become overwhelming. Therefore, I will provide a framework to help the reader efficiently navigate the results section. Figure 2.2 shopws the entire set of results that will be presented for each spacing schedule. For each spacing schedule, a parameter estimation plot is created for each of the nine parameters estimated by the structured latent growth curve model used on each generated data set (for a review, see modelling of each generated data set). Parameter estimation plots with black outlines show the results for day-unit parameter and plots with gray outlines show the results for likert-unit parameters. Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and halfway-triquarter delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix B.

Figure 2.2

Set of Parameter Estimation Plots Constructed for Each Spacing Schedule in Experiment 1



Note. A parameter estimation plot is constructed for each parameter of the logistic function (see Equation 2.3). Note that each parameter of the logistic function is modelled as a fixed and random effect along with an error term (ϵ; for a review, see Figure 1.4).

2.2.2 Pre-Processing of Data and Model Convergence

After collecting the output from the simulations, non-converged models (and their corresponding parameter estimates) were removed from subsequent analyses. Table D.1 in Appendix C provides the convergence success rates for each cell in Experiment 1. Model convergence was almost always above 90% and convergence rates rates below 90% only occurred in two cells with five measurements.

2.2.3 Equal Spacing

For equal spacing, Table 2.3 provides a concise summary of the results for the day-unit parameters (see Figure 2.4 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 2.3 and provide elaboration when necessary.

Before presenting the results for equal spacing, I provide a brief description of the concise summary table created for each spacing schedule and shown for equal spacing in Table 2.3. Text within the 'Highest Modelling Accuracy' indicates the nature-of-change value that leads to the highest modelling accuracy for each day-unit parameter. Text within the 'Unbiased' and 'Precise' columns indicates the number of measurements needed to, respectively, obtain unbiased and precise parameter estimation across all manipulated nature-of-change values. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number that, respectively, results in unbiased estimation and the greatest improvements in bias and precision across all day-unit parameters and manipulated nature-of-change values. The 'Error Bar Length' column indicates the average error bar length across all manipulated nature-of-change values that results from using the measurement number listed in the 'Qualitative Description' column.

Table 2.3

Concise Summary of Results for Equal Spacing in Experiment 1

				Description	
Parameter	Highest Modelling Accuracy	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.4A)	β_{fixed} = 180	All cells	All cells	Largest improvements in precision with **NM = 7**	5.64
γ_{fixed} (Figure 2.4B)	β_{fixed} = 180	All cells	No cells	Largest improvements in precision with **NM = 7**	4.37
β_{random} (Figure 2.4C)	β_{fixed} = 180	All cells	No cells	Largest improvements in precision with **NM = 7**	7.74
γ_{random} (Figure 2.4D)	β_{fixed} = 180	**NM $\geq$ 9**	No cells	Largest improvements in bias and precision with **NM = 7**	7.02

Note. ‘Highest Modelling Accuracy’ indicates the curve that results in the highest modelling accuracy. Emboldened text in the ‘Unbiased’ and ‘Qualitative Description’ columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not obtained in the estimation of all day-unit parameters with equal spacing). ‘Error Bar Length’ indicates the average error bar length value across all nature-of-change values that results from using the measurement number in the ‘Qualitative Description‘ column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

2.2.3.1 Nature of Change That Leads to Highest Modelling Accuracy

For equal spacing, Table 2.4 lists the precision values (i.e., error bar lengths) for each day-unit parameter across each nature-of-change value. The 'Total' column indicates the total error bar length, which is a sum of the the lower ('Lower') and upper ('Upper') whisker lengths. Given that the lower and upper whisker lengths are largely equivalent for each parameter, they are largely redundant and so will not be reported for the remainder of the results for equal spacing. Although modelling accuracy is determined by bias and precision, results for bias are not shown because the differences in bias across the nature-of-change values are negligible. Note that error bar lengths are obtained by computing the average length across all manipulated number of measurements. The columns shaded in gray indicate the nature of change where precision is highest (i.e., shortest error bar lengths) for equal spacing. For equal spacing, precision is lowest with a nature-of-change value of 180 for all day-unit parameters with one exception (i.e., midway change; see the 'Highest Modelling Accuracy' column in Table 2.3).

One important result to discuss concerns the error bar length of the random-effect triquarter-halfway elevation parameter (γ_{random}) across the nature-of-change values. Precision is highest (i.e, shortest error bars) when the halfway-elevation point occurs midway through the measurement period (i.e., 180 days) for each day-unit parameter except the random-effect triquarter-halfway elevation parameter (γ_{random}). For the random-effect triquarter-halfway elevation parameter (γ_{random}), precision is lowest (i.e., longest error bars) when the halfway-elevation point occurred midway through the measurement period (i.e., 180 days). An inspection of the error bar lengths for the random-effect triquarter-halfway elevation parameter (γ_{random}) in Figure 2.4D reveals that the lower precision (i.e., longer error bars) observed for γ_{random} with a nature of change of 180 results from a longer upper whisker with five measurements.

To understand why precision for the random-effect triquarter-halfway elevation parameter (γ_{random}) is lower with a nature-of-change value of 180, I looked at the

Table 2.4

Error Bar Lengths Across Nature-of-Change Values Under Equal Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.4A)	4.42	4.12	8.54	2.46	2.32	4.78	4.09	4.16	8.25
γ_{fixed} (Figure 2.4B)	4.84	4.69	9.53	4.95	3.7	8.65	4.79	4.65	9.44
β_{random} (Figure 2.4C)	4.74	3.88	8.62	3.96	3.55	7.51	4.77	4.05	8.82
γ_{random} (Figure 2.4D)	3.00	5.52	8.52	3.00	13.05[a]	16.05	3.00	5.78	8.78

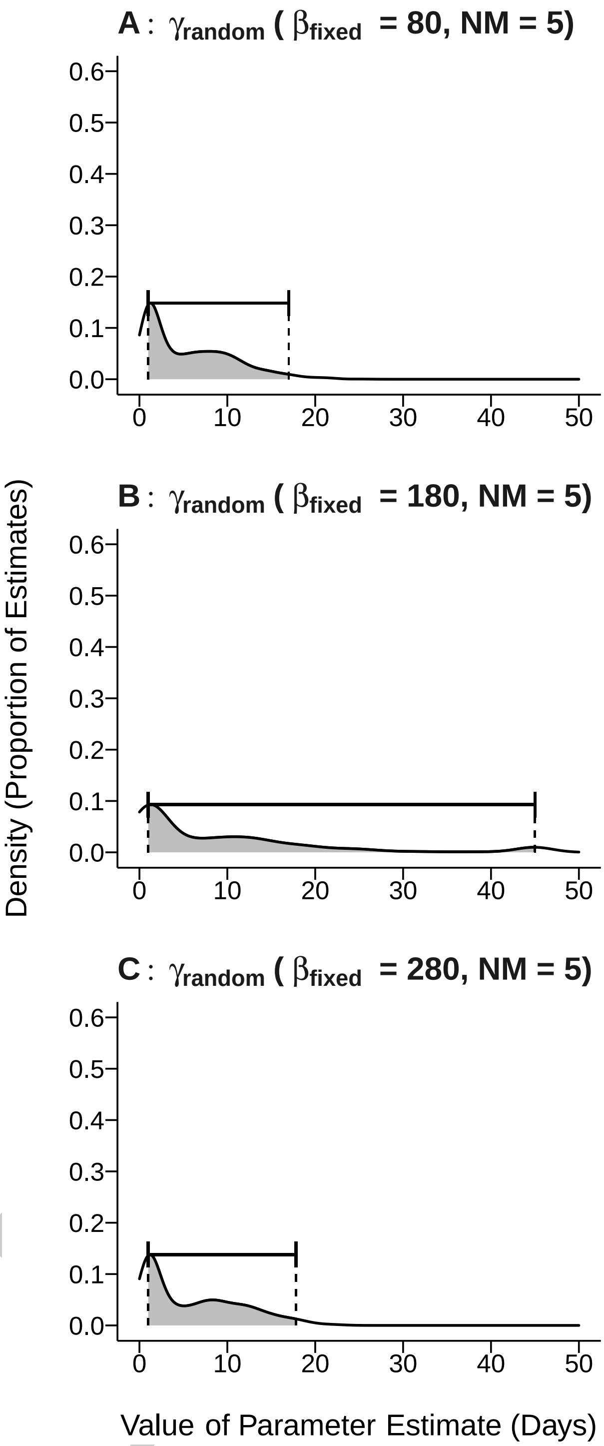
Note. 'Total' indicates the total error bar length, which is a sum of the lower ('Lower') and upper ('Upper') whisker lengths. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in \{5, 7, 9, 11\}$). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

distribution of estimated values. Figure 2.3 shows the distribution of values (i.e., density plots) estimated for the random-effect triquarter-halfway elevation parameter (γ_{random}) for each nature-of-change level with five measurements. Panel A shows the density plot with an early halfway point (i.e., β_{fixed} = 80). Panel B shows the density plot with a midway halfway point (i.e., $\beta_{fixed} = 180$). Panel C shows the density plot with a late halfway point (i.e., $\beta_{fixed} = 280$). Regions shaded in in gray represent the the middle 95% of estimated values and the width of the shaded regions is indicated by the length of the horizontal error bars. As originally confirmed by Table 2.4, Figure 2.3B shows that precision is indeed lowest (i.e., longer error bars) with a nature of change of 180. In looking across the density plots in Figure 2.3, precision is lowest (i.e., longest error bars) for the random-effect triquarter-halfway parameter (γ_{random}) with a nature-of-change value of 180 because of the existence of high-value outliers.

In summary, under equal spacing, modelling accuracy for all the day-unit parameters (with one exception) is greatest when the nature-of-change value set by the fixed-effect days-to-halfway elevation parameter (β_{fixed}) has a value of 180. As one exception, modelling accuracy (as indicated by precision) is lower for the random-effect triquarter-halfway elevation parameter (γ_{random}) with a nature-of-change value of 180 because of high-value estimates.

Figure 2.3

Density Plots of the Random-Effect Halfway-Triquarter Delta (γ_{random}; Figure 2.4D) With Equal Spacing in Experiment 1 (95% Error Bars)



Note. Regions shaded in in gray represent the the middle 95% of estimated values and the width of the shaded regions is indicated by the length of the horizontal error bars. The error bar length if longest when the nature-of-change value is 180. γ_{random} = random-effect triquarter-halfway delta parameter, with population value of 4.00, NM = number of measurements.

2.2.3.2 Bias

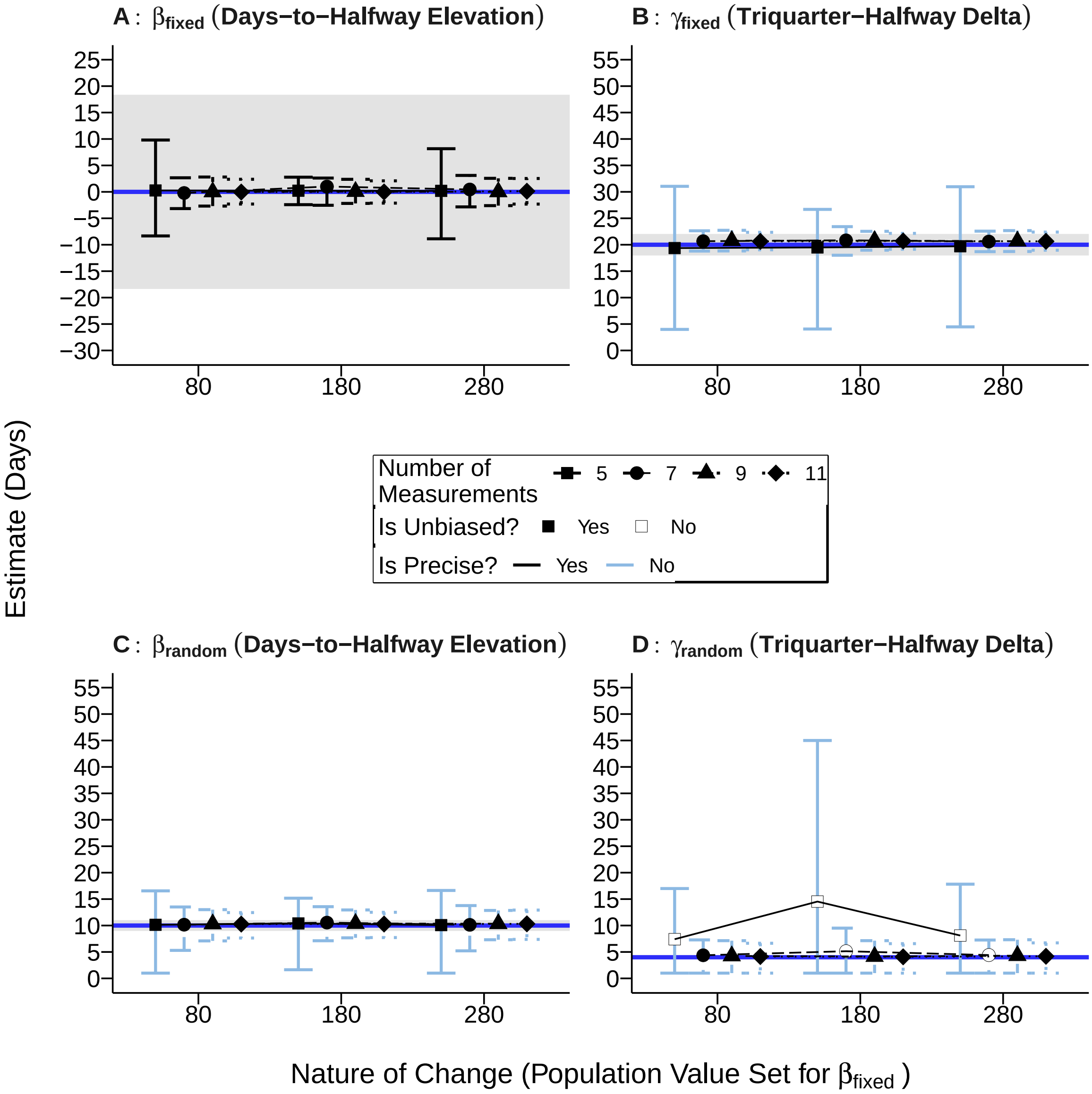
Before presenting the results for bias, I provide a description of the set of parameter estimation plots shown in Figure 2.4 and in the results sections for the other spacing schedules in Experiment 1. Figure 2.4 shows the parameter estimation plots for each day-unit parameter and Table 2.5 provides the partial ω^2 values for each independent variable of each day-unit parameter. In Figure 2.4, blue horizontal lines indicate the population values for each parameter (with population values of $\beta_{fixed} \in \{80, 180, 280\}$, $\beta_{random} = 10.00$, $\gamma_{fixed} = 20.00$, and $\gamma_{random} = 4.00$). Gray bands indicate the ±10% margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Panels A–B show the parameter estimation plots for the fixed- and random-effect days-to-halfway elevation parameters (β_{fixed} and β_{random}, respectively). Panels C–D show the parameter estimation plots for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively). Note that random-effect parameter units are in standard deviation units. Importantly, across all population values used for the fixed-effect days-to-halfway elevation parameter (β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180.

With respect to bias for equal spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.4A): no cells.

Figure 2.4

Parameter Estimation Plots for Day-Unit Parameters With Equal Spacing in Experiment 1



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table ?? for specific values estimated for each parameter and Table 2.5 for ω^2 effect size values.

Table 2.5
Partial ω^2 Values for Manipulated Variables With Equal Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.4A)	0.02	0.00	0.01
β_{random} (Figure 2.4B)	0.29	0.02	0.02
γ_{fixed} (Figure 2.4C)	0.36	0.01	0.03
γ_{random} (Figure 2.4D)	0.21	0.03	0.04

Note. NM = number of measurements $\in$ {5, 7, 9, 11}, NC = nature of change (population value set for β_{fixed} $\in$ {80, 180, 280}), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 2.4B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.4C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 2.4D): five measurements with all manipulated nature-of-change values and seven measurements with nature-of-change values of 180 and 280.

In summary, with equal spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values is unbiased using nine or more measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.3.

2.2.3.3 Precision

With respect to precision for equal spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 2.4B): all cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.4C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 2.4D): all cells.

In summary, with equal spacing, estimation across all manipulated nature-of-change values is only precise for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) with five or more measurements. No manipulated measurement number results in precise estimation of the fixed-effect triquarter-halfway delta parameter (γ_{fixed}) or the random-effect day-unit parameters (see the 'Precise' column of Table 2.3).

2.2.3.4 Qualitative Description

Although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurements numbers. With respect to bias under equal spacing, the largest improvements in bias across all manipulated nature-of-change values result from using the following measurement numbers for the following day-unit parameters (note that only the random-effect triquarter halfway delta parameter [γ_{random}] had instances of high bias):

- random-effect triquarter-halfway delta parameters (γ_{random}): seven measurements.

With respect to precision under equal spacing, the largest improvements precision in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) are obtained with following measurement numbers:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements, which results in a maximum error bar length of 4.37 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements, which results in a maximum error bar length of 7.74 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements,which results in a maximum error bar length of 7.02 days.

Therefore, for equal spacing, seven measurements results leads to the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the emboldened text in the 'Qualitative Description' column of Table 2.3).

2.2.3.5 Summary of Results

In summarizing the results for equal spacing, modelling accuracy is greatest across all day-unit parameters with a nature-of-change value of 180, with the random-effect days-to-halfway elevation parameter (γ_{random}) being an exception (see highest modelling accuracy). Unbiased estimation of all the day-unit parameters across all manipulated nature-of-change values results from using nine or more measurements (see bias). Precise estimation of all the day-unit parameters is never obtained with any manipulated measurement number (see precision). Although it may be discouraging that no manipulated measurement number under equal spacing results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement numbers. With equal spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values are obtained using seven measurements (see Qualitative Description).

2.2.4 Time-Interval Increasing Spacing

For time-interval increasing spacing, Table 2.6 provides a concise summary of the results for the day-unit parameters (see Figure 2.5 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 2.6 and provide elaboration when necessary (for a description of Table 2.6, see concise summary table).

Table 2.6

Concise Summary of Results for Time-Interval Increasing Spacing in Experiment 1

				Description	
Parameter	Highest Modelling Accuracy	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.5A)	β_{fixed} = 80	All cells	NM $\geq$ 7	Largest improvement in precision with NM = 7	8.38
γ_{fixed} (Figure 2.5B)	β_{fixed} = 80	All cells	No cells	Largest improvement in precision with **NM = 9**	3.45
β_{random} (Figure 2.5C)	β_{fixed} = 80	NM $\geq$ 7	No cells	Largest improvement in bias and precision with NM = 7	9.47
γ_{random} (Figure 2.5D)	β_{fixed} = 80	**NM $\geq$ 9**	No cells	Largest improvements in bias and precision with **NM = 9**	5.97

Note. 'Highest Modelling Accuracy' indicates the curve that results in the highest modelling accuracy. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not obtained in the estimation of all day-unit parameters with time-interval increasing spacing). 'Error Bar Length' indicates the average error bar length value across all nature-of-change values that results from using the measurement number in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

2.2.4.1 Nature of Change That Leads to Highest Modelling Accuracy

For time-interval increasing spacing, Table 2.7 lists the precision values (i.e., error bar lengths) for each day-unit parameter across each nature-of-change value. The 'Total' column indicates the total error bar length, which is a sum of the the lower ('Lower') and upper ('Upper') whisker lengths. Given that the lower and upper whisker lengths are largely equivalent for each parameter, they are largely redundant and so will not be reported for the remainder of the results for time-interval increasing spacing. Although modelling accuracy is determined by bias and precision, results for bias are not shown because the differences in bias across the nature-of-change values are negligible. Note that error bar lengths are obtained by computing the average length across all manipulated number of measurements. The columns shaded in gray indicate the nature of change where precision is highest (i.e., shortest error bar lengths). For time-interval increasing spacing, precision is lowest (i.e., longest error bars) with a nature-of-change value of 80 for all day-unit parameters (i.e., early halfway point; see the 'Highest Modelling Accuracy' in Table 2.6).

2.2.4.2 Bias

With respect to bias for time-interval increasing spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.5A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 2.5B): no cells
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.5C): five measurements with a nature-of-change value of 280.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.5C): five mea-

surements with all nature-of-change values and seven measurements with nature-of-change values of 180 and 280.

Table 2.7

Error Bar Lengths Across Nature-of-Change Values Under Time-Interval Increasing Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.5A)	3.04	2.76	5.80	3.90	6.72	10.62	17.87	14.84	32.71
γ_{fixed} (Figure 2.5B)	1.59	2.81	4.40	4.39	3.21	7.60	9.00	6.38	15.38
β_{random} (Figure 2.5C)	3.55	3.25	6.80	4.41	4.18	8.59	6.20	9.60	15.81
γ_{random} (Figure 2.5D)	3.00	3.34	6.34	3.00	4.10	7.10	3.00	7.09	10.09

Note. 'Total' indicates the total error bar length, which is a sum of the lower ('Lower') and upper ('Upper') whisker lengths. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in$ {5, 7, 9, 11}). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

In summary, with time-interval increasing spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values is unbiased using nine or more measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.6.

2.2.4.3 Precision

With respect to precision for time-interval increasing spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

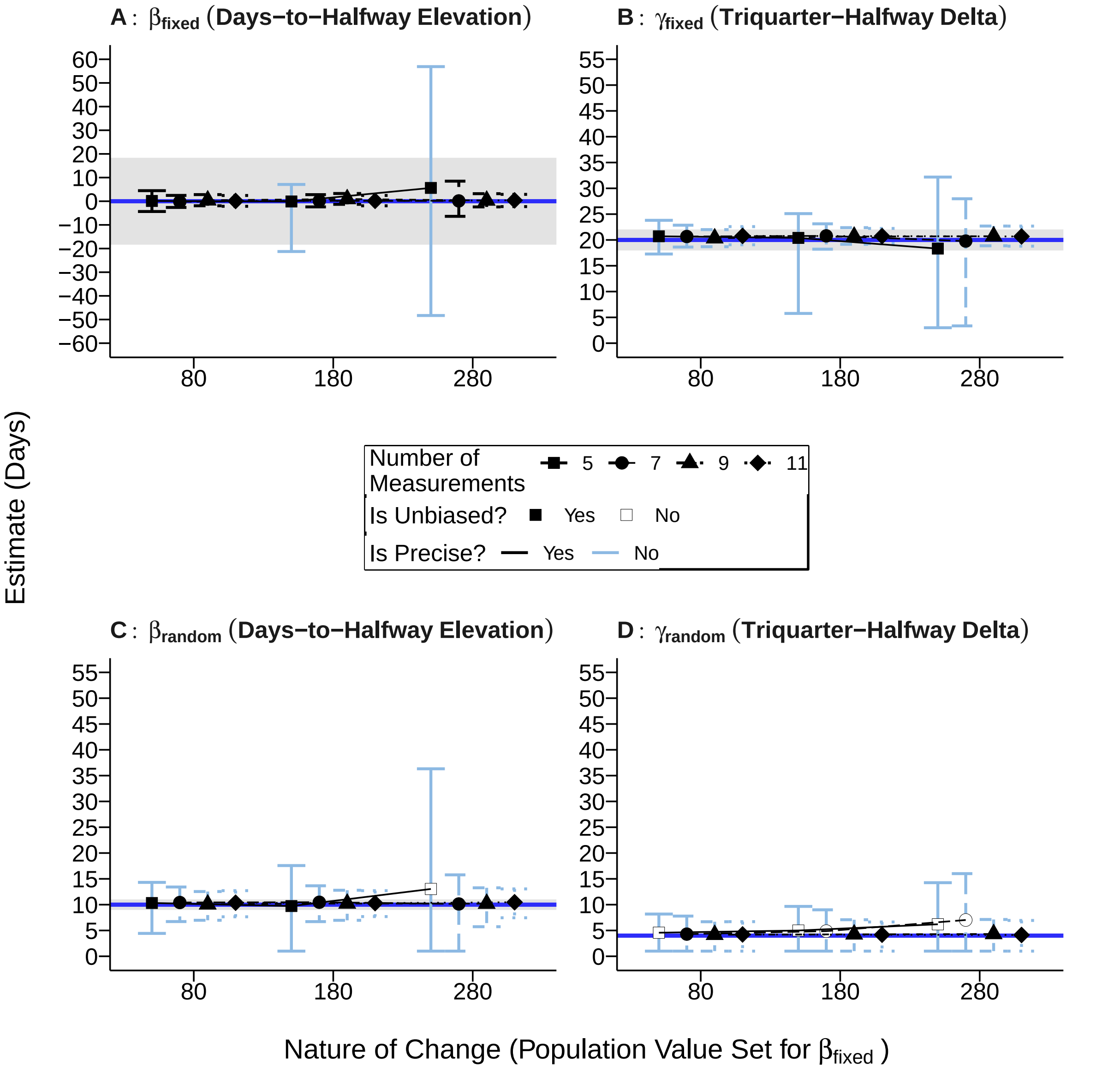
- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.5A): five measurements with nature-of-change values of 180 and 280.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.5B): all cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.5C): all cells.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.5D): all cells.

In summary, with time-interval increasing spacing, precise estimation of the fixed-effect day-unit parameters across all manipulated nature-of-change values results with seven or more measurements, but no manipulated measurement number results in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 2.6).

In summary, with time-interval increasing spacing, estimation across all manipulated nature-of-change values is only precise for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) with seven or more measurements. No manipulated measurement number results in precise estimation of the fixed-effect triquarter-halfway delta parameter (γ_{fixed}) or the random-effect day-unit parameters (see the 'Precise' column of Table 2.6).

Figure 2.5

Parameter Estimation Plots for Day-Unit Parameters With Time-Interval Increasing Spacing in Experiment 1



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80.00, 180.00, 280.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note

that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table ?? for specific values estimated for each parameter and Table 2.8 for ω^2 effect size values.

Table 2.8
Partial ω^2 Values for Manipulated Variables With Time-Interval Decreasing Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.5A)	0.43	0.30	0.50
β_{random} (Figure 2.5B)	0.12	0.04	0.05
γ_{fixed} (Figure 2.5C)	0.26	0.21	0.22
γ_{random} (Figure 2.5D)	0.12	0.05	0.04

Note. NM = number of measurements $\in \{5, 7, 9, 11\}$, NC = nature of change (population value set for $\beta_{fixed} \in \{80, 180, 280\}$), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

2.2.4.4 Qualitative Description

For time-interval increasing spacing in Figure 2.5, although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurements numbers. With respect to bias under time-interval increasing spacing, the largest improvements across all manipulated nature-of-change values in bias occur with the following measurement numbers for the random-effect day-unit parameters:

- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements.

- random-effect triquarter-halfway delta parameters (γ_{random}): nine measurements.

With respect to precision under time-interval increasing spacing, the largest improvements precision in the estimation of all day-unit parameters across all manipulated nature-of-change values result with following measurement numbers:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}): seven measurements, which results in an average error bar length of 8.38 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): nine measurements, which results in an average error bar length of 3.45 days.
- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements, which results in an average error bar length of 9.47 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): nine measurements, which results in an average error bar length of 5.97 days.

Therefore, for time-interval increasing spacing, nine measurements leads to the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the 'Qualitative Description' column in Table 2.6).

2.2.4.5 Summary of Results

In summarizing the results for time-interval increasing spacing, modelling accuracy is highest for each day-unit parameter with a nature-of-change value of 80 (i.e., early halfway point; see highest modelling accuracy). Estimation of all day-unit parameters is unbiased across all manipulated nature-of-change values using nine or more measurements (see bias). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement (see precision). Although it may be discouraging that no manipulated measurement number under time-interval increasing spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement numbers. With time-interval increasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values are obtained using nine measurements (see qualitative description).

2.2.5 Time-Interval Decreasing Spacing

For time-interval decreasing spacing, Table 2.9 provides a concise summary of the results for the day-unit parameters (see Figure 2.6 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 2.9 and provide elaboration when necessary (for a description of Table 2.9, see concise summary table).

Table 2.9
Concise Summary of Results for Time-Interval Decreasing Spacing in Experiment 1

				Description	
Parameter	Highest Modelling Accuracy	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.6A)	β_{fixed} = 280	All cells	NM $\geq$ 9	Largest improvements in precision with **NM = 9**	4.88
γ_{fixed} (Figure 2.6B)	β_{fixed} = 280	NM $\geq$ 7	No cells	Largest improvement in precision with **NM = 9**	3.40
β_{random} (Figure 2.6C)	β_{fixed} = 280	NM $\geq$ 7	No cells	Largest improvement in bias and precision with **NM = 9**	6.15
γ_{random} (Figure 2.6D)	β_{fixed} = 280	**NM $\geq$ 9**	No cells	Largest improvements in bias and precision with **NM = 9**	5.96

Note. 'Highest Modelling Accuracy' indicates the curve that results in the highest modelling accuracy. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not obtained in the estimation of all day-unit parameters with time-interval decreasing spacing). 'Error Bar Length' indicates the average error bar length value across all nature-of-change values that results from using the measurement number in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

2.2.5.1 Nature of Change That Leads to Highest Modelling Accuracy

For time-interval decreasing spacing, Table 2.10 lists the error bar lengths for each day-unit parameter and nature-of-change value. Although modelling accuracy is determined by bias and precision, results for bias are not discussed or computed because the differences in bias across the nature-of-change values are negligible. Given that the lower and upper whisker lengths are largely equivalent for each parameter, they are largely redundant and so will not be reported for the remainder of the results for time-interval decreasing spacing. Note that error bar lengths are computed by computing the average length across all manipulated number-of-measurement values. The column shaded in gray indicates the nature-of-change value that results in the shortest error bar lengths under equal spacing. For time-interval decreasing spacing, precision is lowest (i.e., longest error bars) with a nature-of-change value of 280 for all day-unit parameters (i.e., late halfway point; see the 'Highest Modelling Accuracy' in Table 2.9).

Table 2.10
Error Bar Lengths Across Nature-of-Change Values Under Time-Interval Decreasing Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.6A)	30.51	15.73	46.24	7.64	3.67	11.31	3.28	2.56	5.84
γ_{fixed} (Figure 2.6B)	9.70	6.11	15.81	4.88	3.14	8.02	1.79	2.69	4.48
β_{random} (Figure 2.6C)	6.09	11.26	17.35	4.70	3.90	8.60	3.60	3.13	6.73
γ_{random} (Figure 2.6D)	3.00	6.57	9.57	3.00	4.20	7.20	3.00	3.24	6.24

Note. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in$ {5, 7, 9, 11}). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

2.2.5.2 Bias

With respect to bias for time-interval decreasing spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.6A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 2.6B): five measurements with a nature-of-change value of 80.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.6C): five measurements with a nature-of-change value of 80.
- random-effect halfway-triquarter delta parameter (γ_{random}; Figure 2.6D): five measurements across all manipulated nature-of-change values and seven measurements with nature-of-change values of 80 and 180.

In summary, with time-interval decreasing spacing, unbiased estimation can be obtained for all day-unit parameters across all manipulated nature-of-change values using nine or more measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.9.

2.2.5.3 Precision

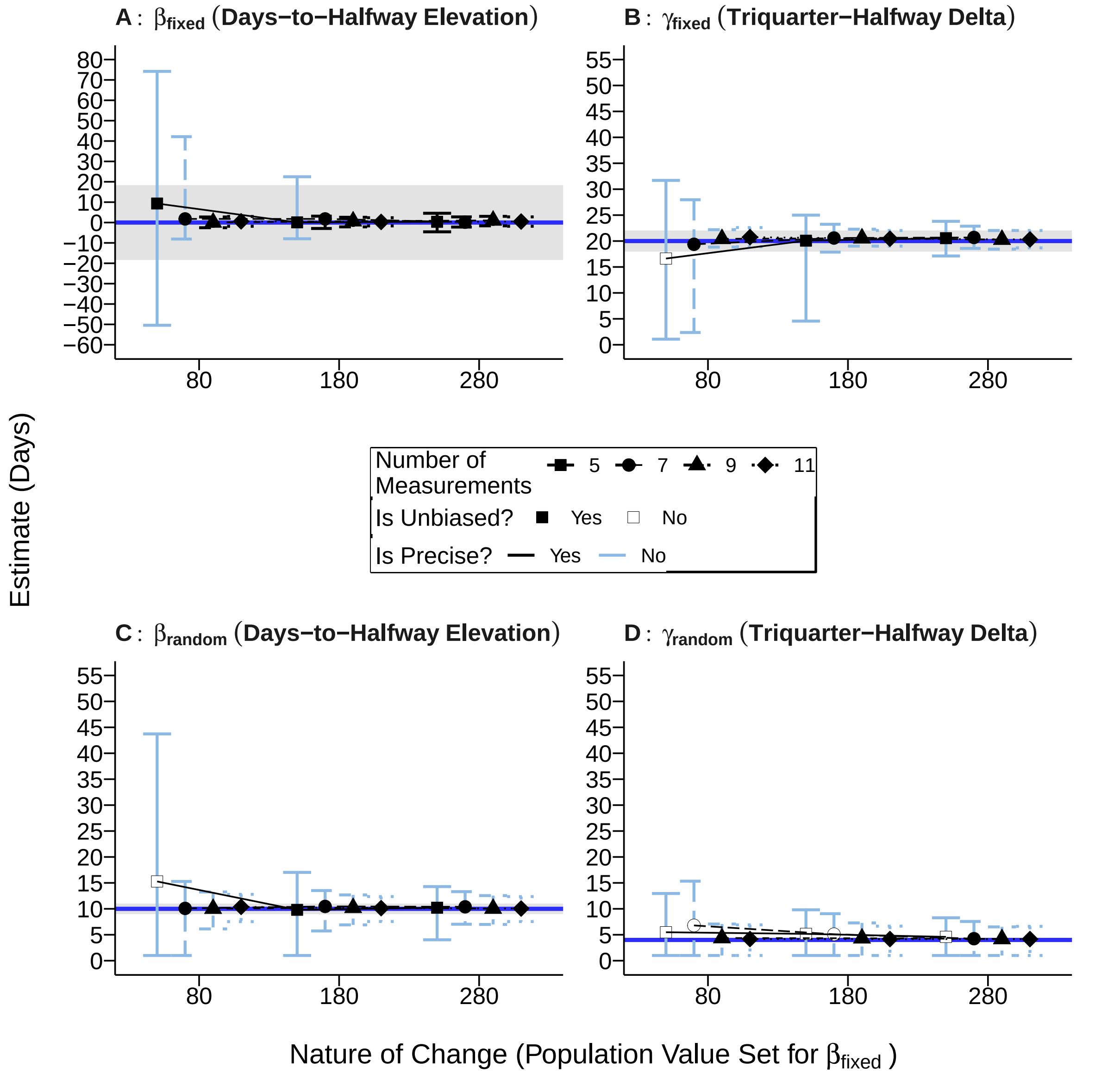
With respect to precision for time-interval decreasing spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.6A): five measurements with nature-of-change values of 80 and 180 an seven measurements with a nature-of-change value of 80.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.6B): all cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.6C): all cells.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.6D): all cells.

In summary, with time-interval increasing spacing, estimation across all manipulated nature-of-change values is only precise for the estimation of the fixed-effect days-to-halfway elevation parameter (β_{fixed}) with nine or more measurements. No manipulated measurement number results in precise estimation of the fixed-effect triquarter-halfway delta parameter (γ_{fixed}) or the random-effect day-unit parameters (see the 'Precise' column of Table 2.9).

Figure 2.6

Parameter Estimation Plots for Day-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 1



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80.00, 180.00, 280.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note

that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table ?? for specific values estimated for each parameter and Table 2.11 for ω^2 effect size values.

Table 2.11
Partial ω^2 Values for Manipulated Variables With Time-Interval Decreasing Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.6A)	0.20	0.10	0.22
β_{random} (Figure 2.6B)	0.13	0.04	0.05
γ_{fixed} (Figure 2.6C)	0.27	0.19	0.21
γ_{random} (Figure 2.6D)	0.11	0.03	0.03

Note. NM = number of measurements $\in$ {5, 7, 9, 11}, NC = nature of change (population value set for $\beta_{fixed} \in$ {80, 180, 280}), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

2.2.5.4 Qualitative Description

For time-interval decreasing spacing in Figure 2.6, although no manipulated measurement number results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) are obtained using moderate measurements numbers. With respect to bias under time-interval decreasing spacing, the largest improvements across all manipulated nature-of-change values in bias occur with the following measurement numbers for the random-effect day-unit parameters:

- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements

- random-effect triquarter-halfway delta parameters (γ_{random}): nine measurements

With respect to precision under time-interval decreasing spacing, the largest improvements precision in the estimation of all day-unit parameters across all manipulated nature-of-change values are obtained with following measurement numbers:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}): seven measurements, which results in a maximum error bar length of 20.42 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): nine measurements, which results in a maximum error bar length of 3.4 days.
- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements, which results in a maximum error bar length of 9.45 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): nine measurements, which results in a maximum bar length of 5.96 days.

Therefore, for time-interval decreasing spacing, nine measurements leads to the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the emboldened text in the 'Qualitative Description' column in Table 2.9).

2.2.5.5 Summary of Results

In summarizing the results for time-interval decreasing spacing, modelling accuracy is highest for each day-unit parameter with a nature-of-change value of 280 (i.e., late halfway point; see highest modelling accuracy). Unbiased estimation of the day-unit parameters across all manipulated nature-of-change values results from using nine or more measurements (see bias). Precise estimation of all the day-unit parameters is never obtained using any of the manipulated measurement numbers (see precision). Although it may be discouraging that no manipulated measurement number under time-interval decreasing spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement numbers. With time-interval decreasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values are obtained using nine measurements (see qualitative description).

2.2.6 Middle-and-Extreme Spacing

For middle-and-extreme spacing, Table 2.12 provides a concise summary of the results for the day-unit parameters (see Figure 2.7 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 2.12 and provide elaboration when necessary (for a description of Table 2.12, see concise summary table).

Table 2.12

Concise Summary of Results for Middle-and-Extreme Spacing in Experiment 1

				Description	
Parameter	Highest Modelling Accuracy	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 2.7A)	β_{fixed} = 180	All cells	NM $\geq$ 7	Largest improvements in precision with NM = 7	14.10
γ_{fixed} (Figure 2.7B)	β_{fixed} = 180	NM $\geq$ 7	No cells	Largest improvements in bias and precision with NM = 7	6.27
β_{random} (Figure 2.7C)	β_{fixed} = 180	NM $\geq$ 9	No cells	Largest improvements in bias and precision with **NM = 9**	9.02
γ_{random} (Figure 2.7D)	β_{fixed} = 180	**NM = 11**	No cells	Largest improvements in bias and precision with NM = 7	7.92

Note. ‘Highest Modelling Accuracy’ indicates the curve that results in the highest modelling accuracy. Emboldened text in the ‘Unbiased’ and ‘Qualitative Description’ columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not obtained in the estimation of all day-unit parameters with middle-and-extreme spacing). ‘Error Bar Length’ indicates the average error bar length value across all nature-of-change values that results from using the measurement number in the ‘Qualitative Description‘ column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

2.2.6.1 Nature of Change That Leads to Highest Modelling Accuracy

For middle-and-extreme spacing, Table 2.13 lists the error bar lengths for each day-unit parameter and nature-of-change value. Although modelling accuracy is determined by bias and precision, results for bias are not discussed or computed because the differences in bias across the nature-of-change values are negligible. Given that the lower and upper whisker lengths are largely equivalent for each parameter, they are largely redundant and so will not be reported for the remainder of the results for middle-and-extreme spacing. Note that error bar lengths are computed by computing the average length across all manipulated number-of-measurement values. The column shaded in gray indicates the nature-of-change value that results in the shortest error bar lengths under equal spacing. For middle-and-extreme spacing, precision is lowest (i.e., longest error bars) with a nature-of-change value of 180 for all day-unit parameters (i.e., midway halfway point; see the 'Highest Modelling Accuracy' in Table 2.12).

Table 2.13
Error Bar Lengths Across Nature-of-Change Values Under Middle-and-Extreme Spacing in Experiment 1

	Population Value of β_{fixed}								
	80			180			280		
Parameter	Lower	Upper	Total	Lower	Upper	Total	Lower	Upper	Total
β_{fixed} (Figure 2.7A)	22.13	19.89	42.02	2.25	2.21	4.46	20.32	21.74	42.06
γ_{fixed} (Figure 2.7B)	6.50	5.77	12.27	0.87	2.22	3.09	6.73	6.11	12.84
β_{random} (Figure 2.7C)	7.14	16.84	23.97	2.28	2.48	4.76	7.27	15.69	22.96
γ_{random} (Figure 2.7D)	3.00	6.20	9.20	3.00	2.73	5.73	3.00	6.77	9.77

Note. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 80, 180, 280; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. Note that error bar lengths were calculated by computing the average error bar length value across all number-of-measurement (NM) values (NM $\in \{5, 7, 9, 11\}$). Columns shaded in gray indicate the nature-of-change value that results in the shortest error bar and whisker lengths.

2.2.6.2 Bias

With respect to bias for middle-and-extreme spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.6A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 2.7B): five measurements with nature-of-change values of 80 and 280.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.7C): five and seven measurements with nature-of-change values of 80 and 280.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.7D): five, seven, and nine measurements with nature-of-change values of 80 and 280.

In summary, with middle-and-extreme spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values is unbiased using 11 measurements, which is indicated by the emboldened text in the 'Unbiased' column of Table 2.12.

2.2.6.3 Precision

With respect to precision for middle-and-extreme spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

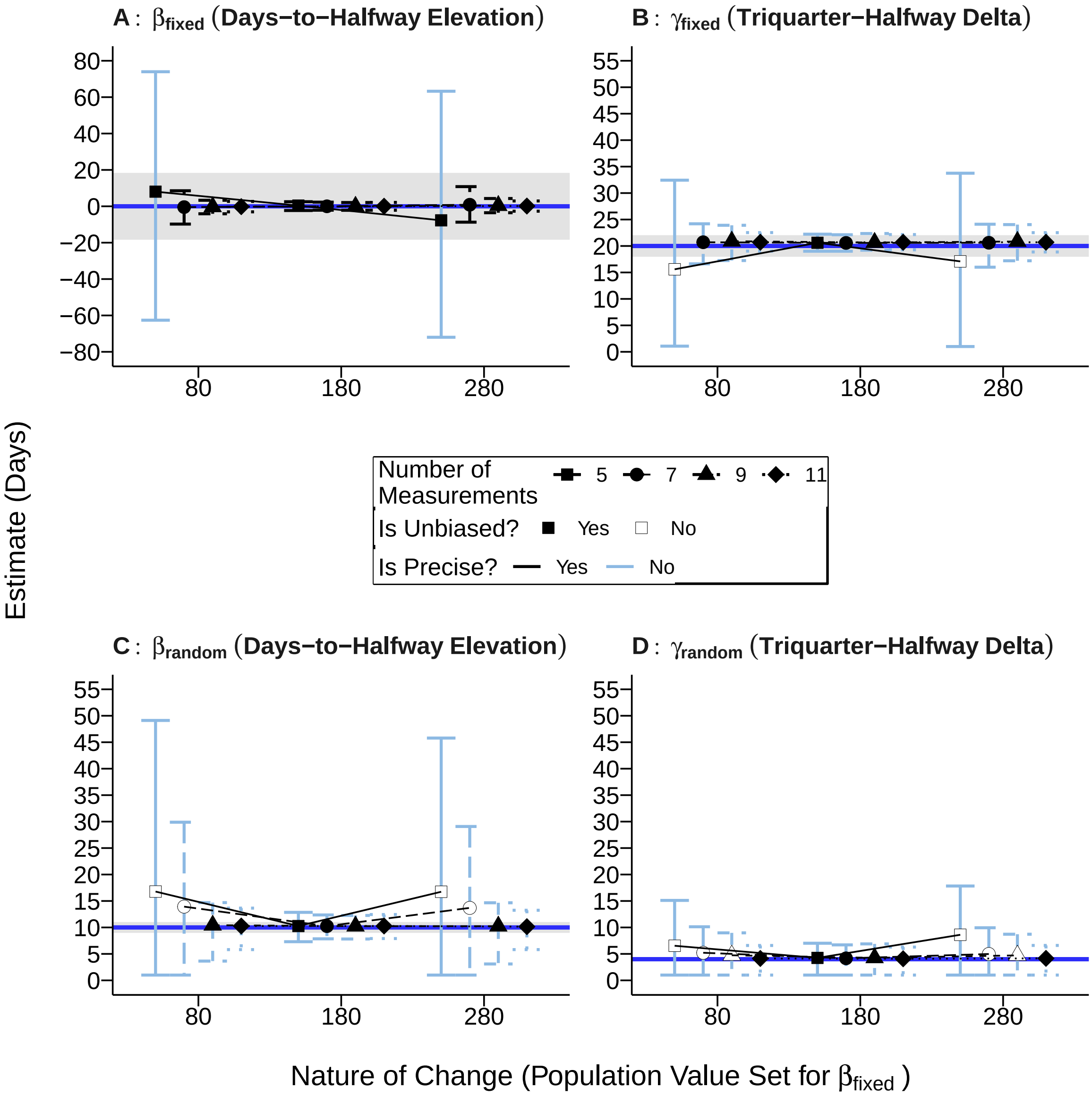
- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 2.7A): five measurements with nature-of-change values of 80 and 280.

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}; Figure 2.7B): five and seven, an nine measurements with nature-of-change values of 80 and 280 (shown on x-axis).
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 2.7C): all cells.
- random-effect triquarter-halfway delta parameter (γ_{random}; Figure 2.7D): all cells.

In summary, with middle-and-extreme spacing, precise estimation of the fixed-effect day-unit parameters across all manipulated nature-of-change values is obtained with 11 measurements, but no manipulated measurement number results in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 2.12).

Figure 2.7

Parameter Estimation Plots for Day-Unit Parameters With Middle-and-Extreme Spacing in Experiment 1



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80.00, 180.00, 280.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note

that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table ?? for specific values estimated for each parameter and Table 2.14 for ω^2 effect size values.

Table 2.14

Partial ω^2 Values for Manipulated Variables With Middle-and-Extreme Spacing in Experiment 1

	Effect		
Parameter	NM	NC	NM x NC
β_{fixed} (Figure 2.7A)	0.32	0.09	0.19
β_{random} (Figure 2.7B)	0.12	0.09	0.06
γ_{fixed} (Figure 2.7C)	0.49	0.20	0.32
γ_{random} (Figure 2.7D)	0.07	0.05	0.03

Note. NM = number of measurements $\in \{5, 7, 9, 11\}$, NC = nature of change (population value set for $\beta_{fixed} \in \{80, 180, 280\}$), NM x NC = interaction between number of measurements and population value set for β_{fixed}. β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

2.2.6.4 Qualitative Description

For middle-and-extreme spacing in Figure 2.7, although no manipulated measurement number results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) are obtained using moderate measurements numbers. With respect to bias under middle-and-extreme spacing, the largest improvements across all manipulated nature-of-change values in bias occur with the following measurement numbers for the following day-unit parameters:

- random-effect days-to-halfway elevation parameter (γ_{fixed}): seven measurements

- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements
- random-effect triquarter-halfway delta parameters (γ_{random}): 11 measurements

With respect to precision under middle-and-extreme spacing, the largest improvements precision in the estimation of all day-unit parameters across all manipulated nature-of-change values result with following measurement numbers:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}): seven measurements, which results in a maximum error bar length of 14.1 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements, which results in a maximum error bar length of 5.55 days.
- random-effect days-to-halfway elevation parameter (β_{random}): nine measurements, which results in a maximum error bar length of 20.49 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements, which results in a maximum error bar length of 7.2 days.

Therefore, for middle-and-extreme spacing, nine measurements obtain the greatest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values (see the emboldened text in the 'Qualitative Description' column in Table 2.12).

2.2.6.5 Summary of Results

In summarizing the results for time-interval decreasing spacing, modelling accuracy is highestfor each day-unit parameter with a a nature-of-change value of 180 (i.e., midway halfway point; see highest modelling accuracy). Unbiased estimation of the day-unit parameters across all manipulated nature-of-change values results from using nine or more measurements (see bias). Precise estimation of all the day-unit parameters is never obtained using any of the manipulated measurement numbers (see precision). Although it may be discouraging that no manipulated measurement number under time-interval decreasing spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement numbers. With time-interval decreasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters across all manipulated nature-of-change values are obtained using nine measurements (see qualitative description).

2.2.7 Addressing My Research Questions

2.2.7.1 When the Nature of Change is Suspected, How Should Measurements be Spaced?

Table 2.15 lists the nature-of-change value that each spacing schedule obtains its highest modelling accuracy along with the corresponding precision with which each day-unit parameter is estimated. Text in the 'Highest Modelling Accuracy' column indicates the nature-of-change with which each spacing schedule obtains its modelling accuracy. The 'Error Bar Summary' columns list the error bar lengths obtained for each day-unit parameter using the nature-of-change value listed in the 'Highest Modelling Accuracy' column.[15] Note that the error bar lengths are obtained by computing the average error bar length across all manipulated measurement numbers for the optimal nature-of-change value. Modelling accuracy for each spacing schedule is highest with the following nature-of-change values:

- equal spacing: $\beta_{fixed} = 180$ (i.e., midway halfway point)

[15] Bias values are not presented because the differences across the schedules are negligible.

- time-interval increasing spacing: $\beta_{fixed} = 80$ (i.e., early halfway point)
- time-interval decreasing spacing: $\beta_{fixed} = 280$ (i.e., late halfway point)
- middle-and-extreme spacing: $\beta_{fixed} = 180$ (i.e., midway halfway point)

To understand why the modelling accuracy of each spacing schedule is highest with a specific nature of change, it is important to consider the locations on the curve where each schedule samples data. Figure 2.8 shows the measurement locations (indicates by dots) where each spacing schedule samples data for each manipulated nature of change ($\beta_{fixed} \in \{80, 180, 180\}$). In Figure 2.8A, data are sampled according to the equal spacing schedule. In Figure 2.8B, data are sampled according to the time-interval increasing spacing schedule. In Figure 2.8C, data are sampled according to the time-interval decreasing spacing schedule. In Figure 2.8D, data are sampled according to the middle-and-extreme spacing schedule. Black curves indicate curves for which modelling accuracy is highest, with gray curves indicating curves where modelling accuracy is not at its highest. Error bar lengths printed on each panel to provide a reference indicate the precision with which each day-unit parameter is estimated when modelling accuracy is highest (values have been copied over from Table 2.15).

Table 2.15

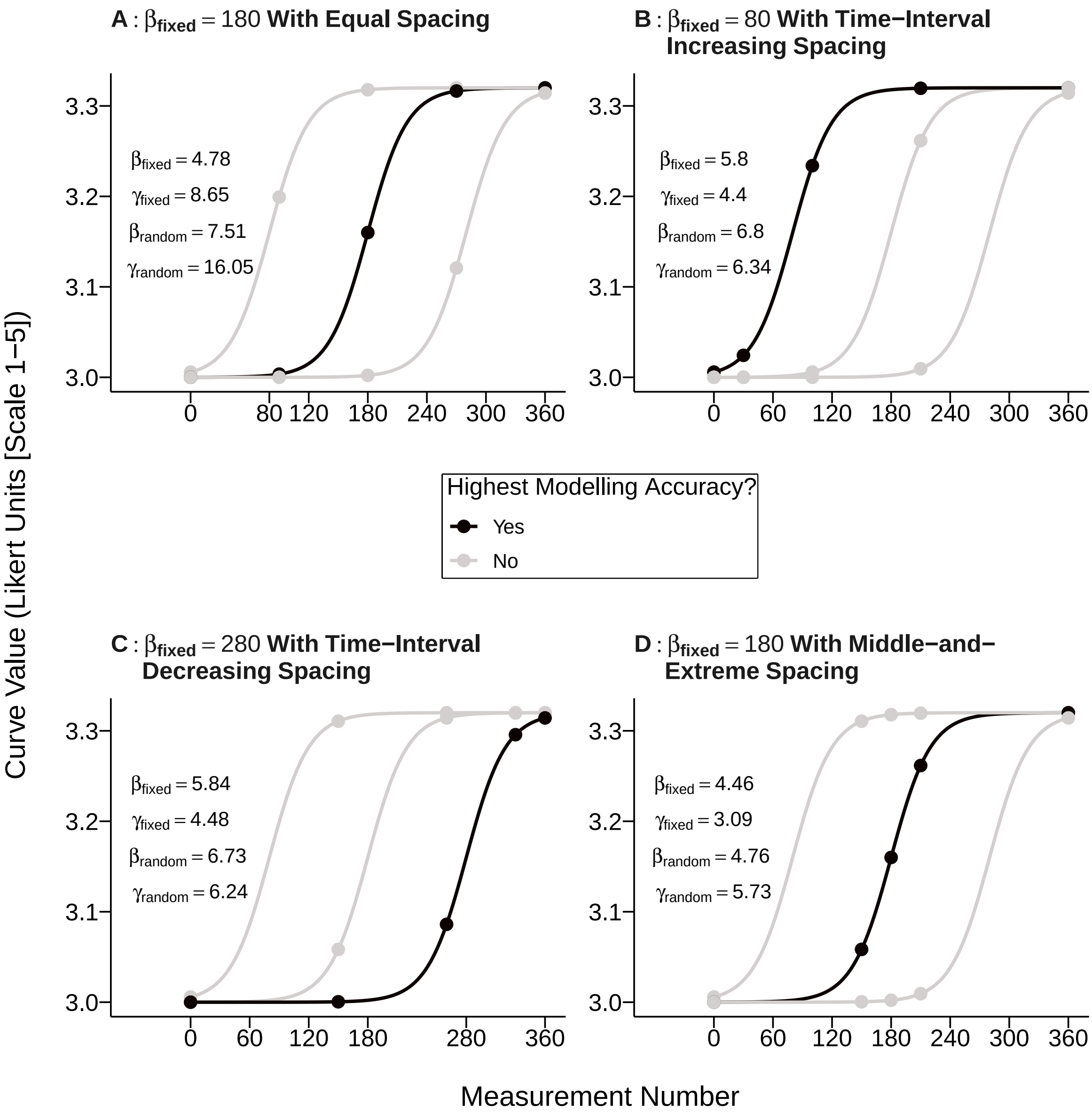
Nature-of-Change Values That Lead to the Highest Modelling Accuracy for Each Spacing Schedule in Experiment 1

		Error Bar Summary			
Spacing Schedule	Highest Modelling Accuracy	β_{fixed}	γ_{fixed}	β_{random}	γ_{random}
Equal (see Figure 2.4 and Table 2.4)	β_{fixed} = 180	4.78	8.65	7.51	16.05
Time-interval increasing (see Figure 2.5 and Table 2.7)	β_{fixed} = 80	5.80	4.40	6.80	6.34
Time-interval decreasing (see Figure 2.6 and Table 2.10)	β_{fixed} = 280	5.84	4.48	6.73	6.24
Middle-and-extreme (see Figure 2.7 and Table 2.13)	β_{fixed} = 180	4.46	3.09	4.76	5.73

Note. 'Highest Modelling Accuracy' indicates the curve that results in the highest modelling accuracy. 'Error Bar Summary' columns lists error bar lengths for each day-unit parameter such that error bar lengths are computed by taking the average error bar length value across all the number-of-measurement (NM) values (NM $\in \{5, 7, 9, 11\}$). Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter $\in \{80, 180, 280\}$; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4.

Figure 2.8

Nature-of-Change Curves for Each Spacing Schedule Have Highest Modelling Accuracy When Measurements are Taken Near Periods of Change

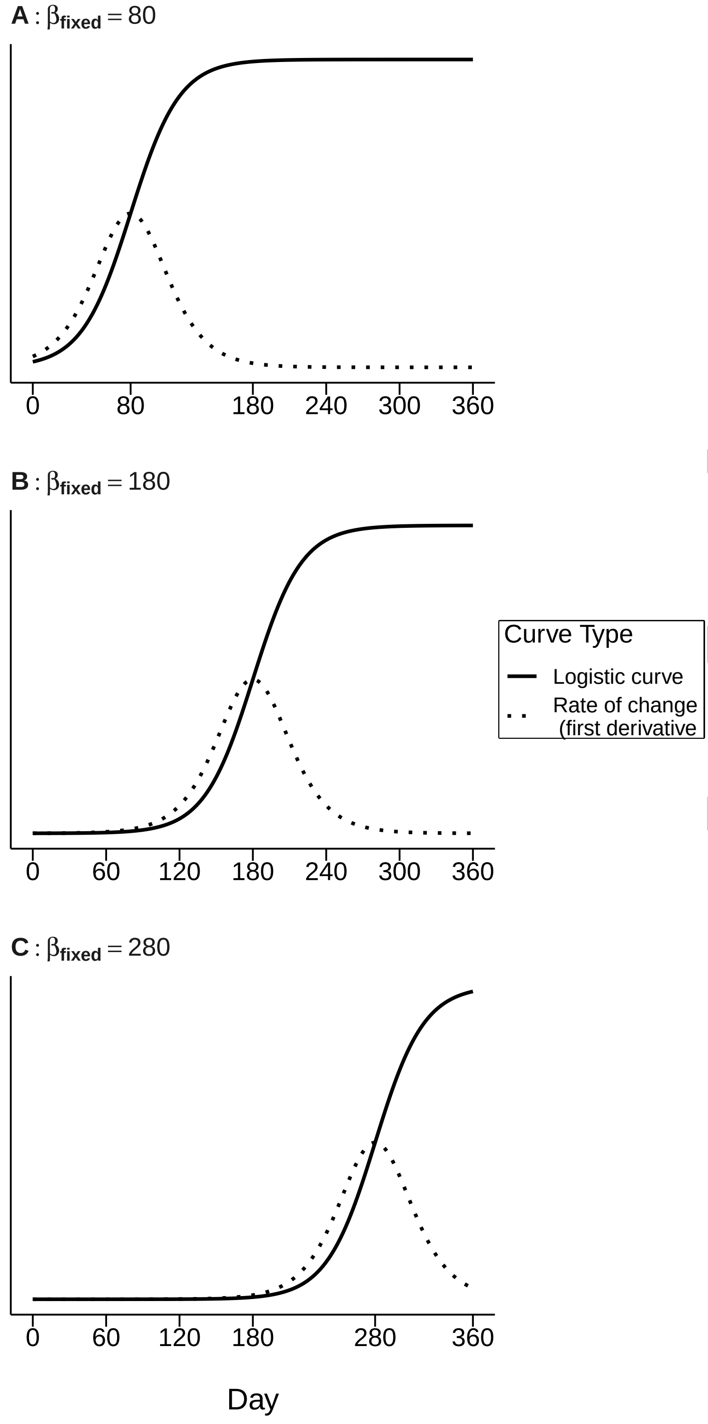


1253 *Note.* Panel A: Measurement sampling locations on each manipulated nature-of-change curve under equal
1254 spacing. Panel B: Measurement sampling locations on each manipulated nature-of-change curve under
1255 time-interval increasing spacing. Panel C: Measurement sampling locations on each manipulated
1256 nature-of-change curve under time-interval decreasing spacing. Panel D: Measurement sampling locations
1257 on each manipulated nature-of-change curve under middle-and-extreme spacing. Black curves indicate the
1258 natures of change that lead to the highest modelling accuracy for each spacing schedule, and so are
1259 optimal. Gray curves indicate the natures of change that lead to suboptimal modelling accuracy for each
1260 spacing schedule, and so are not optimal. Text on each panel indicates the error bar lengths when
1261 modelling accuracy is highest (see Table 2.15).

Before explaining why each spacing schedule has an optimal curve, it is important to define change. For the purpose of this discussion, change occurs when the first derivative of the logistic function has a nonzero value, with larger absolute first derivative values implying greater change. Figure 2.9 shows each nature of change used in Experiment 1 (solid line) along with its corresponding first derivative curve (dotted line). For each nature of change, the first derivative value reaches its peak at the value set for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). In Figure 2.9A, the first derivative is greatest at day 80. In Figure 2.9B, the first derivative is greatest at day 180. In Figure 2.9C, the first derivative is greatest at day 280. Therefore, for each manipulated nature of change, change is greatest at the value set for the fixed-effect days-to-halfway elevation parameter (β_{fixed}).

Figure 2.9

Rate of Change (First Derivative Curve) for Each Nature of Change Curve Manipulated in Experiment 1



Note. Panel A: Logistic curve defined by β_{fixed} = 80, with first-derivative curve peaking at day 80. Panel B: Logistic curve defined by β_{fixed} = 180, with first-derivative curve peaking at day 180. Panel C: Logistic curve defined by β_{fixed} = 280, with first-derivative curve peaking at day 280.

Figure 2.8 provides three reasons to suggest that sampling measurements closer to the period of greatest change increases modelling accuracy. First, for each spacing schedule, more measurements lie closer to the area of greatest change on the optimal black curve than on the suboptimal gray curves. One clear example can be observed for the measurement locations under middle-and-extreme spacing (see Figure 2.8D). In looking across the nature-of-change curves, only the measurement locations of the middle three measurements on each curve are different. For the optimal black nature of change, the middle three measurements are centered on the period of greatest change. For the gray suboptimal nature-of-change curves, the middle three measurements are taken near regions of little change (near-zero first derivative). Therefore, nature-of-change curves are optimal for their respective spacing schedules because measurements are taken closest to the curves' periods of change.

Second, modelling accuracy under time-interval increasing and decreasing spacing is nearly identical because each spacing schedule samples data at the exact same regions of change. In looking at Table 2.15, it is important to realize that the precision values (i.e., error bar lengths) obtained with time-interval increasing and decreasing spacing are nearly identical when modeling accuracy is highest. As an example, the average error bar length obtained for the fixed-effect days-to-halfway elevation parameter (β_{fixed}) is 5.80 days with time-interval increasing spacing and a nature-of-change value of 80 (i.e., early halfway point) and 5.84 days with time-interval decreasing spacing and a nature-of-change value of 280 (i.e., late halfway point). The nearly equivalent precision obtained with time-interval increasing and decreasing spacing occurs because the rates of change (i.e., first derivative values) at the sampled locations are the exact same. Table **??** lists the curve values and measurement days when the time-interval increasing and decreasing spacing schedules sample the same first-derivative values. Note that, because the time-interval increasing and decreasing spacing schedules sample data in opposite orders, the first-derivative values are sampled in opposite orders. In summary, although the time-interval increasing and decreasing spacing schedules sample data on different days on their respective optimal curves, they result in (nearly) identical modelling accuracy because the first-derivative values of the optimal curves at the sampled locations are the exact same.

Table 2.16
Identical First-Derivative Sampling of Time-Interval Increasing and Decreasing Spacing Schedules

	Time-Interval Increasing		Time-Interval Decreasing	
First Derivative Value	Curve Value	Measurement Day	Curve Value	Measurement Day
2.00e-06	3.00	0	3.32	360
8.80e-06	3.00	30	3.32	330
2.83e-04	3.01	100	3.31	260
2.39e-03	3.26	210	3.06	150
2.00e-06	3.32	360	3.00	0

Third, middle-and-extreme spacing obtains higher modelling accuracy than equal spacing by sampling data at periods of greater change. Importantly, both equal and middle-and-extreme spacing obtain their highest modelling accuracy with a curve whose greatest change occurs at 180 days (i.e., midway halfway point), with middle-and-extreme spacing obtaining higher precision (i.e., shorter error bars) than equal spacing (see Figure 2.8 and Table 2.15). An inspection of Figures 2.8A and 2.8D reveals that middle-and-extreme spacing samples measurements at moments of greater change. As an example, consider the measurement locations of equal and middle-and-extreme spacing with five measurements, where only second and fourth measurement locations differ between the schedules. For equal spacing, the second and fourth measurements are respectively sampled on days 90 and 270. For middle-and-extreme spacing, the second and fourth measurements are respectively taken on days 150 and 210. By consulting the first-derivative curve in Figure **??**, change is greater on days 150 and 210 than on days 90 and 270. Therefore, precision across all manipulated measurement numbers is greater (i.e., shorter error bars) with middle-and-extreme spacing than with equal spacing because middle-and-extreme spacing takes measurements closer to periods of change than equal spacing (see Figures 2.8A and 2.8D and Table 2.15).

The idea that modelling accuracy increases when data are sampled during periods of greater change has received considerable discussion and preliminary support. Over the past 20 years, researchers have recommended that measurements be sampled during periods of greater change (Ployhart & Vandenberg, 2010; Siegler, 2013), with one recent simulation study finding evidence support this idea (Timmons & Preacher, 2015). Unfortunately, the evidence from Timmons & Preacher (2015) is preliminary for two reasons. First, the model used to estimate nonlinear change only ever included one random-effect parameter. Given that multilevel models often include several random-effect parameter in practice, the model employed in Timmons & Preacher (2015) may not necessary be realistic. Second, the estimates were obtained by using an impractical starting value procedure: Population values were used as starting values. Because practitioners never know the population value, it is not known whether the results of Timmons & Preacher

(2015) replicate with a realistic starting value procedure.

My simulations in Experiment 1 replicated the finding that modelling accuracy increases from measuring change near periods of change under more realistic conditions. In contrast to the one-random-effect-parameter models used in Timmons & Preacher (2015), my simulations used a four-parameter model where each parameter was modelled as a fixed and random effect. For the starting value procedure, my simulations did not use the population values as starting values, but used the starting value procedure available in OpenMx.

Therefore, three results in Experiment 1 suggest that sampling data closer to periods of change leads to higher modelling accuracy. First, for each spacing schedule, modelling accuracy is highest when measurements are taken closer to periods of change. Second, the time-interval increasing and decreasing spacing schedules obtain nearly identical modelling accuracies for different curves because the sampled locations have the exact same rates of change. Third, middle-and-extreme spacing results in higher modelling accuracy than equal spacing by sampling measurements at periods of greater change. Although several researchers have posited modelling accuracy increases by sampling data closer to periods of change, with one simulation study (to my knowledege) having confirmed this notion under unrealistic modelling conditions, my simulations in Experiment 1 confirm it under realistic modelling conditions.

2.2.7.2 When the Nature of Change is Unknown, How Should Measurements be Spaced?

Table 2.17 provides a summary of the results for each spacing schedule. Text within the 'Unbiased' column indicates the number of measurements needed to obtain unbiased estimation of all the day-unit parameters across all manipulated nature-of-change values for each spacing schedule. Text within the 'Qualitative Description' column indicates the number of measurements that obtains the largest improvements in bias and precision across all manipulated nature-of-change values for each spacing schedule. The 'Error Bar Summary' columns list the error bar lengths obtained for each day-unit parameter of the logistic function using the measurement number listed in the 'Qualitative Description' column. Importantly, the error bar lengths in the 'Error Bar Summary' column are obtained by computing the average length across all manipulated nature-of-change values for the measurement number listed Qualitative Description' column. The following number of measurements are needed to obtain unbiased estimation and the greatest improvements in bias and precision across all manipulated nature-of-change values for all day-unit parameters under each spacing schedule:

- equal spacing: nine or more measurements to obtain unbiased estimation and seven measurements to obtain the greatest improvements in bias and precision.
- time-interval increasing spacing: nine or more measurements to obtain unbiased estimation and nine measurements to obtain the greatest improvements in bias and precision.
- time-interval decreasing spacing: nine or more measurements to obtain unbiased estimation and nine measurements to obtain the greatest improvements in bias and precision.
- middle-and-extreme spacing: 11 measurements to obtain unbiased estimation and nine measurements to obtain the greatest improvements in bias and precision.

Table 2.17

Concise Summary of Results Across All Spacing Schedule Levels in Experiment 1

Spacing Schedule	Unbiased	Precise	Qualitative Description	Error Bar Summary			
				β_{fixed}	γ_{fixed}	β_{random}	γ_{random}
Equal (see Figure 2.4 and Table 2.3)	NM $\geq$ 9	No cells	Largest improvements in bias and precision with NM = 7	5.64	4.37	7.74	7.02
Time-interval increasing (see Figure 2.5 and Table 2.6)	NM $\geq$ 9	No cells	Largest improvements in bias and precision with NM = 9	4.97	3.45	6.31	5.97
Time-interval decreasing (see Figure 2.6 and Table 2.9)	NM $\geq$ 9	No cells	Largest improvements in bias and precision with NM = 9	4.88	3.40	6.15	5.96
Middle-and-extreme (see Figure 2.7 and Table 2.9)	NM = 11	No cells	Largest improvements in bias and precision with NM = 9	6.51	5.55	9.02	7.20

Note. Row shaded in gray indicates the spacing schedules that results in the highest modelling accurac across all manipulated nature-of-change curves. 'Qualitative Description' column indicates the number of measurements that obtains the greatest improvements in bias and precision across all day-unit parameters and manipulated nature-of-change values. 'Error Bar Summary' columns list the error bar lengths that result for each day-unit parameter using the measurement number listed in the 'Qualitative Description' column. Note that error bar lengths were calculated by computing the average length across all manipulated measurement numbers for the nature-of-change value listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter $\in \{80, 180, 280\}$; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

An important point to mention is that the error bar lengths for each day-unit parameter across each spacing schedule are comparable. That is, each spacing schedule obtains similar modelling accuracy when using the number of measurements listed in the 'Qualitative Description' column. Because modelling accuracy is similar across the spacing schedules, then the schedule that requires the fewest number of measurements to obtain the greatest improvements in bias and precision models change most accurately when the nature of change is unknown. With equal spacing using fewer measurements than all the other manipulated spacing schedules to obtain similar modelling accuracy—using seven measurements instead of the nine measurements use by all other spacing schedules—equal spacing is the most effective schedule to use when the nature of change is unknown.

The finding that equal spacing results in the highest modelling accuracy when the nature of change is unknown is not unexpected. Given the previous finding that modelling accuracy increases by sampling data closer to periods of change, then, if the nature of change is unknown, change may occur at any point in time, and so it is prudent to space measurements equally over time.

2.3 Summary of Experiment 1

I designed Experiment 1 to investigate two questions. The first question was how to space measurements when the nature of change is suspected. For each spacing schedule, modelling accuracy was highest when measurements were sampled at periods of greaterc change. Therefore, when the nature of change is suspected, measurements should be taken near periods of change to increase modelling accuracy.

The second question was how to space measurements when the nature of change is unknown. Although no manipulated measurement number under any spacing schedule resulted in accurate modelling of all parameters, the improvements in modelling accuracy began to diminish under each spacing schedule at a specific measurement number. Given that each spacing schedule obtained comparable modelling accuracy when it began to diminish, I concluded that the spacing schedule that used the fewest number of measurements was most effective at modelling change when the nature of change was unknown. With equal spacing using the fewest number of measurements to obtain the greatest improvements in modelling accuracy, equal spacing was the most effective schedule to use when the nature of change was unknown.

3 Experiment 2

In Experiment 2, I investigated the combinations of measurement number and sample size needed to achieve accurate modelling (i.e., unbiased and precise parameter estimation) under different spacing schedules. Before presenting the results of Experiment 2, I will present my design and and analysis goals. For the design, I conducted a 4 (spacing schedule: equal, time-interval increasing, time-interval decreasing, middle-and-extreme) x 4(number of measurements: 5, 7, 9, 11) x 6(sample size: 30, 50, 100, 200, 500, 1000) study. For the analysis, I was interested in determining, for each spacing schedule, the combinations of number of measurements and sample size that achieved accurate modelling (i.e., unbiased and precise parameter estimation). For parsimony, I present the sample size by number of measurements results for each level of spacing schedule.

3.1 Methods

3.1.1 Variables Used in Simulation Experiment

3.1.1.1 Independent Variables

3.1.1.1.1 Spacing of Measurements

For the spacing of measurements, I used the same values as in Experiment 1 of equal, time-interval increasing, time-interval decreasing, and middle-and-extreme spacing (see spacing of measurements for more discussion).

3.1.1.1.2 Number of Measurements

For the number of measurements, I used the same values as in Experiment 1 of 5, 7, 9, and 11 measurements (see number of measurements} for more discussion)).

3.1.1.1.3 Sample Size

Sample size values were borrowed from Coulombe et al. (2016) with one difference. Because my experiments investigated the effects of measurement timing factors on the ability to model nonlinear patterns, which are inherently more complex than linear patterns of change, a sample size value of $N = 1000$ was added as the largest sample size. Therefore, the following values were used for my sample size manipulation: 30, 50, 100, 200, 500, and 1000.

3.1.1.2 Constants

Because the nature of change not manipulated in Experiment 2, I set it to have a constant value across all cells. To keep the nature of change constant across all cells, I set the fixed-effect days-to-halfway elevation parameter (β_{fixed}) to have a value of 180. Another variable set to a constant value across the cells was time structuredness (data were assumed to be time structured). That is, data were generated such that, at each time point, all data were obtained at the exact same time.

3.1.1.3 Dependent Variables

3.1.1.3.1 Convergence Success Rate

The proportion of iterations in a cell where models converged defined the **convergence success rate**.[16] Equation (4.5) below shows the calculation used to compute the convergence success rate:

$$\text{Convergence success rate} = \frac{\text{Number of models that successfully converged in a cell}}{n}, \tag{3.1}$$

where n represents the total number of models run in a cell.

3.1.1.3.2 Bias

Bias was calculated to evaluate the accuracy with which each logistic function parameter was estimated. As shown below in Equation (4.6), *bias* was obtained by calculating the difference between the population value set for a parameter and the average estimated value in each cell.

$$\text{Bias} = \text{Population value for parameter} - \text{Average estimated value} \tag{3.2}$$

Bias was calculated for the fixed- and random-effect parameters of the baseline (θ_{fixed}, θ_{random}), maximal elevation (α_{fixed}, α_{random}), days-to-halfway elevation (β_{fixed}, β_{random}), and the halfway-triquarter delta parameters (γ_{fixed}, γ_{random}).

[16]Specifically, convergence was obtained if the convergence code returned by OpenMx was 0.

3.1.1.3.3 Precision

In addition to computing bias, precision was calculated to evaluate the confidence with which each parameter was estimated in a given cell. *Precision* was obtained by computing the range of values covered by the middle 95% of values estimated for a logistic parameter in each cell. By using the middle 95% of estimated values, a plausible range of population estimates was obtained.

3.1.2 Overview of Data Generation

Data generation was computed the same way as in Experiment 1 (see data generation).

3.1.3 Modelling of Each Generated Data Set

Each generated data set was modelled using the structured latent growth curves outlined in Experiment 1 (see data modelling. For a detailed explanation of how the logistic function was fit into the structural equation modelling framework, see Technical Appendix B.

3.1.4 Analysis of Data Modelling Output and Accompanying Visualizations

Analysis and visualization was conducted as outlined in Experiment 1 (see analysis and visualization).

3.2 Results and Discussion

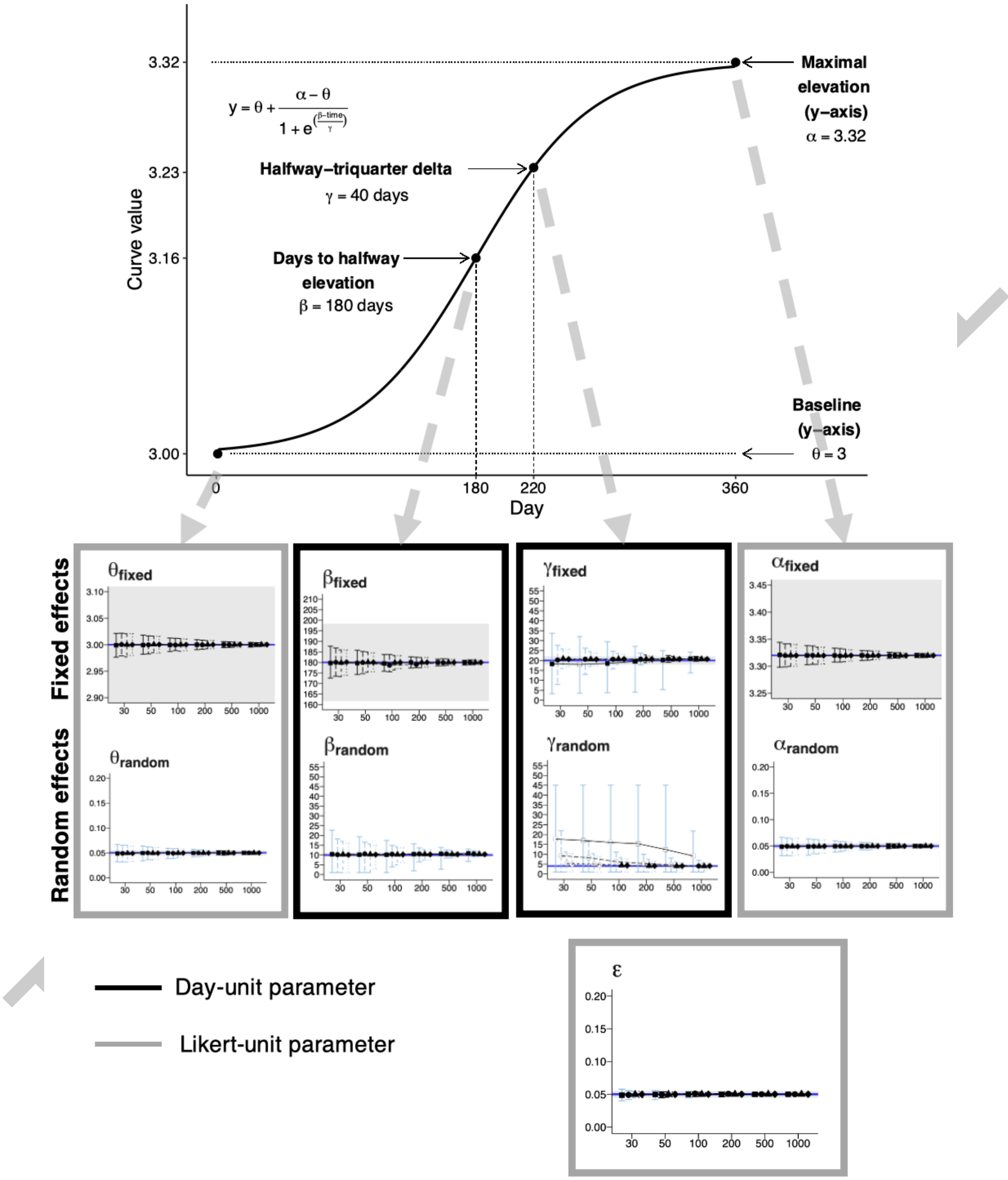
In the sections that follow, I organize the results by presenting them for each spacing schedule (equal, time-interval increasing, time-interval decreasing, middle-and-extreme). Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and halfway-triquarter delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix C).

3.2.1 Framework for Interpreting Results

Because Experiment 2 (like all other experiments) had many cells (i.e., 96 cells in Experiment 2), the number of dependent variables to track in the results section can become overwhelming. Therefore, I will provide a framework to help the reader efficiently navigate the results section. Figure 3.1 shows the entire set of results that will be presented for each spacing schedule. For each spacing schedule, a parameter estimation plot is created for each of the nine parameters estimated by the structured latent growth curve model used on each generated data set (for a review, see modelling of each generated data set). Parameter estimation plots with black outlines show the results for day-unit parameter and plots with gray outlines show the results for likert-unit parameters. Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and halfway-triquarter delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix C.

Figure 3.1

Set of Parameter Estimation Plots Constructed for Each Spacing Schedule in Experiment 2



Note. A parameter estimation plot is constructed for each parameter of the logistic function (see Equation 2.3). Note that each parameter of the logistic function is modelled as a fixed and random effect along with an error term (ϵ; for a review, see Figure 1.4).

3.2.2 Pre-Processing of Data and Model Convergence

After collecting the output from the simulations, non-converged models (and their corresponding parameter estimates) were removed from subsequent analyses. Table D.2 in Appendix D provides the convergence success rates for each cell in Experiment 2. Model convergence was almost always above 90% and convergence rates rates below 90% only occurred in two cells with five measurements.

3.2.3 Equal Spacing

For equal spacing, Table 3.1 provides a concise summary of the results for the day-unit parameters (see Figure 3.2 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 3.1 and provide elaboration when necessary.

Before presenting the results for equal spacing, I provide a brief description of the concise summary table created for each spacing schedule and shown for equal spacing in Table 3.1. ext in the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number-sample size pairing needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with equal spacing). The 'Error Bar Length' column indicates the error bar length that results from using the lower-bounding measurement number-sample size pairing listed in the 'Qualitative Description' column.

Table 3.1

Concise Summary of Results for Equal Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 3.2A)	All cells	All cells	Unbiased and precise estimation in all cells	15.13
γ_{fixed} (Figure 3.2B)	All cells	NM $\geq$ 9 with $N \geq 500$	Largest improvements in precision using **NM = 7 with *N* $\geq$ 200** or **NM = 9 with *N* $\leq$ 100**	9.79
β_{random} (Figure 3.2C)	All cells	No cells	Largest improvements in precision with NM = 7	17.22
γ_{random} (Figure 3.2D)	**NM $\geq$ 7 with *N* = 1000** or **NM $\geq$ 9 with *N* $\geq$ 200** or **NM = 11 with *N* = 100**	No cells	Largest improvements in bias and precision using NM = 7 with *N* $\geq$ 100 or NM = 9 with *N* $\leq$ 50	10.08

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with equal spacing). 'Error Bar Length' column indicates the maximum error bar length that results from using the measurement number-sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

3.2.3.1 Bias

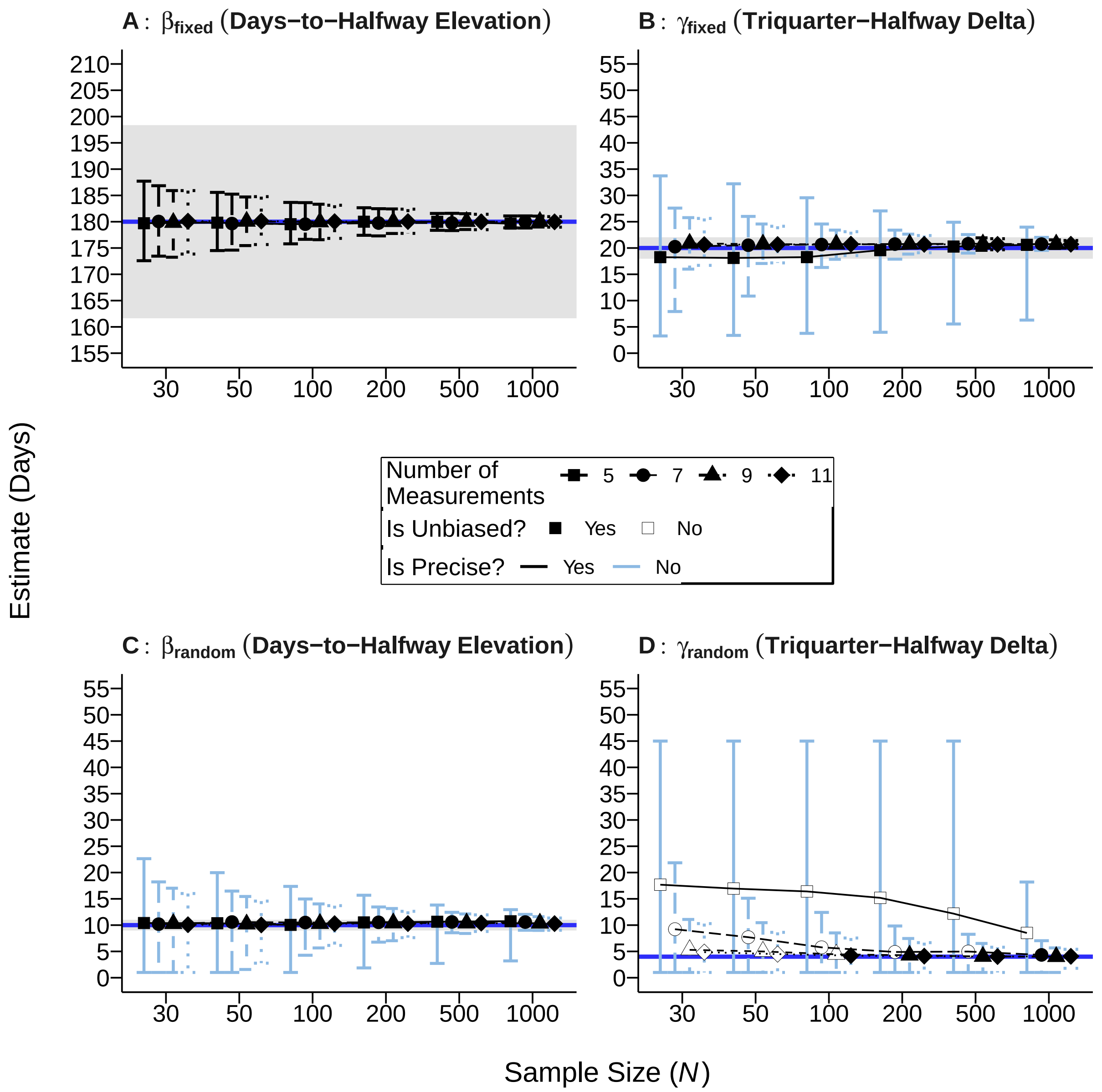
Before presenting the results for bias, I provide a description of the set of parameter estimation plots shown in Figure 3.2 and in the results sections for the other spacing schedules in Experiment 2. Figure 3.2 shows the parameter estimation plots for each day-unit parameter and Table 3.2 provides the partial ω^2 values for each independent variable of each day-unit parameter. In Figure 3.2, blue horizontal lines indicate the population values for each parameter (with population values of β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, and γ_{random} = 4.00). Gray bands indicate the ±10% margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Panels A–B show the parameter estimation plots for the fixed- and random-effect days-to-halfway elevation parameters (β_{fixed} and β_{random}, respectively). Panels C–D show the parameter estimation plots for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively). Note that random-effect parameter units are in standard deviation units.

With respect to bias for equal spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.2A): no cells.

Figure 3.2

Parameter Estimation Plots for Day-Unit Parameters With Equal Spacing in Experiment 2



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that

random-effect parameter units are in standard deviation units. See Table ?? for specific values estimated for each parameter and Table 3.2 for ω^2 effect size values.

Table 3.2

Partial ω^2 Values for Independent Variables With Equal Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.2A)	0.00	0.03	0.00
β_{random} (Figure 3.2B)	0.15	0.28	0.03
γ_{fixed} (Figure 3.2C)	0.31	0.15	0.09
γ_{random} (Figure 3.2D)	0.18	0.03	0.01

Note .NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.2B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.2C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.2D): five measurements across all sample sizes, seven measurements with $N \leq 500$, nine measurements with $N \leq 100$, and 11 measurements with $N \leq 50$.

In summary, with equal spacing, estimation of all the day-unit parameters across all manipulated nature-of-change values is unbiased using nine measurements with $N \geq 200$ or 11 measurements with $N = 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 3.1.

3.2.3.2 Precision

With respect to precision for equal spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.2A): all cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.2B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.2C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 3.2D): all cells.

In summary, with equal spacing, precise estimation can be obtained for the fixed-effect day-unit parameters using at least nine measurements with $N \geq 500$, but no manipulated measurement number-sample size pairing results in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 3.1).

3.2.3.3 Qualitative Description

For equal spacing in Figure 3.2, although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurement number-sample size pairings. With respect to bias under equal spacing, the largest improvements in bias result with the following measurement number-sample size pairings for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively):

- fixed-effect triquarter-halfway delta parameters (γ_{fixed}): seven measurements with $N = 30$.
- random-effect triquarter-halfway delta parameters (γ_{random}): seven measurements

with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under equal spacing, the largest improvements in precision in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) result from using the following measurement number-sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 9.79 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which results in a error bar length of 17.22 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$, which results in a maximum error bar length of 10.08 days.

For an applied researcher, one plausible question might be what measurement number-sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters. In looking across the measurement number-sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters result with the following measurement number-sample size pairing(s): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 3.1).

3.2.3.4 Summary of Results

In summarizing the results for equal spacing, estimation of all day-unit parameters is unbiased using nine measurements with $N \geq 200$ or 11 measurements with $N = 1000$ (see the emboldened text in in the 'Unbiased' column of Table **??**). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement number-sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number-sample size pairing under equal spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement number-sample size pairings. With equal spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters are obtained from using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 3.1).

3.2.4 Time-Interval Increasing Spacing

For time-interval increasing spacing, Table 3.3 provides a concise summary of the results for the day-unit parameters (see Figure 3.3 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 3.3 and provide elaboration when necessary (for a description of Table 3.3, see concise summary table).

Table 3.3

Concise Summary of Results for Time-Interval Increasing Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 3.3A)	All cells	All cells except NM = 5 with N $\leq$ 200	Largest improvements in precision using NM = 7 across all sample sizes	16.77
γ_{fixed} (Figure 3.3B)	All cells	NM $\geq$ 7 with N = 1000 or NM $\geq$ 9 with N = 1000	Largest improvements in precision using **NM = 7 with *N* $\geq$ 200** or NM = 9 with $N \leq$ 100	9.69
β_{random} (Figure 3.3C)	All cells except	No cells	Largest improvements in precision using NM = 7 across all sample sizes	17.85
γ_{random} (Figure 3.3D)	**NM $\geq$ 9 with *N* $\geq$ 200** or **NM = 11 with *N* = 1000**	No cells	Largest improvements in bias and precision using **NM = 5 with *N* $\geq$ 500** or **NM = 9 with *N* $\leq$ 200**	10.15

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with time-interval increasing spacing). 'Error Bar Length' column indicates the maximum error bar length that results from using the measurement number-sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

3.2.4.0.1 Bias

With respect to bias for time-interval increasing spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

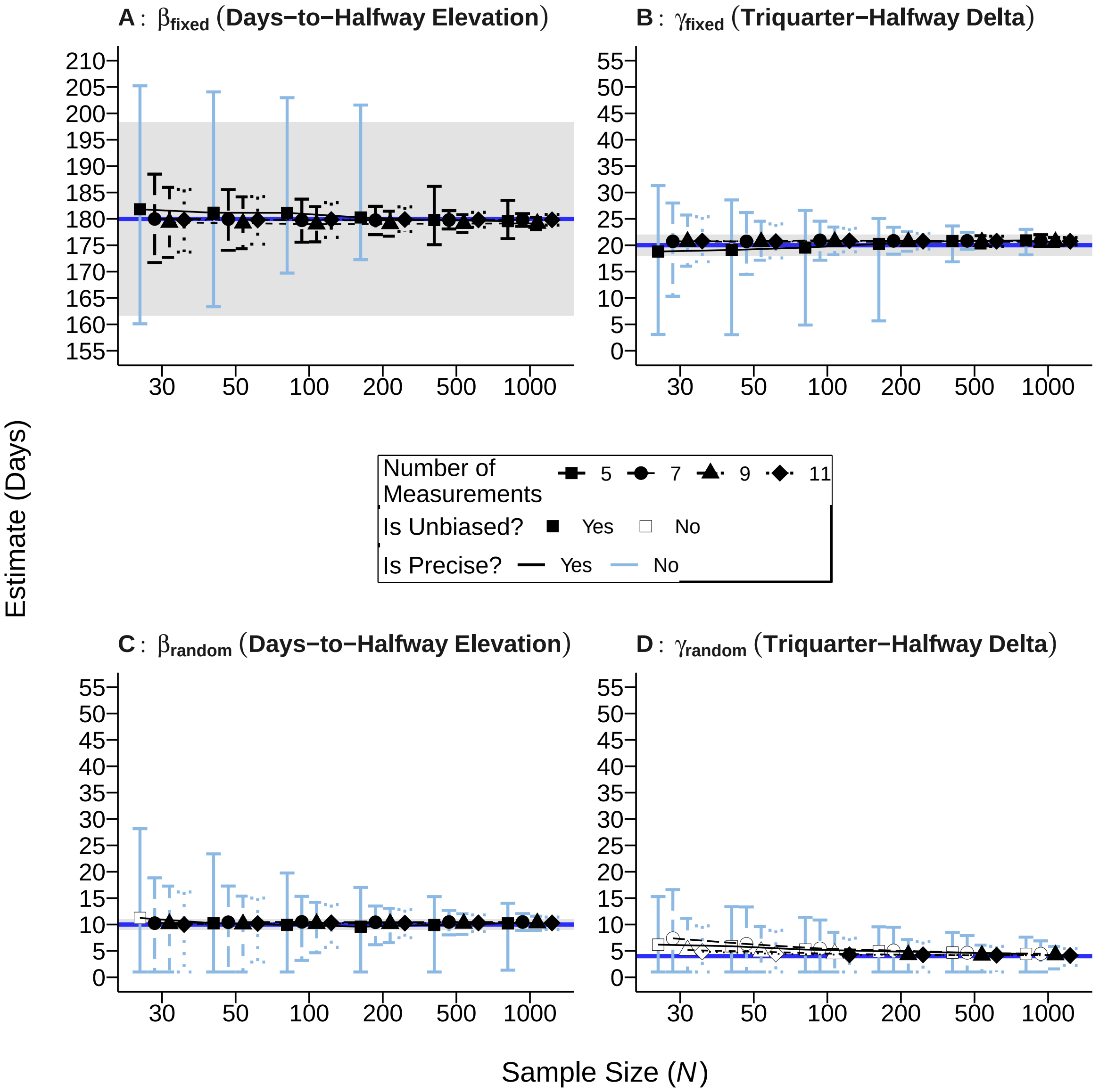
- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.3A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.3B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.3C): NM = 5 with $N = 30$.

- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.3D): five an seven measurements across all sample sizes, nine measurements with $N \leq 100$, and 11 measurements with $N \leq 50$.

In summary, with time-interval increasing spacing, estimation of all the day-unit parameters is unbiased using nine measurements with $N \geq 200$ or 11 measurements with $N = 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 3.3.

Figure 3.3

Parameter Estimation Plots for Day-Unit Parameters With Time-Interval Increasing Spacing in Experiment 2



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that

random-effect parameter units are in standard deviation units. See Table ?? for specific values estimated for each parameter and Table 3.4 for ω^2 effect size values.

Table 3.4
Partial ω^2 Values for Independent Variables With Time-Interval Increasing Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.3A)	0.23	0.15	0.09
β_{random} (Figure 3.3B)	0.15	0.16	0.02
γ_{fixed} (Figure 3.3C)	0.17	0.16	0.07
γ_{random} (Figure 3.3D)	0.07	0.12	0.01

Note. NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

3.2.4.0.2 Precision

With respect to precision for time-interval increasing spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.3A): five measurements with $N \leq 100$.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.3B): five and measurements across all sample sizes, seven measurements with $N \leq 500$, nine and 11 measurements with $N \leq 200$.

- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.3C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 3.3D): all cells.

In summary, with time-interval increasing spacing, precise estimation can be obtained for the fixed-effect day-unit parameters using at least nine measurements with $N \geq 500$, but no manipulated measurement number-sample size pairing results in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 3.3).

3.2.4.0.3 Qualitative Description

For time-interval increasing spacing in Figure 3.3, although no manipulated measurement number-sample size pairing results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurement number-sample size pairings. With respect to bias under time-interval increasing spacing, the largest improvements in bias result with the following measurement number-sample size pairings for random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively):

- random-effect triquarter-halfway delta parameters (γ_{random}): five measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-interval increasing spacing, the largest improvements in precision in the estimation of each day-unit parameter result from using the following measurement number-sample size pairings:

- days-to-halfway elevation parameter (β_{fixed}): seven measurements with $N \geq 30$, which results in a maximum error bar length of 9.69 days.
- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error

bar length of 9.69 days.

- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which results in a error bar length of 17.85 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): five measurements with $N \geq 500$ or nine measurements with $N \leq 200$, which results in a maximum error bar length of 10.15 days.

For an applied researcher, one plausible question might be what measurement number-sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters when using time-interval increasing spacing. In looking across the measurement number-sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters with time-interval increasing spacing result from using the following measurement number-sample size pairing(s): five measurements with $N \geq 500$ or nine measurements with $N \leq 200$ (see the emboldened text in the 'Qualitative Description' column of Table 3.3).

3.2.4.1 Summary of Results

In summarizing the results for time-interval increasing spacing, estimation of all day-unit parameters is unbiased nine measurements with $N \geq 200$ or 11 measurements with $N = 100$ (see bias). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement number-sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number-sample size pairing under time-interval increasing spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement number-sample size pairings. With time-interval increasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters are obtained from using five measurements with $N \geq 500$ or nine measurements with $N \leq 200$ (see qualitative description).

3.2.5 Time-Interval Decreasing Spacing

For time-interval decreasing spacing, Table 3.5 provides a concise summary of the results for the day-unit parameters (see Figure 3.4 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 3.5 and provide elaboration when necessary (for a description of Table 3.5, see concise summary table).

Table 3.5

Concise Summary of Results for Time-Interval Decreasing Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 3.4A)	All cells	All cells except NM = 5 with $N \leq 500$	Largest improvements in precision using NM = 7 across all sample sizes	17.42
γ_{fixed} (Figure 3.4B)	All cells	NM = 7 with N = 1000 or NM $\geq$ 9 with $N \geq 500$	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or NM = 9 with $N \leq 100$	9.62
β_{random} (Figure 3.4C)	All cells except NM = 5 with N = 50	No cells	Largest improvements in precision using NM = 7 across all sample sizes	17.44
γ_{random} (Figure 3.4D)	**NM = 11 with $N \geq$ 100**	No cells	Largest improvements in bias and precision using **NM = 5 with $N \geq$ 500** or **NM = 9 with $N \leq$ 200**	10.32

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with time-interval decreasing spacing). 'Error Bar Length' column indicates the maximum error bar length that results from using the measurement number-sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

3.2.5.1 Bias

With respect to bias for time-interval decreasing spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

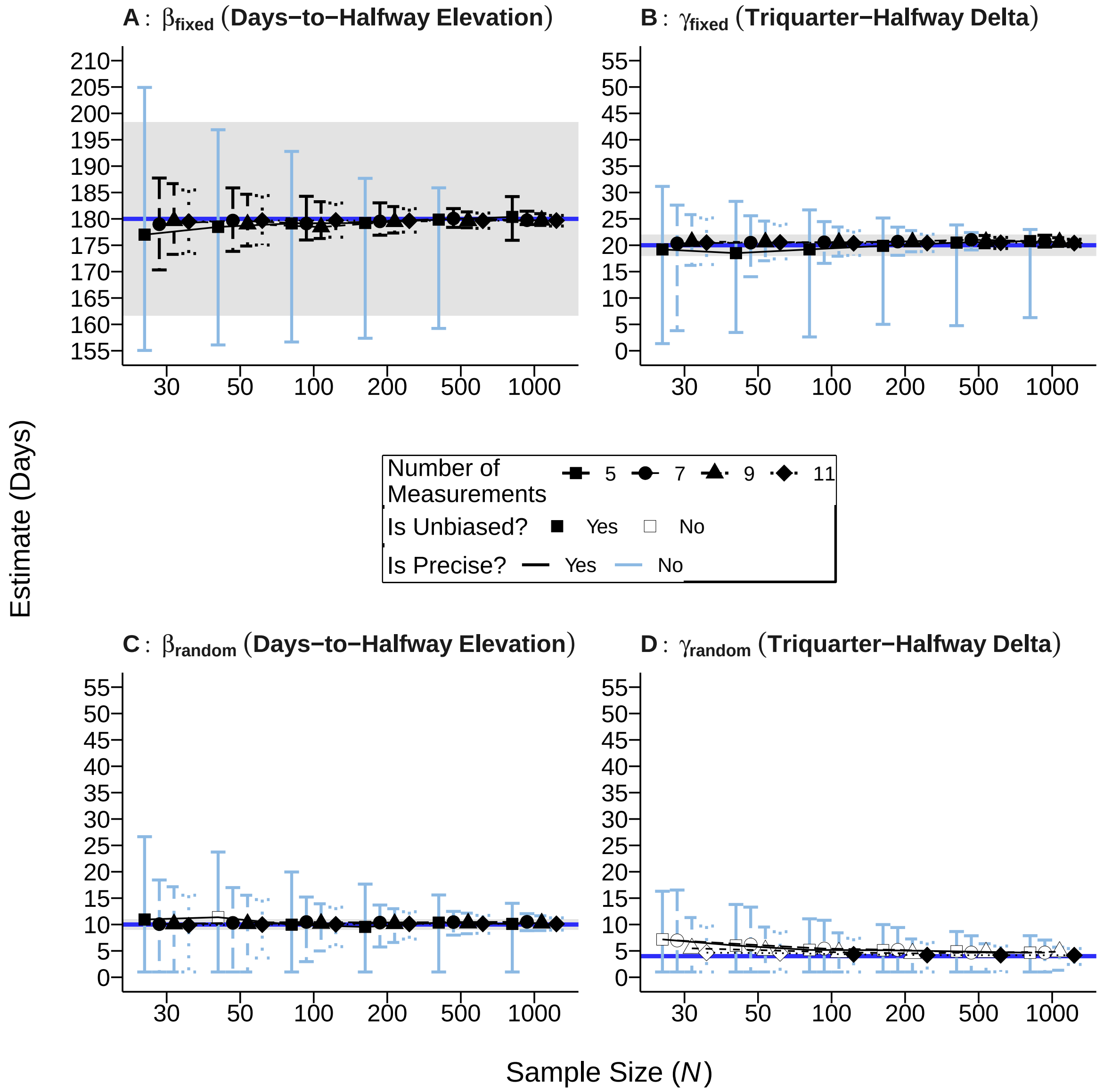
- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.4B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): NM = 5 with $N = 30$.

- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.4D): five, seven, and nine measurements across all sample sizes an 11 measurements with $N \leq 50$, and 11 measurements with $N \leq 50$.

In summary, with time-interval decreasing spacing, estimation of all the day-unit parameters is unbiased using 11 measurements with $N \geq 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 3.5.

Figure 3.4

Parameter Estimation Plots for Day-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 2



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that

random-effect parameter units are in standard deviation units. See Table ?? for specific values estimated for each parameter and Table 3.6 for ω^2 effect size values.

Table 3.6

Partial ω^2 Values for Independent Variables With Time-Interval Decreasing Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.4A)	0.05	0.03	0.01
β_{random} (Figure 3.4B)	0.14	0.12	0.01
γ_{fixed} (Figure 3.4C)	0.07	0.04	0.01
γ_{random} (Figure 3.4D)	0.05	0.09	0.00

Note. NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

3.2.5.2 Precision

With respect to precision for time-interval decreasing spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): five measurements with $N \leq 500$.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.4B): five measurements across all sample sizes, seven measurements with $N \leq 500$, and nine and 11 measurements with $N \leq 200$.

- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 3.4D): all cells.

In summary, with time-interval decreasing spacing, precise estimation can be obtained for the fixed-effect day-unit parameters using at least seven measurements with $N = 1000$ or nine measurements $N \leq 500$. For the random-effect day-unit parameters, no manipulated measurement number-sample size pairing results in precise estimation (see the 'Precise' column of Table 3.5).

3.2.5.3 Qualitative Description

For time-interval decreasing spacing in Figure 3.4, although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurement number-sample size pairings. With respect to bias under time-interval decreasing spacing, the largest improvements in bias result with the following measurement number-sample size pairings for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively):

- random-effect triquarter-halfway delta parameters (γ_{random}): five measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-interval decreasing spacing, the largest improvements in precision in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) result from using the following measurement number-sample size pairings:

- days-to-halfway elevation parameter (β_{fixed}): seven measurements with $N \geq 30$, which results in a maximum error bar length of 9.62 days.

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 9.62 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which results in a error bar length of 17.44 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): five measurements with $N \geq 500$ or nine measurements with $N \leq 200$, which results in a maximum error bar length of 10.32 days.

For an applied researcher, one plausible question might be what measurement number-sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-interval decreasing spacing. In looking across the measurement number-sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters with time-interval decreasing spacing result with the following measurement number-sample size pairing(s): five measurements with $N \geq 500$, seven measurements with $N \geq 200$, or nine measurements with $N \leq 200$ (see the emboldened text in the 'Qualitative Description' column of Table 3.5).

3.2.5.4 Summary of Results

In summarizing the results for time-interval decreasing spacing, estimation of all day-unit parameters is unbiased 11 measurements with $N \geq 10$ (see bias). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement number-sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number-sample size pairing under time-interval decreasing spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement number-sample size pairings. With time-interval decreasing spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters are obtained from using five measurements with $N \geq 500$ or nine measurements with $N \leq 200$ (see qualitative description).

3.2.6 Middle-and-Extreme Spacing

For middle-and-extreme spacing, Table 3.7 provides a concise summary of the results for the day-unit parameters (see Figure 3.5 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 3.7 and provide elaboration when necessary (for a description of Table 3.7, see concise summary table).

Table 3.7

Concise Summary of Results for Middle-and-Extreme Spacing in Experiment 2

			Description	
Parameter	Unbiased	Precise	Qualitative Descroption	Error Bar Length
β_{fixed} (Figure 3.5A)	All cells	All cells	Largest improvements in precision using using **NM = 5**	14.96
γ_{fixed} (Figure 3.5B)	All cells	All number of measurements with $N \geq 500$	Largest improvements in precision using **NM = 5**	9.92
β_{random} (Figure 3.5C)	All cells	No cells	Largest improvements in precision using **NM = 5**	15.94
γ_{random} (Figure 3.5D)	**NM** $\in$ **{5, 9} with *N*** $\geq$ **100** or **NM** $\in$ **{7, 11} with *N*** $\leq$ **50**	No cells	Largest improvements in precision using **NM = 5**	10.13

Note. Text in the ‘Unbiased’ and ‘Precise’ columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the ‘Unbiased’ and ‘Qualitative Description’ columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with middle-and-extreme spacing). ‘Error Bar Length’ column indicates the maximum error bar length that results from using the measurement number-sample size recommendation listed in the ‘Qualitative Description’ column Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

3.2.6.0.1 Bias

With respect to bias for middle-and-extreme spacing, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

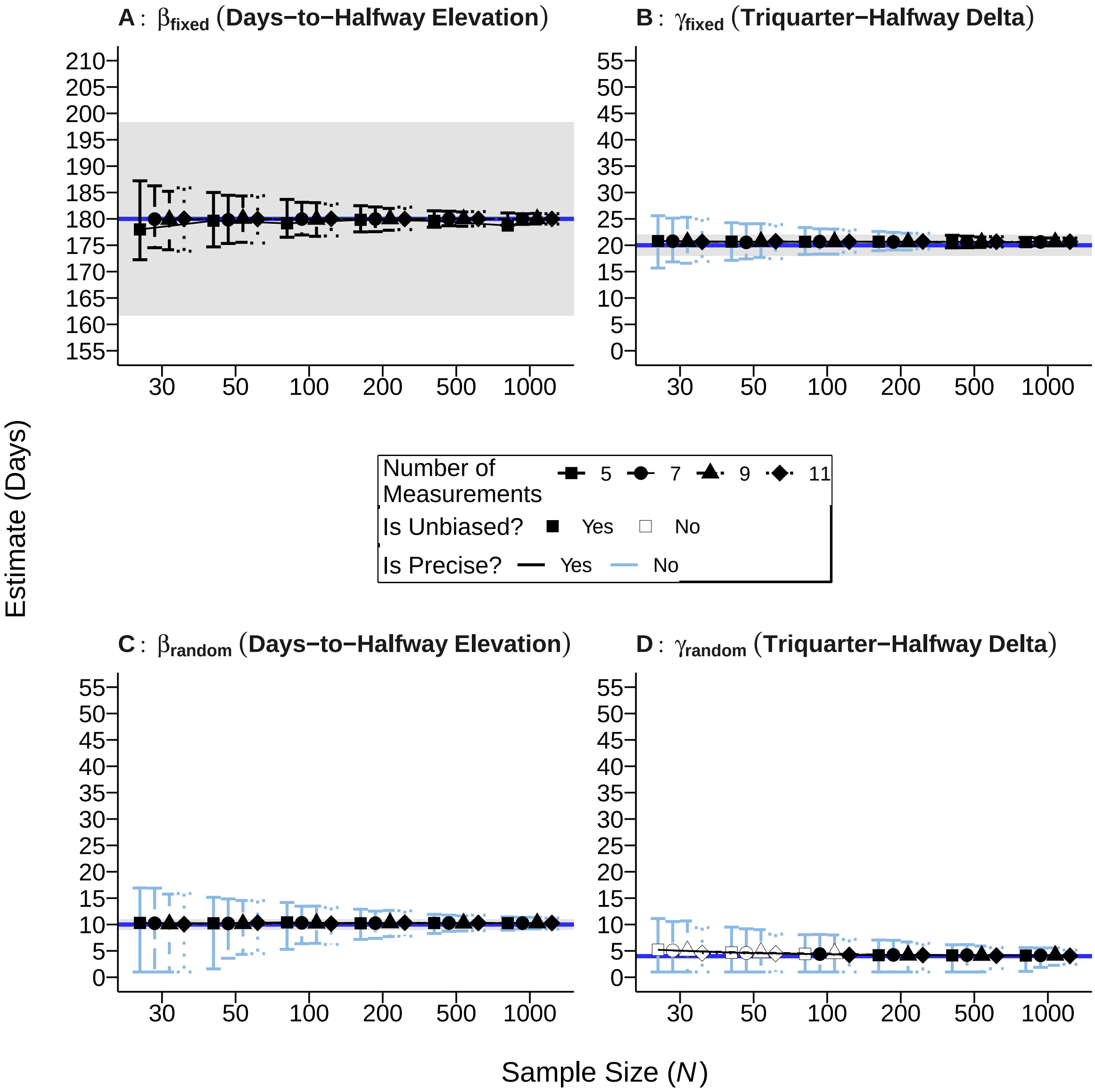
- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.4B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): no cells.

- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 3.4D): five and nine measurements with $N \leq 100$ and seven an 11 with $N \leq 50$.

In summary, with middle-and-extreme spacing, estimation of all the day-unit parameters is unbiased using five and nine measurements with $N \leq 100$ and seven an 11 with $N \leq 50$, which is indicated by the emboldened text in the 'Unbiased' column of Table 3.7.

Figure 3.5

Parameter Estimation Plots for Day-Unit Parameters With Middle-and-Extreme Spacing in Experiment 2



1835 *Note.* Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}).
1836 Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel
1837 C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D:
1838 Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue
1839 horizontal lines in each panel represent the population value for each parameter. Population values for each
1840 day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray
1841 bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average
1842 parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of
1843 estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside
1844 the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff
1845 (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that

random-effect parameter units are in standard deviation units. See Table ?? for specific values estimated for each parameter and Table 3.8 for ω^2 effect size values.

Table 3.8

Partial ω^2 Values for Independent Variables With Middle-and-Extreme Spacing in Experiment 2

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 3.5A)	0.05	0.03	0.01
β_{random} (Figure 3.5B)	0.14	0.12	0.01
γ_{fixed} (Figure 3.5C)	0.07	0.04	0.01
γ_{random} (Figure 3.5D)	0.05	0.09	0.00

Note. NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 1000), NM x S = interaction between number of measurements and sample size, β_{fixed} = fixed-effect days-to-halfway elevation parameter, γ_{fixed} = fixed-effect halfway-triquarter delta parameter, β_{random} = random-effect days-to-halfway elevation parameter, and γ_{random} = random-effect halfway-triquarter delta parameter.

3.2.6.0.2 Precision

With respect to precision for middle-and-extreme spacing, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 3.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 3.4B): all measurements numbers with $N \geq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 3.4C): all cells.
- random-effect halfway-triquarter delta parameter (γ_{random}; Figure 3.4D): all cells.

In summary, with middle-and-extreme spacing, precise estimation can be obtained for the fixed-effect day-unit parameters using at least five measurements with $N \geq 500$. For the random-effect day-unit parameters, no manipulated measurement number-sample size pairing results in precise estimation (see the 'Precise' column of Table 3.7).

3.2.6.0.3 Qualitative Description

For middle-and-extreme spacing in Figure 3.5, although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurement number-sample size pairings. With respect to bias under middle-and-extreme spacing, it is negligible under all manipulated measurement number-sample size pairings and so listing pairings that result in the greatest improvements in bias is of little value. With respect to precision under middle-and-extreme spacing, the largest improvements in precision in the estimation of all day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) result from using the following measurement number-sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): five measurements across all sample sizes, which results in a maximum error bar length of 9.92 days.
- random-effect days-to-halfway elevation parameter (β_{random}): five measurements across all sample sizes, which results in a maximum error bar length of 15.94 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): five measurements across all sample sizes, which results in a maximum error bar length of 10.13 days.

For an applied researcher, one plausible question might be what measurement number-sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with middle-and-extreme spacing. In looking across the measurement number-sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters with middle-and-extreme spacing result from using five measurments with any sample size (see the emboldened text in the 'Qualitative Description' column of Table 3.7).

3.2.6.1 Summary of Results

In summarizing the results for middle-and-extreme spacing, estimation of all day-unit parameters is unbiased using five and nine measurements with $N \leq 100$ and seven an 11 with $N \leq 50$ (see bias). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement number-sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number-sample size pairing under time-interval decreasing spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement number-sample size pairings. With middle-and-extreme spacing, the largest improvements in bias and precision in the estimation of all day-unit parameters are obtained from using five measurements any sample size (see qualitative description).

3.3 What Measurement Number-Sample Size Pairings Should be Used With Each Spacing Schedule?

In Experiment 2, I was interested in determining the measurement number-sample size pairings that resulted in high modelling accuracy (unbiased and precise parameter estimation) for each spacing schedule. Table 3.9 summarizes the results for each spacing schedule in Experiment 2. Text within the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairing needed to, respectively, obtain unbiased an precise estimation of all the day-unit parameters. The ‘Error Bar Length’ column indicates longest error bar lengths that result in the estimation of each day-unit parameter from using the measurement number-sample size pairings liste in the ‘Qualitative Description’ column. Although no measurement number-sample size pairing resulted in high modelling accuracy for any spacing schedule, the greatest improvements in modelling accuracy were made with the following pairings for each spacing schedule (see Table 3.9):

- equal: seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$.
- time-interval increasing: five measurements with $N \geq 500$, seven measurements with $N \geq 200$ or nine measurements with $N \leq 200$.
- time-interval decreasing: five measurements with $N \geq 500$, seven measurements with $N \geq 200$ or nine measurements with $N \leq 200$.
- middle-and-extreme: five measurements with any manipulated sample size.

Because each spacing schedule obtains comparable moelling accuracy as indicated by the similar error bar lengths, two statements can be made. First, using either seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ with any spacing schedule except middle-and-extreme spacing results in similar modelling accuracy. Second, middle-and-extreme spacing results in the highest modelling accuracy.

Table 3.9

Concise Summary of Results Across All Spacing Schedule Levels in Experiment 2

				Error Bar Summary			
Spacing Schedule	Unbiased	Precise	Qualitative Description	β_{fixed}	γ_{fixed}	β_{random}	γ_{random}
Equal (see Figure 3.2 and Table 3.1)	NM $\geq$ 7 with N = 1000 or NM $\geq$ 9 with $N \geq$ 100	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	12.67	9.79	16.02	10.08
Time-interval increasing (see Figure 3.3 and Table 3.3)	NM $\geq$ 9 with $N \geq$ 200 or NM = 11 with N = 1000	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	13.27	9.69	16.28	10.15
Time-interval decreasing (see Figure 3.4 and Table 3.5)	NM = 11 with $N \geq$ 1000	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	13.41	9.62	16.16	10.32
Middle and extreme (see Figure 3.5 and Table 3.7)	**NM $\geq$ 5 with $N \geq$ 200** or **NM $\in \{5, 7\}$ with N = 100**	No cells	Largest improvements in bias and precision with NM = 5	14.96	9.92	15.94	10.13

Note. 'Qualitative Description' column indicates the number of measurements that result in the greatest improvements in bias and precision across all day-unit parameters. 'Error Bar Summary' columns list the longest error bar lengths that result for each day-unit parameter using the measurement number-sample size pairing listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. N = sample size, NM = number of measurements.

The results of Experiment 2 are the first (to my knowledge) to provide measurement number and sample size guidelines for researchers interested in using nonlinear functions to model nonlinear change. Although previous simulation studies have investigated how to accurately model nonlinear change, three characteristics limit these results. First, some studies investigated the issue with fixed-effects models (e.g., Finch, 2017). Given that researchers often model effects as random, findings with fixed-effects effects models are limited in their application. Second, some studies used linear functions to model nonlinear change (e.g., Fine et al., 2019; J. Liu et al., 2022). Given that the parameters of linear functions become uninterpretable when modelling nonlinear change, these models are less useful to practitioners. Third, some studies implemented unrealistic model fitting procedures by dropping a random-effect parameter from the model each time convergence failed (Finch, 2017). By dropping random-effect parameters when model convergence failed, estimation accuracy could not meaningfully evaluated for parameters because values could have been obtained with reduced models.

In summary, the results of Experiment 2 provide measurement number-sample size guidelines for researchers interested in modelling nonlinear change. Importantly, because no measurement number-sample pairing resulted in unbiased and precise estimation of all the day-unit parameters, the guidelines provided by this study are only suggestions to obtain the greatest improvements in bias and precision. Although researchers are encouraged to use larger measurement numbers and sample sizes than suggested in the current guidelines, the improvements in bias and accuracy are likely to be incommensurate with the measurement number and sample size increments.

4 Experiment 3

In Experiment 3, I investigated the measurement number-sample size pairings needed to achieve accurate modelling (i.e., low bias and high precision) under different levels of time structuredness. Before presenting the results of Experiment 3, I will present my design and and analysis goals. For the design, I conducted a 3(time structuredness: time-structured data, time-unstructured data [fast response rate], time-unstructured data [slow response rate]) x 4(number of measurements: 5, 7, 9, 11) x 6(sample size: 30, 50, 100, 200, 500, 1000) study. For the analysis, I was primarily interested in determining, for each level of time structuredness, what measurement number-sample size pairings achieved accurate modelling (i.e., low bias and high precision).

4.1 Methods

4.1.1 Variables Used in Simulation Experiment

4.1.1.1 Independent Variables

4.1.1.1.1 Number of Measurements

For the number of measurements, I used the same values as in Experiment 1 of 5, 7, 9, and 11 measurements (see number of measurements for more discussion)).

4.1.1.1.2 Sample Size

For sample size, I used the same values as in Experiment 2 of 30, 50, 100, 200, 500, and 1000 (see sample size for more discussion).

4.1.1.1.3 Time Structuredness{time-structuredness}

Time structuredness describes the extent to which, at each time point, data are obtained at the exact same time point. The manipulation of time structuredness was adopted from the manipulation used in Coulombe et al. (2016) with a slight modification. In Coulombe et al. (2016), time-unstructured data were generated according to an exponential pattern such that most data were obtained at the beginning of the response window, with a smaller amount of data being obtained towards the end of the response window. Importantly, Coulombe et al. (2016) employed a non-continuous function for generating time-unstructured data: A binning method was employed such that 80% of the data were obtained within a time period equivalent to 12% (fast response rate) or 30% (slow response rate) of the entire response window. Using a response window length of 10 days with a fast response rate, the procedure employed by Coulombe et al. (2016) for generating time-unstructured data would have generated the following percentages of data in each of the four bins (note that, using the data generation procedure for Coulombe et al. (2016), the effective response window length was 4 days instead of 10 days):[17]

1) Bin 1: 60% of the data would be generated in the initial 10% length of the response window (0–0.4 day).
2) Bin 2: 20% of the data would be generated in the next 20% length of the response response window (0.4–1.2 days).
3) Bin 3: 10% of the data would be generated in the next 30% length of the response window (1.2–2.4 days).
4) Bin 4: the remaining 10% of the data would be generated in the remaining 40% length of the response window (2.4–4 days).

Note that, summing the data percentages and time durations from the first two bins

[17] The data generation procedure in (ref:coulombe2016) for a fast response rate assumed that all of the data were collected within the initial 40% length of the nominal response window length (i.e., 4 days in the current example).

yields an 80% cumulative response rate that is obtained in the initial 12% length of the full-length response window of 10 days (i.e., $(\frac{1.2}{10})100\% = 12\%$). Also note that, in Coulombe et al. (2016), a data point in each bin was randomly assigned a measurement time within the bin's time range. In the current example where the full-length response window had a length of 10 days, a data point obtained in the first bin would be randomly assigned a measurement time between 0–0.4.

Although Coulombe et al. (2016) generated time-unstructured data to resemble data collection conditions——response rates have been shown to follow an exponential pattern (Dillman et al. (2014); Pan (2010))——the use of a pseudo-continuous binning function for generating time-unstructured data lacked ecological validity. Therefore, the simulations here used a continuous function to create more realistic versions of time-unstructured data. Specifically, the exponential function shown below in Equation 4.1 was used:

$$y = M(1 - e^{-ax}), \tag{4.1}$$

where x stores the time delay for a measurement at a particular time point, y represents the cumulative response percentage achieved at a given x time delay, a sets the rate of growth of the cumulative response percentage over time, and M sets the range of possible y values. Two important points need to be made with respect to the M parameter (range of possible y values) and the response window length used in the current simulations. First, because the range of possible values for the cumulative response percentage (y) is 0–1 (data can be collected from a 0% to a maximum of 100% of respondents; $\{y : 0 \leq y \leq 1\}$), the M parameter had a value of 1 (M = 1). Second, the response window length in the current simulations was 36 days, and so the range of possible time delay values was between 0–36 ($\{x : 0 \leq x \leq 36\}$).[18]

To replicate the time structuredness manipulation in Coulombe et al. (2016) using the continuous exponential function of Equation 4.1, the growth rate parameter (a) had to be calibrated to achieve a cumulative response rate of 80% after either 12% or 30% of the response window length of 36 days. The derivation below solves for a, with Equation 4.2 showing the equation for computing a.

$$
\begin{aligned}
y &= M(1 - e^{-ax}) \\
y &= M - Me^{-ax} \\
y &= 1 - e^{-ax} \\
e^{-ax} &= 1 - y \\
-ax\log(e) &= \log(1 - y) \\
a &= \frac{\log(1 - y)}{-x} \qquad (4.2)
\end{aligned}
$$

Because the target response rate was 80%, y took on a value of .80 (y = .80). Given that the response window length in the current simulations was 36 days, x took on a value of 4.32 (12% of 36) when time-unstructured data were defined by a fast response rate and 10.80 (30% of 36) when time-unstructured data were defined by a slow response rate.

[18] A value of 36 days was used because the generation of time-unstructured data had to remain independent of the manipulation of measurement number (i.e., the response window lengths used in generating time-unstructured data could not vary with the number of measurements). To ensure the manipulations of measurement number and time structuredness remained independent, the reponse window length had to remain constant for all measurement number conditions with equal spacing. Looking at Table 2.1, the longest possible response window that fit within all measurement number conditions with equal spacing was the interval length of the 11-measurement condition (i.e., 36 days).

Using Equation 4.2 yielded the following growth rate parameter values for fast and slow response rates (a_{fast}, a_{slow}):

$$a_{fast} = \frac{\log(1 - .80)}{-4.32} = 0.37$$
$$a_{slow} = \frac{\log(1 - .80)}{-10.80} = 0.15$$

Therefore, to obtain 80% of the data with a fast response rate (i.e., in 4.32 days), the growth parameter (a) needed to have a value of 0.37 ($a_{fast} = 0.37$) and, to obtain 80% of the data with a slow response rate (i.e., in 10.80 days), the growth parameter (a) needed to have a value of 0.15 ($a_{slow} = 0.15$). Using the above growth rate values derived for the fast and slow response growth rate parameters (a_{fast}, a_{slow}), the following functions were generated for fast and slow response rates:

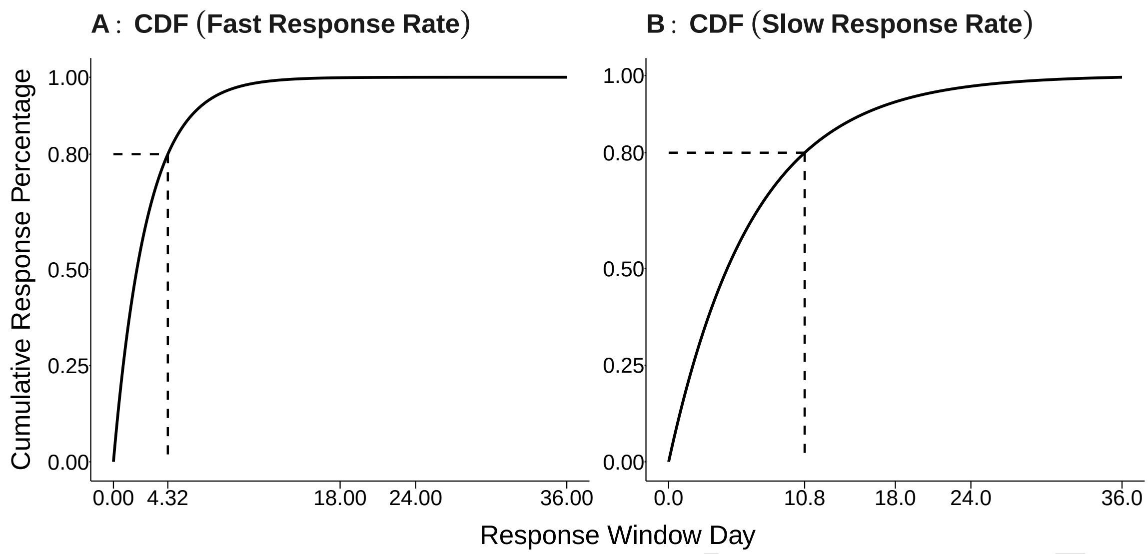
$$f_{fast}(x) = M(1 - e^{a_{fast}x}) = M(1 - e^{-0.37x}) \text{ and} \tag{4.3}$$

$$f_{slow}(x) = M(1 - e^{a_{slow}x}) = M(1 - e^{-0.15x}). \tag{4.4}$$

Using Equations 4.3–4.4, Figure 10 shows the resulting cumulative distribution functions (CDF) for time-unstructured data that show the cumulative response percentage as a function of time. Panel A shows the cumulative distribution function for a fast response rate (Equation 4.3), where an 80% response rate was obtained in 4.32 days. Panel B shows the cumulative distribution function for a slow response rate (Equation 4.4), where an 80% response rate was obtained in 10.80 days.

Figure 4.1

Cumulative Distribution Functions (CDF) With Fast and Slow Response Rates



Note. Panel A: Cumulative distribution function for a fast response rate (Equation 4.3), where an 80% response rate is obtained in 4.32 days. Panel B: Cumulative distribution function for a slow response rate (Equation 4.4), where an 80% response rate is obtained in 10.80 days.

4.1.1.2 Constants

Because the nature of change not manipulated in Experiment 3, I set it to have a constant value across all cells. To keep the nature of change constant across all cells, I set the fixed-effect days-to-halfway elevation parameter (β_{fixed}) to have a value of 180. Another variable set to a constant value across the cells was measurement spacing (equal spacing was used).

4.1.1.3 Dependent Variables

4.1.1.3.1 Convergence Success Rate

The proportion of iterations in a cell where models converged defined the **convergence success rate**.[19] Equation (4.5) below shows the calculation used to compute the

[19]Specifically, convergence was obtained if the convergence code returned by OpenMx was 0.

convergence success rate:

$$\text{Convergence success rate} = \frac{\text{Number of models that successfully converged in a cell}}{n}, \tag{4.5}$$

where n represents the total number of models run in a cell.

4.1.1.3.2 Bias

Bias was calculated to evaluate the accuracy with which each logistic function parameter was estimated. As shown below in Equation (4.6), *bias* was obtained by calculating the difference between the population value set for a parameter and the average estimated value in each cell.

$$\text{Bias} = \text{Population value for parameter} - \text{Average estimated value} \tag{4.6}$$

Bias was calculated for the fixed- and random-effect parameters of the baseline (θ_{fixed}, θ_{random}), maximal elevation (α_{fixed}, α_{random}), days-to-halfway elevation (β_{fixed}, β_{random}), and the halfway-triquarter delta parameters (γ_{fixed}, γ_{random}).

4.1.1.3.3 Precision

In addition to computing bias, precision was calculated to evaluate the confidence with which each parameter was estimated in a given cell. *Precision* was obtained by computing the range of values covered by the middle 95% of values estimated for a logistic parameter in each cell. By using the middle 95% of estimated values, a plausible range of population estimates was obtained.

4.1.2 Overview of Data Generation

Data generation was computed the same way as in Experiment 1 (see data generation) with one addition to the procedure needed for time structuredness. The section that follows details how time structuredness was simulated.

4.1.2.0.1 Simulation Procedure for Time Structuredness

To simulate time-unstructured data, response rates at each collection point followed an exponential pattern described by either a fast or slow response rate (for a review, see time structuredness). Importantly, data generated for each person at each time point had to be sampled according to a probability density function defined by either the fast or slow response rate cumulative distribution function. In the current context, a *probability density function* describes the probability of sampling any given time delay value x where the range of time delay values is 0–36 ($\{x : 0 \leq x \leq 36\}$). To obtain the probability density functions for fast and slow response rates, the response rate function shown in Equation (**??**) was differentiated with respect to x to obtain the function shown below in Equation 4.7[20]:

$$
\begin{aligned}
f' = \frac{\partial f(x)}{\partial x} &= \frac{\partial}{\partial x} M(1 - e^{-ax}). \\
&= M(e^{-ax}a)
\end{aligned}
\tag{4.7}
$$

To compute the probability density function for the fast response rate cumulative distribution function, the growth rate parameter a was set to 0.37 in Equation 4.7 to obtain

[20]Euler's notation for differentiation is used to represent derivatives. In words, $\frac{\partial f(x)}{\partial x}$ means that the derivative of the function $f(x)$ is taken with respect to x.

the following function in Equation 4.8:

$$f'_{fast}(x) = M(e^{-a_{fast}x}a_{fast}) = M(e^{-0.37x}0.37). \tag{4.8}$$

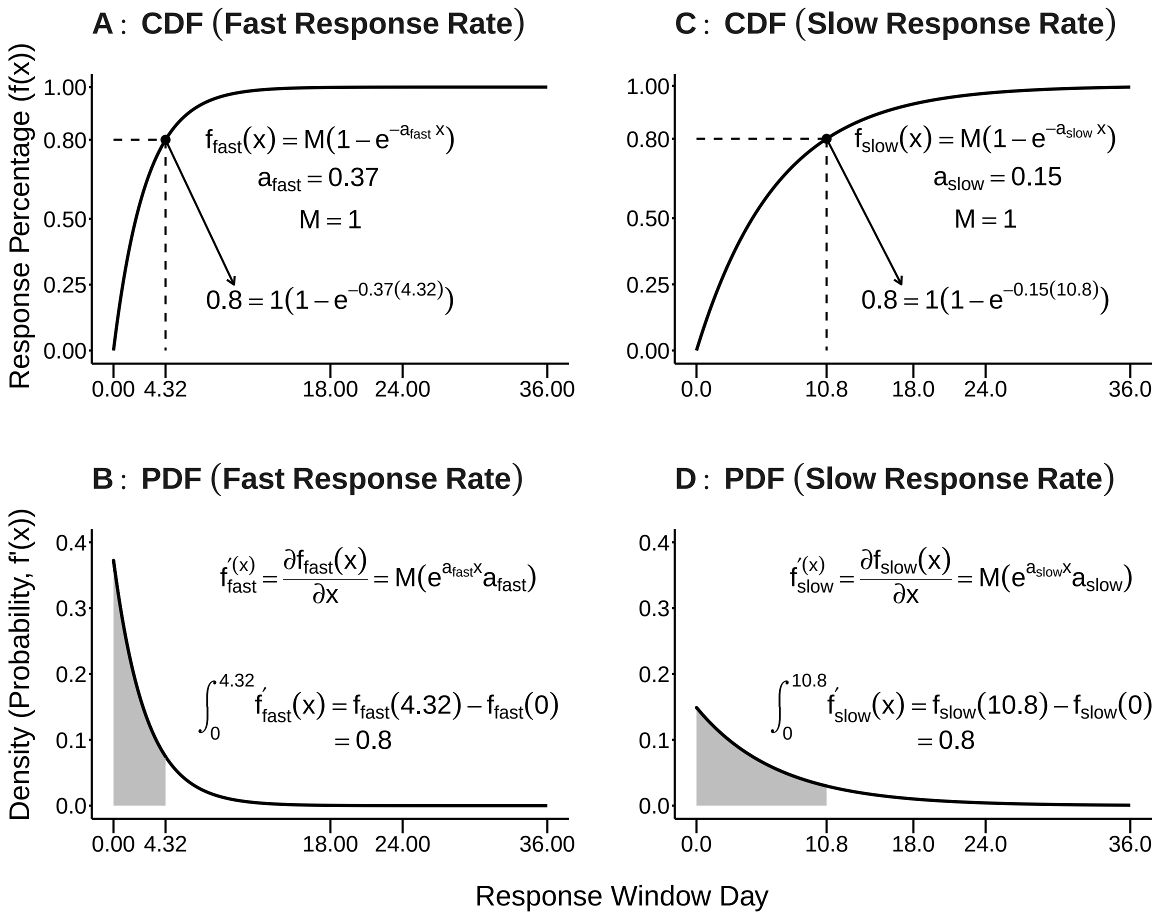
To compute the probability density function for the slow response rate cumulative distribution function, the growth rate parameter a was set to 0.15 in Equation 4.7 to obtain the following function in Equation 4.9:

$$f'_{slow}(x) = M(e^{-0.15}a_{slow}) = M(e^{-0.15}0.15). \tag{4.9}$$

Figure 4.2 shows the fast and slow response cumulative distribution functions (CDF) and their corresponding probability density functions (PDF). Panel A shows the cumulative distribution function for the fast response rate (with a growth parameter value a set to 0.37; see Equation 4.3) and Panel B shows the probability density function that results from computing the derivative of the fast response rate cumulative distribution function with respect to x (see Equation 4.8). Panel C shows the cumulative distribution function for the slow response rate (with a growth parameter value a set to 0.15; see Equation 4.4)) and Panel D shows the probability density function that results from computing the derivative of the slow response rate cumulative distribution function with respect to x (see Equation 4.9 and section on time structuredness for more discussion). For the fast response rate functions, an 80% response rate is obtained after 4.32 days or, equivalently, 80% of the area underneath the probability density function is obtained at 4.32 days ($\int_0^{4.32} f'_{fast}(x) = 0.80$; the integral from 0 to 4.32 of the probability density function for a fast response rate $f'(x)_{fast}$ is 0.80). For the slow response rate functions, an 80% response rate is obtained after 10.80 days or, equivalently, 80% of the area underneath the probability density function is obtained at 10.80 days ($\int_0^{10.80} f'_{slow}(x) = 0.80$; the integral from 0 to 10.80 of the probability density function for a slow response rate $f'(x)_{slow}$ is 0.80).

Figure 4.2

Cumulative Distribution Functions (CDF) and Probability Density Functions (PDF) for Fast and Slow Response Rates



Note. Panel A: Cumulative distribution function for the fast response rate (with a growth parameter value a set to 0.37; see Equation 4.3). Panel B: Probability density function that results from computing the derivative of the fast response rate cumulative distribution function with respect to x (see Equation 4.8). Panel C: Cumulative distribution function for the slow response rate (with a growth parameter value a set to 0.15; see Equation 4.4). Panel D: Probability density function that results from computing the derivative of the slow response rate cumulative distribution function with respect to x (see Equation 4.9 and Time Structuredness for more discussion on time structuredness). For the fast response rate functions, an 80% response rate is obtained after 4.32 days or, equivalently, 80% of the area underneath the probability density function is obtained at 4.32 days ($\int_0^{4.32} f'_{fast}(x) = 0.80$). For the slow response rate functions, an 80% response rate is obtained after 10.80 days or, equivalently, 80% of the area underneath the probability density function is obtained at 10.80 days ($\int_0^{10.80} f'_{slow}(x) = 0.80$).

Having computed probability density functions for fast and slow response rates, time delays could be generated to create time-unstructured data. To generate time-unstructured data for a person at a given time point, a time delay was first generated by sampling values according to the probability density function defined by either a fast or slow response rate (Equations 4.8–4.9). The sampled time delay was then added to the value of the current measurement day, with the combined measurement day then being plugged into the logistic function (Equation 2.3) along with a set of person-specific parameter values to generate an observed score at a given time point for a given person.

4.1.3 Modelling of Each Generated Data Set

Each generated data set was modelled using the structured latent growth curves outlined in Experiment 1 (see data modelling. For a detailed explanation of how the logistic function was fit into the structural equation modelling framework, see Technical Appendix B.

4.1.4 Analysis of Data Modelling Output and Accompanying Visualizations

Analysis and visualization was conducted as outlined in Experiment 1 (see analysis and visualization).

4.2 Results and Discussion

In the sections that follow, I organize the results by presenting them for each spacing schedule (equal, time-interval increasing, time-interval decreasing, middle-and-extreme). Importantly, only the results for the day-unit parameters will be presented (i.e., fixed- and random-effect days-to-halfway elevation and halfway-triquarter delta parameters [β_{fixed}, β_{random}, γ_{fixed}, γ_{random}, respectively]). The results for the likert-unit parameters (i.e., fixed- and random-effect baseline and maximal elevation parameters [θ_{fixed}, θ_{random}, α_{fixed}, α_{random}, respectively]) were largely trivial and so are presented in Appendix C).

For each level of time structuredness, I first provide a concise summary of the results and then provide a detailed report of the estimation accuracy of each day-unit parameter of the logistic function. Because the lengths of the detailed reports are considerable, I provide concise summaries before the detailed reports to establish a framework to interpret the detailed reports. The detailed report of each measurement spacing schedule will summarize the results of each day-unit's parameter estimation plot, report partial ω^2 values, and then provide a qualitative summary.

4.2.1 Pre-Processing of Data and Model Convergence

After collecting the output from the simulations, non-converged models (and their corresponding parameter estimates) were removed from subsequent analyses. Table D.3 in Appendix B provides the convergence success rates for each cell in Experiment 3. Model convergence was almost always above 90% and convergence rates rates below 90% only occurred in two cells with five measurements.

4.2.2 Time-Structured Data

For time-structured data, Table 4.1 provides a concise summary of the results for the day-unit parameters (see Figure 3.2 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 4.1 and provide elaboration when necessary.

Before presenting the results for equal spacing, I provide a brief description of the concise summary table created for each spacing schedule and shown for equal spacing in Table 4.1. ext in the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the measurement number-sample size pairing needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with equal spacing). The 'Error Bar Length' column indicates the error bar length that results from using the lower-bounding measurement number-sample size pairing listed in the 'Qualitative Description' column.

Table 4.1

Concise Summary of Results for Time-Structured Data in Experiment 3

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 4.3A)	All cells	All cells	Unbiased and precise estimation in all cells	15.13
γ_{fixed} (Figure 4.3B)	All cells	NM $\geq$ 9 with N = 500	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	9.79
β_{random} (Figure 4.3C)	All cells	No cells	Largest improvements in precision with NM = 7	17.22
γ_{random} (Figure 4.3D)	**NM $\geq$ 9 with $N \geq$ 200**	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	10.08

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with equal spacing). 'Error Bar Length' column indicates the maximum error bar length that results from using the measurement number-sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

4.2.2.0.1 Bias

Before presenting the results for bias, I provide a description of the set of parameter estimation plots shown in Figure 4.3 and in the results sections for the other spacing schedules in Experiment 2. Figure 4.3 shows the parameter estimation plots for each day-unit parameter and Table 3.2 provides the partial ω^2 values for each independent variable of each day-unit parameter. In Figure 4.3, blue horizontal lines indicate the population values for each parameter (with population values of β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, and γ_{random} = 4.00). Gray bands indicate the ±10% margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Panels A–B show the parameter estimation plots for the fixed- and random-effect days-to-halfway elevation parameters (β_{fixed} and β_{random}, respectively). Panels C–D show the parameter estimation plots for the fixed- and random-effect triquarter-halfway delta parameters (γ_{fixed} and γ_{random}, respectively). Note that random-effect parameter units are in standard deviation units.

With respect to bias for time-structured data, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

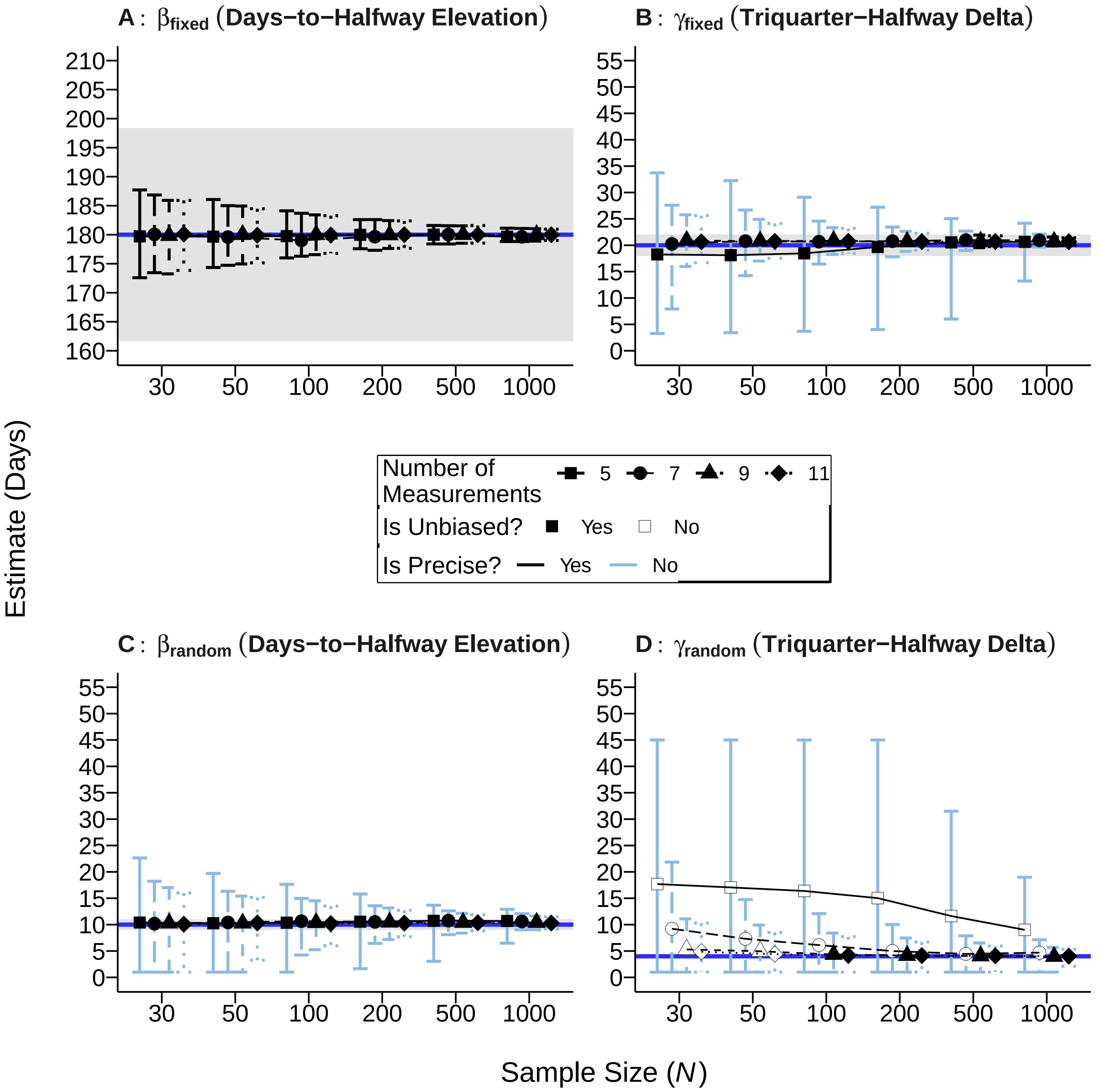
- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.3A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.3B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.3C): no cells.

- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 4.3D): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 100$.

In summary, with time-structured data, estimation of all the day-unit parameters across all manipulated nature-of-change values is unbiased using at least nine measurements with $N \geq 200$, which is indicated by the emboldened text in the 'Unbiased' column of Table 4.1.

Figure 4.3

Parameter Estimation Plots for Day-Unit Parameters With Time-Structured Data in Experiment 3



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff

(i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table F.3 for specific values estimated for each parameter and Table 4.2 for ω^2 effect size values.

Table 4.2

Partial ω^2 Values for Manipulated Variables With Time-Structured Data in Experiment 3

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 4.3A)	0.00	0.02	0.00
β_{random} (Figure 4.3B)	0.14	0.27	0.03
γ_{fixed} (Figure 4.3C)	0.25	0.12	0.07
γ_{random} (Figure 4.3D)	0.18	0.03	0.01

NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 100), NM x S = interaction between number of measurements and sample size.

4.2.2.0.2 Precision

With respect to precision for time-structured data, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.3A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.3B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.3C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 4.3D): all cells.

In summary, with time-structured data, precise estimation can be obtained for the fixed-effect day-unit parameters using at least nine measurements with $N \geq 500$, but no manipulated measurement number-sample size pairing results in precise estimation of the random-effect day-unit parameters (see the ‘Precise’ column of Table 4.1).

4.2.2.0.3 Qualitative Description

For time-structured data in Figure 4.3, although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurement number-sample size pairings. With respect to bias under time-structured data, the largest improvements in bias result with the following measurement number-sample size pairings for the random-effect triquarter-halfway delta parameter (γ_{fixed}):

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-structured data, the largest improvements in precision for the estimation of all the day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) result from using the following measurement number-sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 9.79 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which results in a error bar length of 17.22 days.
- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 10.08 days.

For an applied researcher, one plausible question might be what measurement number-sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-structured data. In looking across the measurement number-sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters result with the following measurement number-sample size pairing(s): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 4.1).

4.2.2.1 Summary of Results

In summarizing the results for time-structured data, estimation of all day-unit parameters is unbiased using least nine measurements with $N \geq 200$ (see bias). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement number-sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number-sample size pairing under equal spacing results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement number-sample size pairings. With time-structured data, the largest improvements in bias and precision in the estimation of all day-unit parameters are obtained using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see qualitiative description).

4.2.3 Time-Unstructured Data Characterized by a Fast Response Rate

For time-unstructured data characterized by a fast response rate, Table 4.3 provides a concise summary of the results for the day-unit parameters (see Figure 4.4 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 4.3 and provide elaboration when necessary (for a description of Table 4.3, see concise summary).

Table 4.3

Concise Summary of Results for Time-Unstructured Data (Fast Response Rate) in Experiment 3

			Description	
Parameter	Unbiased	Precise	Qualitative Description	Error Bar Length
β_{fixed} (Figure 4.4A)	All cells	All cells	Unbiased and precise estimation in all cells	15.35
γ_{fixed} (Figure 4.4B)	All cells	NM $\geq$ 9 with $N \geq$ 500	Largest improvements in precision using **NM = 7 with *N* $\geq$ 200** or **NM = 9 with *N* $\leq$ 100**	10.25
β_{random} (Figure 4.4C)	All cells	No cells	Largest improvements in precision with NM = 7	17.47
γ_{random} (Figure 4.4D)	**NM $\geq$ 7 with *N* = 1000** or **NM $\geq$ 9 with *N* $\geq$ 200** or **NM = 11 with *N* = 100**	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq$ 200 or NM = 9 with $N \leq$ 100	10.51

Note. Text in the 'Unbiased' and 'Precise' columns indicates the measurement number-sample size pairings that, respectively, result in unbiased and precise estimation. Emboldened text in the 'Unbiased' and 'Qualitative Description' columns indicates the number of measurements needed to, respectively, obtain unbiased estimates and the greatest improvements in bias and precision across all day-unit parameters (acceptable precision not achieved in the estimation of all day-unit parameters with equal spacing). 'Error Bar Length' column indicates the maximum error bar length that results from using the measurement number-sample size recommendation listed in the 'Qualitative Description' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4. NM = number of measurements.

4.2.3.0.1 Bias

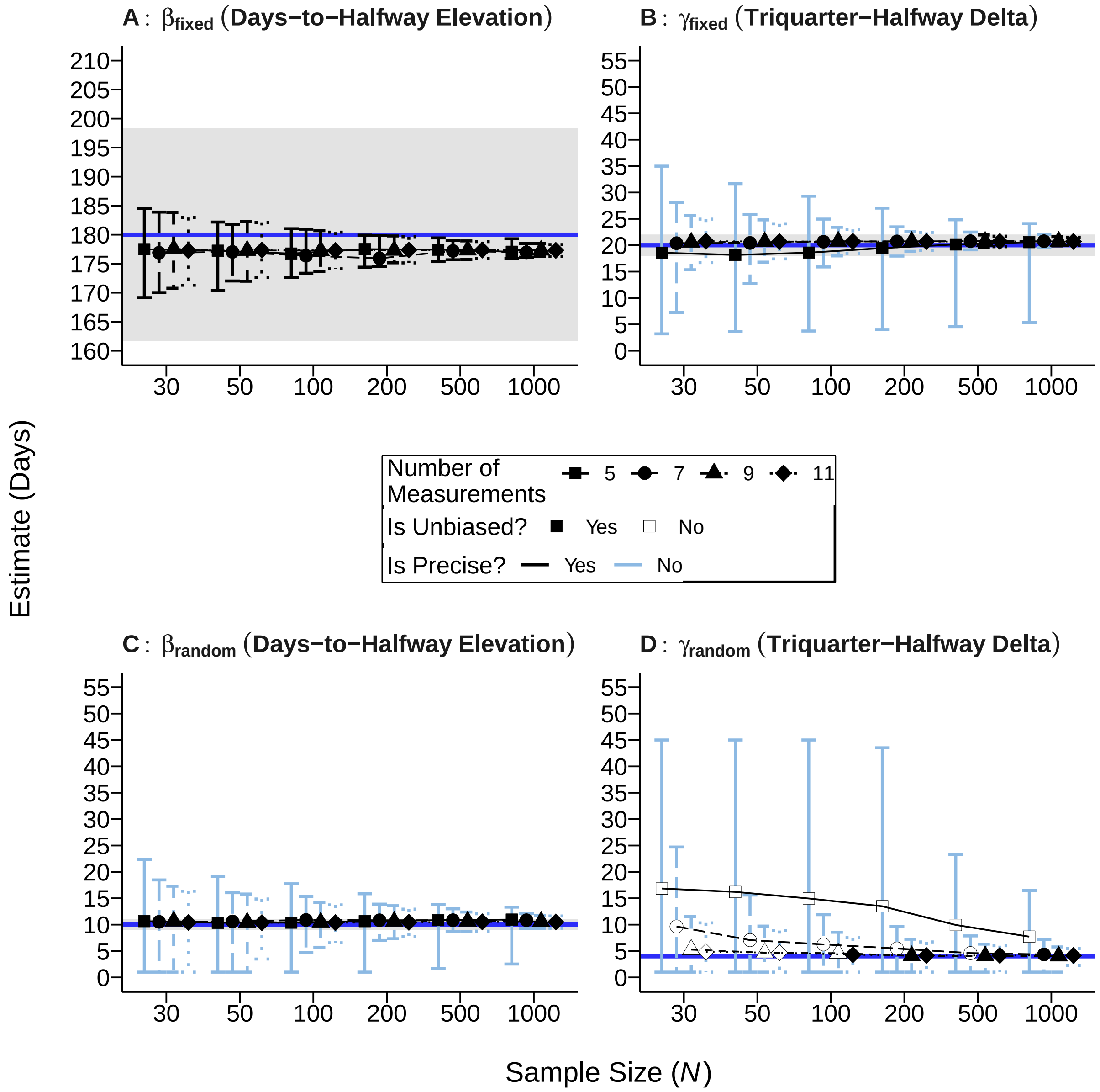
With respect to bias for time-unstructured data characterized by a fast response rate, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.4B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.4C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 4.4D): five measurements across all sample sizes, seven measurements with $N \leq 500$, nine measurements with $N \geq 100$, and 11 measurements with $N \leq 50$.

In summary, with time-unstructured data characterized by a fast response rate, estimation of all the day-unit parameters across all manipulated nature-of-change values is unbiased using at least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 4.3.

Figure 4.4

Parameter Estimation Plots for Day-Unit Parameters With Time-Unstructured Data Characterized by a Fast Response Rate in Experiment 3



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff

(i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table F.3 for specific values estimated for each parameter and Table 4.4 for ω^2 effect size values.

Table 4.4
Partial ω^2 Values for Manipulated Variables With Time-Structured Data in Experiment 3

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 4.4A)	0.00	0.02	0.00
β_{random} (Figure 4.4B)	0.15	0.27	0.03
γ_{fixed} (Figure 4.4C)	0.29	0.14	0.08
γ_{random} (Figure 4.4D)	0.17	0.04	0.01

NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 100), NM x S = interaction between number of measurements and sample size.

4.2.3.0.2 Precision

With respect to precision for time-unstructured data characterized by a fast response rate, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.4A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.4B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.4C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 4.4D): all cells.

In summary, with time-unstructured data characterized by a fast response rate, precise estimation can be obtained for the fixed-effect day-unit parameters using at least nine measurements with $N \geq 500$, but no manipulated measurement number-sample size pairing results in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 4.3).

4.2.3.0.3 Qualitative Description

For time-unstructured data characterized by a fast response rate (see Figure 4.4), although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurement number-sample size pairings. With respect to bias under time-unstructured data characterized by a fast response rate, the largest improvements in bias result with the following measurement number-sample size pairings for the random-effect triquarter-halfway delta parameter (γ_{fixed}):

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-unstructured data characterized by a fast response rate, the largest improvements in precision for the estimation of all the day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) result from using the following measurement number-sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 10.25 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which results in a error bar length of 17.47 days.

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 10.51 days.

For an applied researcher, one plausible question might be what measurement number-sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-unstructured data characterized by a fast response rate. In looking across the measurement number-sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters result with the following measurement number-sample size pairing(s): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 4.3).

4.2.3.1 Summary of Results

In summarizing the results for time-unstructured data characterized by a fast response rate, estimation of all day-unit parameters is unbiased using least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$ (see bias). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement number-sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number-sample size pairing under time-unstructured data characterized by a fast response rate results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement number-sample size pairings. With time-unstructured data characterized by a fast response rate, the largest improvements in bias and precision in the estimation of all day-unit parameters are obtained using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see qualitiative description).

4.2.4 Time-Unstructured Data Characterized by a Slow Response Rate

For time-unstructured data characterized by a slow response rate, Table 4.5 provides a concise summary of the results for the day-unit parameters (see Figure 4.5 for the corresponding parameter estimation plots). The sections that follow will present the results for each column of Table 4.5 and provide elaboration when necessary (for a description of Table 4.5, see concise summary).

Table 4.5

Concise Summary of Results for Time-Unstructured Data (Slow Response Rate) in Experiment 3

			Summary	
Parameter	Unbiased	Precise	Qualitative Summary	Error Bar Length
β_{fixed} (Figure 4.5A)	All cells	All cells	Low bias and high precision in all cells	16.68
γ_{fixed} (Figure 4.5B)	All cells except NM = 5 with N = 50	NM = 7 with N = 200 or NM = 9 with $N \leq 500$	Largest improvements in precision using **NM = 7 with $N \geq$ 200** or **NM = 9 with $N \leq$ 100**	10.53
β_{random} (Figure 4.5C)	No cells except NM = 5 with N = 30 and NM = 11 with $N \leq 50$	No cells	Largest improvements in precision with NM = 7	18.44
γ_{random} (Figure 4.5D)	No cells	No cells	Largest improvements in bias and precision using NM = 7 with $N \geq 200$ or M = 9 with $N \leq 100$	10.9

Note.

Bolded text in the 'Low Bias' and 'Qualitative Summary' columns indicates the measurement number-sample size pairing needed to, respectively, achieve low bias and the greatest improvements in bias and precision across all day-unit parameters (high precision not achieved in the estimation of all day-unit parameters with time-unstructured data characterized by a slow response rate). 'Error Bar Length' indicates the longest error bar length that results from using the measurement number-sample size pairings in the 'Qualitative Summary' column. Parameter names and population values are as follows: β_{fixed} = fixed-effect days-to-halfway elevation parameter = 180; γ_{fixed} = fixed-effect halfway-triquarter delta parameter = 20; β_{random} = random-effect days-to-halfway elevation parameter = 10; γ_{random} = random-effect halfway-triquarter delta parameter = 4.

4.2.4.0.1 Bias

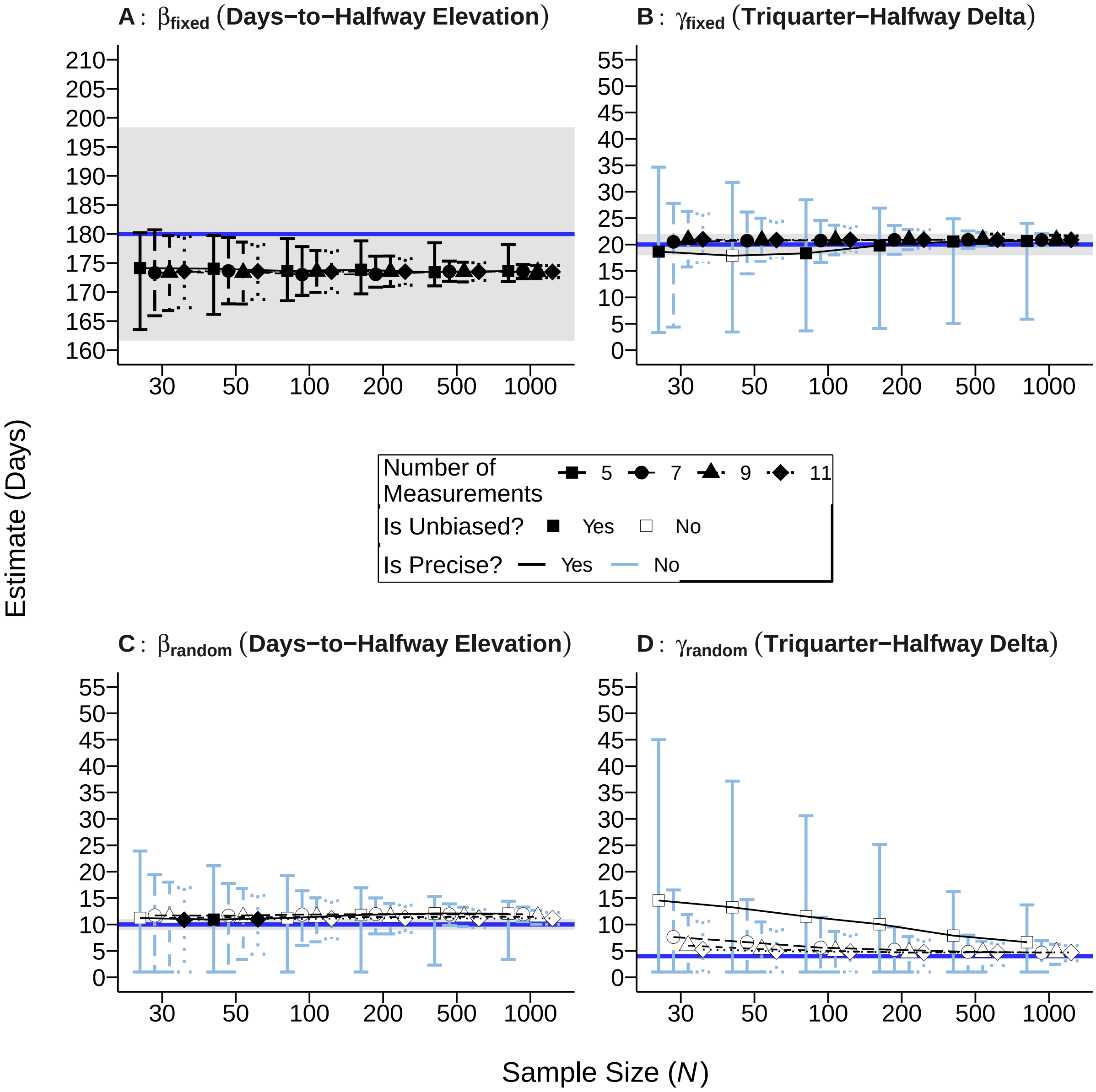
With respect to bias for time-unstructured data characterized by a slow response rate, estimates are biased (i.e., above the acceptable 10% cutoff) for each day-unit parameter in the following cells:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.5A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.5B): no cells.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.5C): no cells.
- random-effect triquarter-halfway elevation parameter (γ_{random}; Figure 4.5D): five measurements across all sample sizes, seven measurements with $N \leq 500$, nine measurements with $N \geq 100$, and 11 measurements with $N \leq 50$.

In summary, with time-unstructured data characterized by a slow response rate, estimation of all the day-unit parameters across all manipulated nature-of-change values is unbiased using at least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$, which is indicated by the emboldened text in the 'Unbiased' column of Table 4.5.

Figure 4.5

Parameter Estimation Plots for Day-Unit Parameters With Time-Unstructured Data Characterized by a Slow Response Rate in Experiment 3



Note. Panel A: Parameter estimation plot for the fixed-effect days-to-halfway elevation parameter (β_{fixed}). Panel B: Parameter estimation plot for the fixed-effect triquarter-halfway elevation parameter (γ_{fixed}). Panel C: Parameter estimation plot for the random-effect days-to-halfway elevation parameter (β_{random}). Panel D: Parameter estimation plot for the random-effect triquarter-halfway elevation parameter (γ_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: β_{fixed} = 180.00, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff

(i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table F.3 for specific values estimated for each parameter and Table 4.6 for ω^2 effect size values.

Table 4.6
Partial ω^2 Values for Manipulated Variables With Time-Unstructured Data Characterized by a Slow Response Rate in Experiment 3

	Effect		
Parameter	NM	S	NM x S
β_{fixed} (Figure 4.5A)	0.00	0.02	0.00
β_{random} (Figure 4.5B)	0.15	0.27	0.03
γ_{fixed} (Figure 4.5C)	0.29	0.14	0.08
γ_{random} (Figure 4.5D)	0.17	0.04	0.01

NM = number of measurements (5, 7, 9, 11), S = sample size (30, 50, 100, 200, 500, 100), NM x S = interaction between number of measurements and sample size.

4.2.4.0.2 Precision

With respect to precision for time-unstructured data characterized by a slow response rate, estimates are imprecise (i.e., error bar length with at least one whisker length exceeding 10% of a parameter's population value) in the following cells for each day-unit parameter:

- fixed-effect days-to-halfway elevation parameter (β_{fixed}; Figure 4.5A): no cells.
- fixed-effect halfway-triquarter delta parameter (γ_{fixed}; Figure 4.5B): five and seven measurements across all sample sizes and nine and 11 measurements with $N \leq 200$.
- random-effect days-to-halfway elevation parameter (β_{random}; Figure 4.5C): all cells.
- random-effect halfway-triquarter delta parameter [γ_{random}] in Figure 4.5D): all cells.

In summary, with time-unstructured data characterized by a slow response rate, precise estimation can be obtained for the fixed-effect day-unit parameters using at least nine measurements with $N \geq 500$, but no manipulated measurement number-sample size pairing results in precise estimation of the random-effect day-unit parameters (see the 'Precise' column of Table 4.5).

4.2.4.0.3 Qualitative Description

For time-unstructured data characterized by a slow response rate (see Figure 4.5), although no manipulated measurement number results in precise estimation of all the day-unit parameters, the largest improvements in precision (and bias) result from using moderate measurement number-sample size pairings. With respect to bias under time-unstructured data characterized by a slow response rate, the largest improvements in bias result with the following measurement number-sample size pairings for the random-effect triquarter-halfway delta parameter (γ_{fixed}):

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 100$ or nine measurements with $N \leq 50$.

With respect to precision under time-unstructured data characterized by a slow response rate, the largest improvements in precision for the estimation of all the day-unit parameters (except the fixed-effect days-to-halfway elevation parameter [β_{fixed}]) result from using the following measurement number-sample size pairings:

- fixed-effect triquarter-halfway delta parameter (γ_{fixed}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 10.53 days.
- random-effect days-to-halfway elevation parameter (β_{random}): seven measurements across all manipulated sample sizes, which which results in a error bar length of 18.44 days.

- random-effect triquarter-halfway delta parameter (γ_{random}): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$, which results in a maximum error bar length of 10.9 days.

For an applied researcher, one plausible question might be what measurement number-sample size pairing(s) results in the greatest improvements in bias and precision in the estimation of all day-unit parameters with time-unstructured data characterized by a fast response rate. In looking across the measurement number-sample size pairings in the above lists, it becomes apparent that greatest improvements in bias and precision in the estimation of all day-unit parameters result with the following measurement number-sample size pairing(s): seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see the emboldened text in the 'Qualitative Description' column of Table 4.5).

4.2.4.1 Summary of Results

In summarizing the results for time-unstructured data characterized by a slow response rate, estimation of all day-unit parameters is least seven measurements with $N = 1000$, nine measurements with $N \geq 200$, or 11 measurements with $N \geq 100$ (see bias). Precise estimation is never obtained in the estimation of all day-unit parameters with any manipulated measurement number-sample size pairing (see precision). Although it may be discouraging that no manipulated measurement number-sample size pairing under time-unstructured data characterized by a slow response rate results in precise estimation of all day-unit parameters, the largest improvements in precision (and bias) across all day-unit parameters are obtained with moderate measurement number-sample size pairings. With time-unstructured data characterized by a slow response rate, the largest improvements in bias and precision in the estimation of all day-unit parameters are obtained using seven measurements with $N \geq 200$ or nine measurements with $N \leq 100$ (see qualitiative description).

4.2.5 How Does Time Structuredness Affect Modelling Accuracy?

4.3 Summary

4.4 References

Adolf, J. K., & Fried, E. I. (2019). Ergodicity is sufficient but not necessary for group-to-individual generalizability. *Proceedings of the National Academy of Sciences*, *116*(14), 6540–6541. https://doi.org/10.1073/pnas.1818675116

Aguinis, H., & Bakker, R. M. (2021). Time is of the essence: Improving the conceptualization and measurement of time. *Human Resource Management Review*, *31*(2), 100763. https://doi.org/10.1016/j.hrmr.2020.100763

Aydin, B., Leite, W. L., & Algina, J. (2014). The Consequences of ignoring variability in measurement occasions within data collection waves in latent growth ,odels. *Multivariate Behavioral Research*, *49*(2), 149–160. https://doi.org/10.1080/00273171.2014.887901

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*, *28*(2), 135–167. https://doi.org/10.3102/10769986028002135

Beal, D. J. (2015). ESM 2.0: State of the art and future potential of experience sampling methods in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 383–407. https://doi.org/10.1146/annurev-orgpsych-032414-111335

Bergman, L., & Magnusson, D. (1990). *General issues about data quality in longitudinal research* (L. Bergman & D. Magnusson, Eds.; pp. 1–31). Cambridge University Press. shorturl.at/enwxM

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, *17*(12), 656–660. https://doi.org/10.1073/pnas.17.2.656

Blozis, S. A. (2004). Structured latent curve models for the study of change in multivariate repeated measures. *Psychological Methods*, *9*(3), 334–353. https://doi.org/10.1037/1082-989x.9.3.334

Bodenmann, G., Atkins, D. C., Schär, M., & Poffet, V. (2010). The association between daily stress and sexual activity. *Journal of Family Psychology*, *24*(3), 271–279. https://doi.org/10.1037/a0019365

Boker, S. M., Neale, M. C., Maes, H. H., Wilde, M. J., Spiegel, M., Brick, T. R., Estabrook, R., Bates, T. C., Mehta, P., Oertzen, T. von, Gore, R. J., Hunter, M. D., Hackett, D. C., Karch, J., Brandmaier, A., Pritikin, J. N., Zahery, M., & Kirkpatrick, R. M. (2020). *OpenMx user guide release 2.18.1.*

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Converting among effect sizes. In *Introduction to meta-analysis* (1st ed., pp. 45–49). John Wiley & Sons. https://doi.org/10.1002/9780470743386.ch7

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*(2), 431–449. https://doi.org/10.1037/a0038047

Brown, M. B., & Forsythe, A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, *69*(346), 364–367. https://doi.org/

10.1080/01621459.1974.10482955

Browne, M. W. (1993). Structured latent curve models. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate analysis: Future directions 2* (1st ed., pp. 171–197). Elsevier. https://doi.org/10.1016/B978-0-444-81531-6.50016-7

Browne, M. W., & Du Toit, S. H. C. (1991). *Models for learning data* (pp. 47–68). American Psychological Association. https://doi.org/10.1037/10099-004

Burchinal, M., & Appelbaum, M. I. (1991). Estimating individual developmental functions: Methods and their assumptions. *Child Development*, *62*(1), 23. https://doi.org/10.2307/1130702

Chen, C. X., Martin, M., & Merchant, K. A. (2014). The effect of measurement timing on the information content of customer satisfaction measures. *Management Accounting Research*, *25*(3), 187–205. https://doi.org/10.1016/j.mar.2013.12.003

Choe, C. H. (2005). *Computational ergodic theory*. Springer-Verlag. https://doi.org/10.1007/b138894

Cohen, A. (1993). Organizational commitment and turnover: A meta-analysis. *Academy of Management Journal*, *36*(5), 1140–1157. https://doi.org/10.2307/256650

Cole, D. A., & Maxwell, S. E. (2009). Statistical methods for risk-outcome research: Being sensitive to longitudinal structure. *Annual Review of Clinical Psychology*, *5*(1), 71–96. https://doi.org/10.1146/annurev-clinpsy-060508-130357

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 558–577. https://doi.org/10.1037/0021-843x.112.4.558

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, *57*(1), 505–528. https://doi.org/10.1146/annurev.psych.57.102904.190146

Coulombe, P. (2016). *Partially and fully time-unstructured residual variance-covariance matrices in growth curve modeling: Consequences of ignoring variability in times of assessment* (Publication No. 10155460). [Doctoral dissertation, University of New Mexico]; ProQuest Dissertations and Theses Global.

Coulombe, P., Selig, J. P., & Delaney, H. D. (2016). Ignoring individual differences in times of assessment in growth curve modeling. *International Journal of Behavioral Development*, *40*(1), 76–86. https://doi.org/10.1177/0165025415577684

Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, *58*(1), 615–637. https://doi.org/10.1146/annurev.psych.58.110405.085520

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356

Dalal, R. S., Bhave, D. P., & Fiset, J. (2014). Within-person variability in job performance. *Journal of Management*, *40*(5), 1396–1436. https://doi.org/10.1177/0149206314532691

Day, D. V., & Sin, H.-P. (2011). Longitudinal tests of an integrative model of leader development: Charting and understanding developmental trajectories. *The Leadership Quarterly*, *22*(3), 545–560. https://doi.org/10.1016/j.leaqua.2011.04.011

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.

Dormann, C., & Griffin, M. A. (2015). Optimal time lags in panel studies. *Psychological Methods*, *20*(4), 489–505. https://doi.org/10.1037/met0000041

Dormann, C., & Ven, B. van de. (2014). Timing in methods for studying psychosocial factors at work. In *Psychosocial factors at work in the asia pacific* (1st ed., pp. 89–116). Springer Dordrecht. https://doi.org/10.1007/978-94-017-8975-2_4

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989x.12.2.121

Finch, W. H. (2017). Investigation of parameter estimation accuracy for growth curve modeling with categorical indicators. *Methodology*, *13*(3), 98–112. https://doi.org/10.1027/1614-2241/a000134

Fine, K. L., & Grimm, K. J. (2020). Examination of nonlinear and functional mixed-effects models with nonparametrically generated data. *Multivariate Behavioral Research*, 1–18. https://doi.org/10.1080/00273171.2020.1754746

Fine, K. L., Suk, H. W., & Grimm, K. J. (2019). An examination of a functional mixed-effects modeling approach to the analysis of longitudinal data. *Multivariate Behavioral Research*, *54*(4), 475–491. https://doi.org/10.1080/00273171.2018.1520626

Fisher, C. D. (2008). What if we took within-person variability seriously? *Industrial and Organizational Psychology*, *1*(2), 185–189. https://doi.org/10.1111/j.1754-9434.2008.00036.x

Fisher, J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27). https://doi.org/10.1073/pnas.1711978115

Fournier, M., d'Arripe-Longueville, F., Rovere, C., Easthope, C. S., Schwabe, L., El Methni, J., & Radel, R. (2017). Effects of circadian cortisol on the development of a health habit. *Health Psychology*, *36*(11), 1059–1064. https://doi.org/10.1037/hea0000510

Fried, Y., & Slowik, L. H. (2004). Enriching goal-setting theory with time: An integrated approach. *Academy of Management Review*, *29*(3), 404–422. https://doi.org/10.5465/amr.2004.13670973

Gasimova, F., Robitzsch, A., Wilhelm, O., & Hülür, G. (2014). A Hierarchical bayesian model with correlated residuals for investigating stability and change in intensive longitudinal data settings. *Methodology*, *10*(4), 126–137. https://doi.org/10.1027/1614-2241/a000083

George, J. M., & Jones, G. R. (2000). The role of time in theory and theory building. *Journal of Management*, *26*(4), 657–684. https://doi.org/10.1177/014920630002600404

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management*, *26*(3), 463–488. https://doi.org/10.1177/014920630002600305

Grimm, K., & Widaman, K. (2010). Residual structures in latent growth curve modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(3), 424–442. https://doi.org/10.1080/10705511.2010.489006

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science*, *16*(1), 70–76. https://doi.org/10.1111/j.0956-7976.2005.00782.x

Hamaker, Ellen L. (2012). Why researchers should think" within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of studying research methods for daily life* (pp. 43–61). The Guilford Press. shorturl.at/lmGX9

Hamaker, Ellen L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research*, *40*(2), 207–233. https://doi.org/10.1207/s15327906mbr4002_3

Hom, P. W., Caranikas-Walker, F., Prussia, G. E., & Griffeth, R. W. (1992). A meta-analytical structural equations analysis of a model of employee turnover. *Journal of Applied Psychology*, *77*(6), 890–909. https://doi.org/10.1037/0021-9010.77.6.890

Howell, D. C. (2009). *Statistical methods for psychology* (7th ed.). Wadsworth Publishing.

Huh, D., Kaysen, D. L., & Atkins, D. C. (2015). Modeling cyclical patterns in daily college drinking data with many zeroes. *Multivariate Behavioral Research*, *50*(2), 184–196. https://doi.org/10.1080/00273171.2014.977433

Igic, I., Keller, A. C., Elfering, A., Tschan, F., Kälin, W., & Semmer, N. K. (2017). Ten-year trajectories of stressors and resources at work: Cumulative and chronic effects on health and well-being. *Journal of Applied Psychology*, *102*(9), 1317–1343. https://doi.org/10.1037/apl0000225

Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science*, *25*(2), 334–339. https://doi.org/10.1177/0956797613502676

Kraft, D. (1994). Algorithm 733: TOMP–Fortran modules for optimal control calculations. *ACM Transactions on Mathematical Software*, *20*(3), 262–281. https://doi.org/10.1145/192115.192124

Kunisch, S., Bartunek, J. M., Mueller, J., & Huy, Q. N. (2017). Time in strategic change research. *Academy of Management Annals*, *11*(2), 1005–1064. https://doi.org/10.5465/annals.2015.0133

Larsen, R. J., & Kasimatis, M. (1990). Individual differences in entrainment of mood to the weekly calendar. *Journal of Personality and Social Psychology*, *58*(1), 164–171. https://doi.org/10.1037/0022-3514.58.1.164

Lawrence, T. B., Winn, M. I., & Jennings, P. D. (2001). The temporal dynamics of institutionalization. *Academy of Management Review*, *26*(4), 624–644. https://doi.org/10.5465/amr.2001.5393901

Liu, J., & Perera, R. A. (2021). Estimating knots and their association in parallel bilinear spline growth curve models in the framework of individual measurement occasions. *Psychological Methods*. https://doi.org/10.1037/met0000309

Liu, J., Perera, R. A., Kang, L., Kirkpatrick, R. M., & Sabo, R. T. (2019). Obtaining interpretable parameters from reparameterizing longitudinal models: Transformation matrices between growth factors in two parameter-spaces. *arXiv Preprint arXiv:1911.09939*.

Liu, J., Perera, R. A., Kang, L., Sabo, R. T., & Kirkpatrick, R. M. (2022). Obtaining interpretable parameters from reparameterized longitudinal models: Transformation matrices between growth Factors in two parameter spaces. *Journal of Educational and Behavioral Statistics*, *47*(2), 167–201. https://doi.org/10.3102/

10769986211052009

Liu, Y., Liu, H., Li, H., & Zhao, Q. (2015). The effects of individually varying times of observations on growth parameter estimations in piecewise growth model. *Journal of Applied Statistics*, *42*(9), 1843–1860. https://doi.org/10.1080/02664763.2015.1014884

Loehlin, J. C., & Beaujean, A. A. (2017). *Latent variable models*. Routledge. https://doi.org/10.4324/9781315643199

Magee, C. A., Caputi, P., & Lee, J. K. (2016). Distinct longitudinal patterns of absenteeism and their antecedents in full-time Australian employees. *Journal of Occupational Health Psychology*, *21*(1), 24–36. https://doi.org/10.1037/a0039138

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*(1), 23–44. https://doi.org/10.1037/1082-989x.12.1.23

Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research*, *46*(5), 816–841. https://doi.org/10.1080/00273171.2011.606716

Medaglia, J. D., Jeronimus, B. F., & Fisher, A. J. (2019). Reply to Adolf and Fried: Conditional equivalence and imperatives for person-level science. *Proceedings of the National Academy of Sciences*, *116*(14), 6542–6543. https://doi.org/10.1073/pnas.1820221116

Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, *10*(3), 259–284. https://doi.org/10.

1037/1082-989x.10.3.259

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*(1), 23–43. https://doi.org/10.1037/1082-989x.5.1.23

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*(1), 107–122. https://doi.org/10.1007/bf02294746

Mill, J. S. (2011). Of the law of universal causation. In *A system of logic, tatiocinative and inductive: Being a connected view of the principles of evidence, and the methods of scientific investigation* (Vol. 1, pp. 392–424). Cambridge University Press. (Original work published in 1843). https://doi.org/10.1017/cbo9781139149839.021

Miller, M. L., & Ferrer, E. (2017). The effect of sampling-time variation on latent growth curve models. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(6), 831–854. https://doi.org/10.1080/10705511.2017.1346476

Miraglia, M., Alessandri, G., & Borgogni, L. (2015). Trajectory classes of job performance. *Career Development International*, *20*(4), 424–442. https://doi.org/10.1108/cdi-03-2015-0032

Mitchell, M. A., & Maxwell, S. E. (2013). A comparison of the cross-sectional and sequential designs when assessing longitudinal mediation. *Multivariate Behavioral Research*, *48*(3), 301–339. https://doi.org/10.1080/00273171.2013.784696

Mitchell, T. R., & James, L. R. (2001). Building better theory: Time and the specification of when things happen. *Academy of Management Review*, *26*(4), 530–547. https://doi.org/10.5465/amr.2001.5393889

Mo Wang, & Bodner, T. E. (2007). Growth mixture modeling. *Organizational Research Methods*, *10*(4), 635–656. https://doi.org/10.1177/1094428106289397

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspective*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Molenaar, P. C. M. (2008). *Consequences of the ergodic theorems for classical test theory, factor analysis, and the analysis of developmental processes* (pp. 90–104). SAGE Publications. https://doi.org/10.4135/9781412976589.n5

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*(2), 112–117. https://doi.org/10.1111/j.1467-8721.2009.01619.x

Murphy, D. L., Beretvas, S. N., & Pituch, K. A. (2011). The effects of autocorrelation on the curve-of-factors growth model. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(3), 430–448. https://doi.org/10.1080/10705511.2011.582399

Murre, J. M. J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLOS ONE*, *10*(7), e0120644. https://doi.org/10.1371/journal.pone.0120644

Muthén, B. O. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. *Psychometrika*.

Navarro, J., Roe, R. A., & Artiles, M. I. (2015). Taking time seriously: Changing practices and perspectives in work/organizational psychology. *Journal of Work and*

Organizational Psychology, *31*(3), 135–145. https://doi.org/10.1016/j.rpto.2015.07.002

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*(3), 263–280. https://doi.org/10.1002/ejsp.2420150303

Newman, D. A. (2008). *Missing data techniques and low response rates: The role of systematic nonresponse parameters* (C. E. Lance & R. J. Vandenberg, Eds.; pp. 7–36). Routledge. https://doi.org/10.4324/9780203867266

Newsom, J. T., & Smith, N. A. (2020). Performance of latent growth curve models with binary variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(6), 888–907. https://doi.org/10.1080/10705511.2019.1705825

Nixon, A. E., Mazzola, J. J., Bauer, J., Krueger, J. R., & Spector, P. E. (2011). Can work make you sick? A meta-analysis of the relationships between job stressors and physical symptoms. *Work & Stress*, *25*(1), 1–22. https://doi.org/10.1080/02678373.2011.569175

O'Laughlin, K. D., Martin, M. J., & Ferrer, E. (2018). Cross-sectional analysis of longitudinal mediation processes. *Multivariate Behavioral Research*, *53*(3), 375–402. https://doi.org/10.1080/00273171.2018.1454822

O'Rourke, H. P., Fine, K. L., Grimm, K. J., & MacKinnon, D. P. (2021). The Importance of time metric precision when implementing bivariate latent change score ,odels. *Multivariate Behavioral Research*, 1–19. https://doi.org/10.1080/00273171.2021.1874261

Okada, K. (2013). Is omega squared less biased? A Ccmparison of three major effect size indices in one-way anova. *Behaviormetrika*, *40*(2), 129–147. https://doi.org/10.2333/bhmk.40.129

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. https://doi.org/10.1037/h0043424

Ostroff, C., Kinicki, A. J., & Clark, M. A. (2002). Substantive and operational issues of response bias across levels of analysis: An example of climate-satisfaction relationships. *Journal of Applied Psychology*, *87*(2), 355–368. https://doi.org/10.1037/0021-9010.87.2.355

Pan, B. (2010). Online travel surveys and response patterns. *Journal of Travel Research*, *49*(1), 121–135. https://doi.org/10.1177/0047287509336467

Petersen, K. E. (1983). *Ergodic theory*. Cambridge University Press. https://doi.org/10.1017/cbo9780511608728

Ployhart, R. E., & Vandenberg, R. J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, *36*(1), 94–120. https://doi.org/10.1177/0149206309352110

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Preacher, K. J., & Hancock, G. R. (2015). Meaningful aspects of change as novel random coefficients: A general method for reparameterizing longitudinal models. *Psycho-*

logical Methods, *20*(1), 84–101. https://doi.org/10.1037/met0000028

Preacher, K., Wichman, A., MacCallum, R., & Briggs, N. (2008). *Latent growth curve modeling*. SAGE Publications. https://doi.org/10.4135/9781412984737

Ram, N., Brose, A., & Molenaar, P. C. M. (2013). *Dynamic factor analysis: Modeling person-specific process* (T. D. Little, Ed.; Vol. 2, pp. 441–457). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199934898.013.0021

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed., Vol. 1). SAGE Publications. shorturl.at/imFN7

Riketta, M. (2008). The causal relation between job attitudes and performance: A meta-analysis of panel studies. *Journal of Applied Psychology*, *93*(2), 472–481. https://doi.org/10.1037/0021-9010.93.2.472

Robert, C., & Casella, G. (2010). *Introducing monte carlo methods with r* (1st ed.). Springer New York. https://doi.org/10.1007/978-1-4419-1576-4

Roe, R. A. (2008). Time in applied psychology. *European Psychologist*, *13*(1), 37–52. https://doi.org/10.1027/1016-9040.13.1.37

Roe, R. A. (2014). Test validity from a temporal perspective: Incorporating time in validation research. *European Journal of Work and Organizational Psychology*, *23*(5), 754–768. https://doi.org/10.1080/1359432x.2013.804177

Roe, R. A., Gockel, C., & Meyer, B. (2012). Time and change in teams: Where we are and where we are moving. *European Journal of Work and Organizational Psychology*, *21*(5), 629–656. https://doi.org/10.1080/1359432x.2012.729821

RStudio Team. (2020). *RStudio: Integrated development environment for r*. RStudio, PBC. http://www.rstudio.com/

Schoot, R. van de, Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

Shapiro, A., & Browne, M. W. (1987). Analysis of covariance structures under elliptical distributions. *Journal of the American Statistical Association*, *82*(400), 1092–1097. https://doi.org/10.1080/01621459.1987.10478544

Shipp, A. J., & Cole, M. S. (2015). Time in individual-level organizational studies: What is it, how is it used, and why isn't it exploited more often? *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 237–260. https://doi.org/10.1146/annurev-orgpsych-032414-111245

Siegler, R. S. (2013). *Microgenetic analyses of learning* (D. Kuuhn, R. S. Siegler, W. Damon, & R. M. Lerner, Eds.; 6th ed., Vol. 2, pp. 464–510). John Wiley & Sons. shorturl.at/AHRT6

Sonnentag, S. (2012). Time in organizational research: Catching up on a long neglected topic in order to improve theory. *Organizational Psychology Review*, *2*(4), 361–368. https://doi.org/10.1177/2041386612442079

Spector, P. E. (2019). Do not cross me: Optimizing the use of cross-sectional designs. *Journal of Business and Psychology*, *34*(2), 125–137. https://doi.org/10.1007/s10869-018-09613-8

Stanley, D. J., & Spence, J. R. (2014). Expectations for eeplications. *Perspectives on Psychological Science*, *9*(3), 305–318. https://doi.org/10.1177/1745691614528518

Steel, R. P., Hendrix, W. H., & Balogh, S. P. (1990). Confounding effects of the turnover base rate on relations between time lag and turnover study outcomes: An extension of meta-analysis findings and conclusions. *Journal of Organizational Behavior*, *11*(3), 237–242. https://doi.org/10.1002/job.4030110306

Steel, R. P., & Ovalle, N. K. (1984). A review and meta-analysis of research on the relationship between behavioral intentions and employee turnover. *Journal of Applied Psychology*, *69*(4), 673–686. https://doi.org/10.1037/0021-9010.69.4.673

Stockdale, G. D. (2007). *Factors affecting goodness of fit of the quasi-simplex, linear growth curve, and latent difference score models to oppositive data structures: A simulation study* (Publication No. 3303209). [Doctoral dissertation, University of California]; ProQuest Dissertations and Theses Global.

Tiberio, S. S. (2008). *The effects of misspecified measurement intervals in multivariate latent differential equation models* (Publication No. 3441759). [Doctoral dissertation, University of Notre Dame]; ProQuest Dissertations and Theses Global.

Timmons, A. C., & Preacher, K. J. (2015). The importance of temporal design: How do measurement intervals affect the accuracy and efficiency of parameter estimates in longitudinal research? *Multivariate Behavioral Research*, *50*(1), 41–55. https://doi.org/10.1080/00273171.2014.961056

Vantilborgh, T., Hofmans, J., & Judge, T. A. (2018). The time has come to study dynamics at work. In *Journal of Organizational Behavior* (No. 9; Vol. 39, pp. 1045–1049). Wiley Online Library. https://doi.org/10.1002/job.2327

Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a unified framework for the study of between-person and within-person structures: Building a bridge between two research paradigms. *Multivariate Behavioral Research*, *49*(3), 193–213. https://doi.org/10.1080/00273171.2014.889593

Wang, L. (Peggy)., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63–83. https://doi.org/10.1037/met0000030

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, *14*(4), 490–495. https://doi.org/10.5465/amr.1989.4308371

Wu, J.-Y., Kwok, O.-M., & Willson, V. (2014). Using design-based latent growth curve modeling with cluster-level predictor to address dependency. *The Journal of Experimental Education*, *82*(4), 431–454. https://doi.org/10.1080/00220973.2013.876226

Wu, W., Jia, F., Kinai, R., & Little, T. D. (2016). Optimal number and allocation of data collection points for linear spline growth curve modeling. *International Journal of Behavioral Development*, *41*(4), 550–558. https://doi.org/10.1177/0165025416644076

Xia, W., Ye, M., Liu, J., Cao, M., & Sun, X.-M. (2020). Analysis of a nonlinear opinion dynamics model with biased assimilation. *Automatica*, *120*, 109113. https://doi.org/10.1016/j.automatica.2020.109113

Ye, F. (2016). Latent growth curve analysis with dichotomous items: Comparing four approaches. *British Journal of Mathematical and Statistical Psychology*, *69*(1), 43–61. https://doi.org/10.1111/bmsp.12058

Yigit, S., & Mendes, M. (2018). Which effect size measure is appropriate for one-way and two-way ANOVA models? A Monte Carlo simulation study. *Statistical Journal*, *16*(3), 295–313. https://doi.org/10.2333/bhmk.40.129

Zaheer, S., Albert, S., & Zaheer, A. (1999). Time scales and organizational theory. *Academy of Management Review*, *24*(4), 725–741. https://doi.org/10.5465/amr.

1999.2553250

Appendix A: Technical Appendix A: Ergodicity and the Need to Conduct Longitudinal Research

To understand why cross-sectional results are unlikely to agree with longitudinal results for any given analysis, a discussion of data structure is apropos. Consider an example where a researcher obtains data from 50 people measured over 100 time points such that each row contains a p person's data over the 100 time points and each column contains data from 50 people at a t time point. For didactic purposes, all data are assumed to be sampled from a normal distribution. To understand whether findings in any given cross-sectional data set yield the same findings in any given longitudinal data set, the researcher randomly samples one cross-sectional and one longitudinal data set and computes the mean and variance in each set. To conduct a cross-sectional analysis, the researcher randomly samples the data across the 50 people at a given time point and computes a mean of the scores at the sampled time point ($\bar{X}_t$) using Equation A.1 shown below:

$$\bar{X}_t = \frac{1}{P}\sum_{p=1}^{P} x_p, \tag{A.1}$$

where the scores of all P people are summed (x_p) and then divided by the number of people (P). To compute the variance of the scores at the sampled time point (S_t^2), the researcher uses Equation A.2 shown below:

$$\S_t^2 = \frac{1}{P}\sum_{p=1}^{P}(x_p - \bar{X}_t)^2, \tag{A.2}$$

where the sum of squared differences between each person's score (x_p) and the average value at the given t time point ($\bar{X}_t$) is computed and then divided by the number of people (P). To conduct a longitudinal analysis, the researcher randomly samples the data across the 100 time points for a given person and also computes a mean and variance of the scores. To compute the mean across the t time points of the longitudinal data set ($\bar{X}_p$), the researcher uses Equation A.3 shown below:

$$\bar{X}_p = \frac{1}{T}\sum_{t=1}^{T} x_t, \tag{A.3}$$

where the scores at each t time point are summed (x_t) and then divided by the number of time points (T). The researcher also computes a variance of the sampled person's scores across all time points (S_p^2) using Equation A.4 shown below:

$$\S_p^2 = \frac{1}{T}\sum_{t=1}^{T} (x_t - \bar{X}_p)^2, \tag{A.4}$$

where the sum of squared differences between the score at each time point (x_t) and the average value of the p person's scores ($\bar{X}_p$) is computed and then divided by the number of time points (T).

If the researcher wants treat the mean and variance values computed from the cross-sectional and longitudinal data sets as interchangeable, then two conditions outlined by ergodic theory must be satisfied (Molenaar, 2004; Molenaar & Campbell, 2009).[21] First, a given cross-sectional mean and variance can only closely estimate the mean and variance

[21] Note that ergodic theory is an entire mathematical discipline (for an introduction, see Petersen, 1983). In the current context, the most important ergodic theorems are those proven by Birkhoff (1931; for a review, see Choe, 2005, Chapter 3).

of any given person's data (i.e., a longitudinal data set) to the extent that each person's data are generated from a normal distribution with the same mean and variance. If each person's data were generated from a different normal distribution, the computing the mean and variance at a given time point would, at best, describe the values of one person. When each person's data are generated from the same normal distribution, the condition of *homogeneity* is met. Importantly, satisfying the condition of homogeneity does not guarantee that the mean and variance obtained from another cross-sectional data set will closely estimate the mean and variance of any given person (i.e., any given longitudinal data set). The mean and variance values computed from any given cross-sectional data set can only closely estimate the values of any given person to the extent that the cross-sectional mean and variance remain constant over time. If the mean and variance of observations remain constant over time, then the the second condition of *stationarity* is satisfied. Therefore, the researcher can only treat means and variances from cross-sectional and longitudinal data sets as interchangeable if each person's data is generated from the same normal distribution (homogeneity) and if the mean and variance remain constant over time (stationarity). When the conditions of homogeneity and stationarity are satisfied, a process is said to be *ergodic*: Analyses of cross-sectional data sets return the same values as analyses on longitudinal data sets.

Given that psychological studies almost never collect data from only one person, one potential reservation may be that the conditions required for ergodicity only hold when a longitudinal data set contains the data of one person. That is, if the researcher used the full data set containing the data of 100 people sampled over 100 time points and computed 100 cross-sectional means and variances (Equation A.1 and Equation A.2, respectively) and 100 longitudinal means and variances (Equation A.3 and Equation A.4, respectively), wouldn't the average of the cross-sectional means and variances be the same as the average of the longitudinal means and variances? Although the averaging the cross-sectional mean returns the same value as averaging the longitudinal means, the average longitudinal variance remains different from the average cross-sectional variance (for several empirical examples, see J. Fisher et al., 2018). Therefore, the conditions of ergodicity apply even with larger longitudinal and cross-sectional sample sizes.

The guaranteed differences in cross-sectional and longitudinal variance values that result from non-ergodic processes have far-reaching implications. Almost every analysis employed in organizational research—whether it be correlation, regression, factor analysis, mediation analysis, etc.—analyzes variability, and so, when a process is non-ergodic, cross-sectional variability will differ from longitudinal variability, and the results obtained from applying any given analysis on each of the variabilities will differ as a consequence. Because variability is central to so many analyses, the non-equivalence of longitudinal and cross-sectional variances that results from a non-ergodic process explains why discussions of ergodicity often comment that "for non-ergodic processes, an analysis of the structure of IEV [interindividual variability] will yield results that differ from results obtained in an analogous analysis of IAV [intraindividual variability]"(Molenaar, 2004, p. 202).[22]

With an understanding of the conditions required for ergodicity, a brief consid-

[22]It is important to note that a violation of one or both ergodic conditions (homogeneity and stationarity) does not mean that an analysis of cross-sectional variability yields results that have no relation to the results gained from applying the analysis on longitudinal variability (i.e., the causes of cross-sectional variability are independent from the causes of longitudinal variability). An analysis of cross-sectional variability can still give insight into temporal dynamics if the causes of non-ergodicity can be identified (Voelkle et al., 2014; for similar discussion, see Spector, 2019). Thus, conceptualizing ergodicity on a continuum with non-erdogicity and ergodicity on opposite ends provides a more balanced perspective for understanding ergodicity (Adolf & Fried, 2019; Medaglia et al., 2019).

eration of organizational phenomena finds that these conditions are regularly violated. Focusing only on homogeneity (each person's data are generated from the same distribution), several instances in organizational research violate this condition. As examples of homogeneity violations, employees show different patterns of absenteeism over five years (Magee et al., 2016), leadership development over the course of a seminar (Day & Sin, 2011), career stress over the course of 10 years (Igic et al., 2017), and job performance in response to organizational restructuring (Miraglia et al., 2015). With respect to stationarity (constant values for statistical parameters across people over time), several examples can be generated by realizing how calendar events affect psychological processes and behaviours throughout the year. As examples of stationarity violations, consider how salespeople, on average, undoubtedly sell more products during holidays, how employees, on average, take more sick days during the winter months, and how accountants, on average, experience more stress during tax season. With ergodic condition violations commonly occurring in organizational psychology, it becomes fitting to echo the commonly held sentiment that few, if any, psychological processes are ergodic (Curran & Bauer, 2011; J. Fisher et al., 2018; Ellen L. Hamaker, 2012; Molenaar, 2004, 2008; Molenaar & Campbell, 2009; Wang & Maxwell, 2015).

Appendix B: Technical Appendix B: Using Nonlinear Function in the Structural Equation Modelling Framework

B.1 Nonlinear Latent Growth Curve Model Used to Analyze Each Generated Data Set

The sections that follow will first review the framework used to build latent growth curve models and then explain how nonlinear functions can be modified to fit into this framework.

B.1.1 Brief Review of the Latent Growth Curve Model Framework

The latent growth curve model proposed by Meredith & Tisak (1990) is briefly reviewed here (for a review, see K. Preacher et al., 2008). Consider an example where data are collected at five time points ($T = 5$) to yield five observations for each p person ($\mathbf{y_p} = [y_1, y_2, y_3, y_4, y_5$). A simple model to fit is one where change over is defined by a straight line and each person's pattern of change is some variation of this straight line. In modelling parlance, an intercept-slope model is fit where both the intercept and slope are random effects whose values are allowed to vary for each person. Intercept and slope parameters can be algebraically represented by a two-column matrix that represents the effect of each parameter on the outcome variable y at each i time point. Because the effect of the intercept parameter is constant over time, a column of 1s is used to represent its effect. For the slope parameter, a pattern of linear growth can be specified filling the second column with a series of monotonically increasing numbers such as 0–4.[23]The matrix Λ below shows a two-column matrix that specifies the effects for an intercept and slope parameter:

[23] The set of numbers specified for the slope starts at zero because there is presumably no effect of any variable at the first time point.

$$\Lambda = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

To create a model that allows different linear patterns to be fit to each person's data, a weight can be applied to each column of Λ and each weight can vary across individuals.[24]That is, each p person's pattern of change is predicted with a unique set of weights in $\boldsymbol{\iota_p}$ that determines the extent to which each basis column of Λ contributes to that person's change over time. Discrepancies between the values predicted by $\Lambda\boldsymbol{\iota_p}$ and a person's observed scores across all five time points are stored in an error vector $\mathcal{E}_\mathbf{p}$. Thus, a person's observed data ($\mathbf{y_p}$) is constructed using the expression shown below in Equation B.1:

$$y_p = \Lambda\boldsymbol{\iota_p} + \mathcal{E}_\mathbf{p}. \tag{B.1}$$

Note that Equation B.1 defines the general structural equation modelling framework.

B.1.2 Fitting a Nonlinear Function in the Structural Equation Modelling Framework

Unfortunately, the logistic function of Equation 2.3—where each parameter was estimated as a fixed- and random-effect—could not be directly used in a latent growth

[24]The columns of Λ are called basis curves (Blozis, 2004) or basis functions (Browne, 1993; Meredith & Tisak, 1990) because each column specifies a particular component of change.

curve model because it would have violated the linear nature of the structural equation modelling framework (Equation B.1). Structural equation models only permit linear combinations—specifically, the products of matrix-vector and/or matrix-matrix multiplication—and so directly fitting a nonlinear function such as the logistic function in Equation 2.3 would not have been possible.

One solution to fitting the logistic function within the structural equation modelling framework was to implement the structured latent curve modelling approach (Browne, 1993; Browne & Du Toit, 1991; for an excellent review, see K. J. Preacher & Hancock, 2015). Briefly, the structured latent curve modelling approach constructs a Taylor series approximation of a nonlinear function so that the nonlinear function can be fit into the structural equation modelling framework (Equation B.1). The sections that follow will present the structured latent curve modelling approach in four parts such that 1) Taylor series approximations will first be reviewed, 2) a Taylor series approximation will then be constructed for the logistic function, 3) the logistic Taylor series approximation will be modified and fit into the structural equation modelling framework, and 4) the process of parameter estimation will be reviewed.

B.1.2.1 Taylor Series Approximations

A Taylor series uses derivative information of a nonlinear function to construct a linear approximation.[25]Equation B.2 shows the general formula for a Taylor series such

[25]Linear functions are defined as functions where no parameter exists within its own partial derivative. For example, none of the parameters in the polynomial equation of $y = a + bt + ct^2 + dt^3$ exist within their own partial derivative: $\frac{\partial y}{\partial a} = 1$, $\frac{\partial y}{\partial b} = t$, $\frac{\partial y}{\partial c} = t^2$, and $\frac{\partial y}{\partial d} = t^3$. Conversely, the logistic function is nonlinear because β and γ exist in their own partial derivatives. For example, the derivative of the logistic function $y = \theta + \frac{\alpha - \theta}{1 + e^{\frac{\beta - t}{\gamma}}}$ with respect to β is $\frac{(\theta - \alpha)(e^{\frac{\beta - t}{\gamma}})(\frac{1}{\gamma})}{1 + (e^{\frac{\beta - t}{\gamma}})^2}$ and so is nonlinear because it contains β.

that

$$P^N(f(x), a) = \sum_{n=0}^{N} \frac{f^n a}{n!}(x-a)^n, \tag{B.2}$$

where N is the highest derivative order of the function $f(a)$ that is taken beginning from a zero-value derivative order ($n = 0$), a is the point where the Taylor series is derived, and x is the point where the Taylor series is evaluated. As an example, consider $f(x) = \cos(x)$. Note that, across the continuum of x values (i.e., from $-\infty$ to ∞), $\cos(x)$ returns values between -1 and 1 in an oscillatory manner. Computing the second-order Taylor series approximation of $f(x) = \cos(x)$ yields the following function shown in Equation B.3:

$$\begin{aligned} P^2(\cos(x), a) &= \frac{\frac{\partial^0 \cos(a)}{\partial a^0}}{0!}(x-a)^0 + \frac{\frac{\partial^1 \cos(a)}{\partial a^1}}{1!}(x-a)^1 + \frac{\frac{\partial^2 \cos(a)}{\partial a^2}}{2!}(x-a)^2 \\ &= \frac{\cos(0)}{0!}(x-0)^0 - \frac{\sin(0)}{1!}(x-0)^1 - \frac{\cos(0)}{2!}(x-0)^2 \\ &= \frac{1}{1}1 - \frac{0}{1}x - \frac{1}{2}x^2 \\ P^2(\cos(x), 0) &= 1 - \frac{1}{2}x^2. \end{aligned} \tag{B.3}$$

Note that that the second-order Taylor series of $\cos(x)$ perfectly estimates $\cos(x)$ when the point of evaluation x is set equal to the point of derivation a and estimates $\cos(x)$ with an increasing amount of error as the difference between x and a increases (see Example .1).

Example .1. Estimates of Taylor series approximation of $f(x) = \cos(x)$ as the difference between the point of evaluation x and the point of derivation a increases.

Taylor series approximation of $f(x) = \cos(x)$ estimates values that are exactly equal to the values returned by $f(x) = \cos(x)$ when the point of evaluation x is set to the point of derivation a. The example below computes the value predicted by the Taylor series approximation of $f(x) = \cos(x)$ and by $f(x) = \cos(x)$ when $x = a = 0$.

$$
\begin{aligned}
P^2(\cos(x = 0), a = 0) &= \cos(x = 0) \\
1 - \frac{1}{2}x^2 &= \cos(0) \\
1 - \frac{1}{2}0^2 &= 1 \\
1 - 0 &= 1 \\
1 &= 1
\end{aligned}
$$

Taylor series approximation of $f(x) = \cos(x)$ estimates a value that is clearly not equal (neq) to the value returned by $f(x) = \cos(x)$when the difference between the point of evaluation x and the point of derivation a is smaller. The example below computes the value predicted by the Taylor series approximation of $f(x) = \cos(x)$ and by $f(x) = \cos(x)$ when $x = 1$ and $a = 0$.

$$
\begin{aligned}
P^2(\cos(x = 1), 0) &\approx \cos(x = 1) \\
1 - \frac{1}{2}x^2 &\approx \cos(1) \\
1 - \frac{1}{2}1^2 &\approx 0.54 \\
1 - 0.5 &\approx 0.54 \\
0.5 &\approx 0.54
\end{aligned}
$$

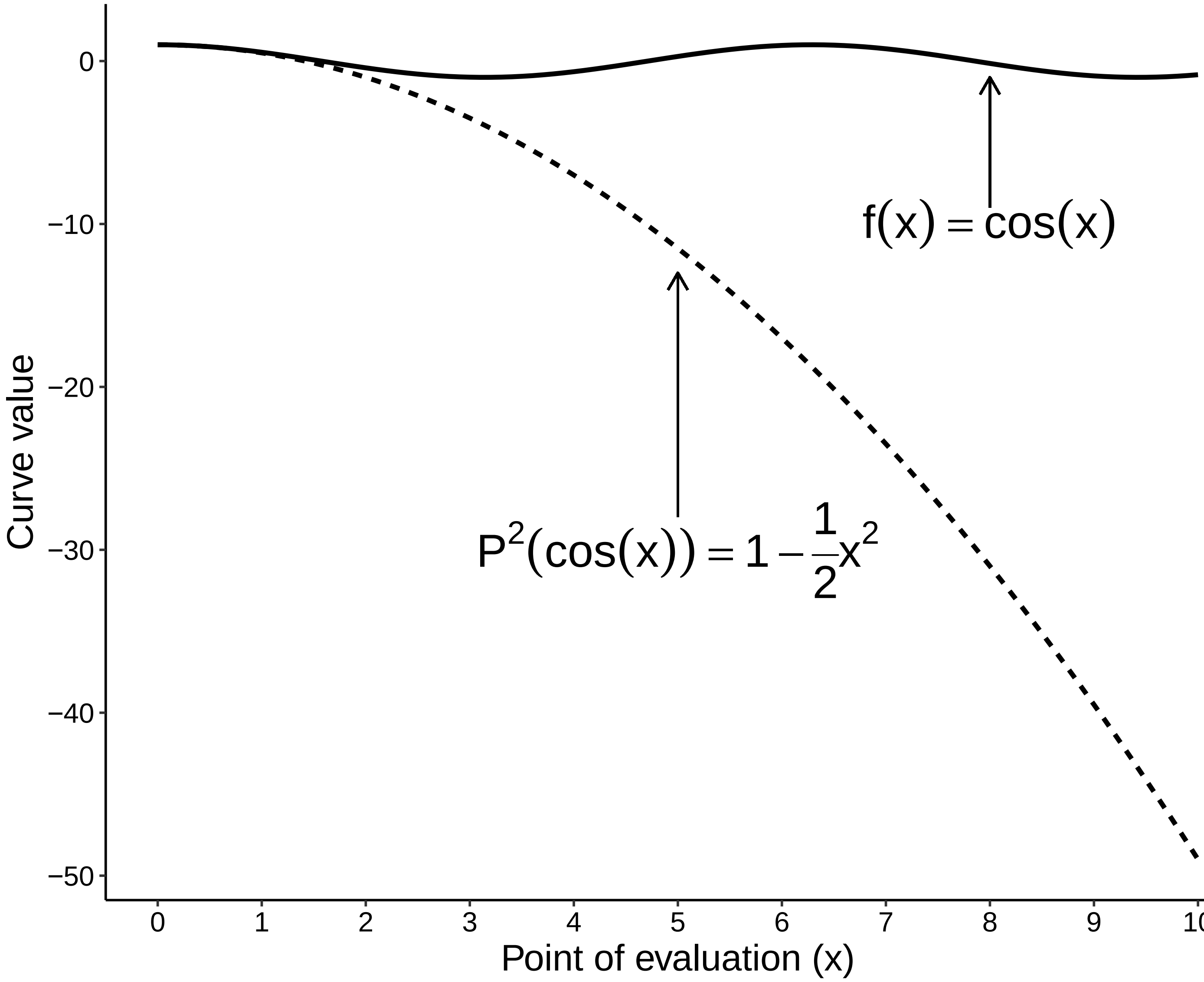
Taylor series approximation of $f(x) = \cos(x)$ estimates a value that is clearly not equal (neq) to the value returned by $f(x) = \cos(x)$ when the difference between the point of evaluation x and the point of derivation a is larger The example below computes the value predicted by the Taylor series approximation of $f(x) = \cos(x)$ and by $f(x) = \cos(x)$ when $x = 4$ and $a = 0$.

$$
\begin{aligned}
P^2(\cos(x = 4), 0) &\neq \cos(x = 4) \\
1 - \frac{1}{2}x^2 &\neq \cos(4) \\
1 - \frac{1}{2}4^2 &\neq -0.65 \\
1 - 16 &\neq -0.65 \\
0.5 &\neq -0.65
\end{aligned}
$$

Figure B.1 plots the nonlinear function of $\cos(x)$ and its second-order Taylor series $P^2(\cos(x)) = 1 - \frac{1}{2}x^2$. The second order Taylor series perfectly estimates $\cos(x)$ when the point of evaluation (x) equals the point of derivation (a; $x = a = 0$), but incurs an increasingly large amount of error as the difference between the point of evaluation and the point of derivation increases. For example, at $x = 10$, $\cos(10) = -0.84$, but the Taylor series outputs a value of -49.50 ($P^2(cos(50)) = 1 - \frac{1}{2}10^2 = -49.50$). Therefore, Taylor series' are approximations because they are locally accurate.

Figure B.1

Estimation Accuracy of Taylor Series Approximation of Nonlinear Function (cos(x))



Note. The second order Taylor series perfectly estimates $\cos(x)$ when the point of evaluation (x) equals the point of derivation (a; $x = a = 0$), but incurs an increasingly large amount of error as the difference between the point of evaluation and the point of derivation increases. For example, at $x = 10$, $\cos(x) = -0.84$, but the Taylor series outputs a value of -49.50 ($P^2(cos(50)) = 1 - \frac{1}{2}10^2 = -49.50$).

B.1.2.2 Taylor Series Approximation of the Logistic Function

Given that a Taylor series provides a linear approximation of a nonlinear function and the structural equation modelling framework is linear, the structured latent curve modelling approach uses Taylor series approximations to construct linear representations of nonlinear functions (Browne, 1993; Browne & Du Toit, 1991). In the current simulations, a Taylor series approximation was constructed for the logistic function (Equation B.5). Note that, because the logistic function had four parameters (θ, α, β, γ), derivatives were computed with respect to each of the parameters. Using a derivative order set to one ($n = 1$), the following Taylor series was constructed for the logistic function (Equation B.4):

$$P^1(L(\Theta, t)) = L + \frac{\partial L}{\partial \theta}(x_\theta - a_\theta)^1 + \frac{\partial L}{\partial \alpha}(x_\alpha - a_\alpha)^1 + \frac{\partial L}{\partial \beta}(x_\beta - a_\beta)^1 + \frac{\partial L}{\partial \gamma_\gamma}(x - a_\gamma)^1, \tag{B.4}$$

where $\mathbf{L}(\Theta, \mathbf{t})$ represents the logistic function shown below in Equation B.5:

$$\mathbf{L}(\Theta, \mathbf{t}) = \theta + \frac{\alpha - \theta}{1 + e^{\frac{\beta - t}{\gamma}}} + \epsilon, \tag{B.5}$$

with $\Theta = [\theta, \alpha, \beta, \gamma]$ and $\mathbf{L}(\Theta, \mathbf{t})$ being a vector of scores at all $\mathbf{t}$ time points. In the current context, because each parameter of the logistic function had a unique meaning (see section on data generation), the point of derivation a differed for each parameter—using the same a value for each parameter to construct the Taylor series approximation of the logistic function would have yielded a practically useless equation. Because the logistic Taylor series approximation (Equation B.4) was deployed in a statistical model (i.e., the structural equation modelling framework), the derivation values (a_θ, a_α, a_β, a_γ) were set to the mean values estimated by the analysis for each parameter. Thus, the derivation values were replaced with the following terms:

- $a_\theta = \hat{\theta}$
- $a_\alpha = \hat{\alpha}$
- $a_\beta = \hat{\beta}$
- $a_\gamma = \hat{\gamma}$

where that a caret ˆ indicates the mean value estimated for a parameter by the analysis. In order to estimate curves for each p person, the values of evaluation (x_θ, x_α, x_β, x_γ) corresponded to the parameter values computed for a given person (θ_p, α_p, β_p, γ_p). Thus, the evaluation values were replaced with the following terms:

- $x_\theta = \theta_p$
- $x_\alpha = \alpha_p$
- $x_\beta = \beta_p$
- $x_\gamma = \gamma_p$

Substituting the above values for the derivation and evaluation values of x and a in the initial logistic Taylor series approximation (Equation B.4) yielded the following expression for the logistic Taylor series approximation (Equation B.6):

$$P^1(L(\Theta, t)) = L(\Theta, t) + \frac{\partial L}{\partial \theta}(\theta_i - \hat{\theta})^1 + \frac{\partial L}{\partial \alpha}(\alpha_i - \hat{\alpha}_i)^1 + \frac{\partial L}{\partial \beta}(\beta - \hat{\beta})^1 + \frac{\partial L}{\partial \gamma_\gamma}(\beta - \hat{\beta})^1. \tag{B.6}$$

Therefore, because the Taylor series was derived using the mean values estimated for each parameter ($\hat{\theta}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$), it provided a perfect approximation of the estimated population curve—the evaluation values for each parameter would have been set to their corresponding mean estimated value. To estimate the curve of any given p person, the evaluation values could be offset from their corresponding derivation value (i.e., mean estimated value for a parameter) by using the set of parameter values computed for that person (θ_p, α_p, β_p, γ_p). Note that, because Taylor series approximations are only locally accurate, the predicted curves for any given p person become increasingly inaccurate curves as the difference between the derivation and evaluation values increases (e.g., $\theta_i - \hat{\theta}$).

B.1.2.3 Fitting the Logistic Taylor Series Approximation Into the Structual Equation Modelling Framework

Although the logistic Taylor series approximation provides an accurate estimation of the logistic function, the function in (Equation B.6) is modified in the structured latent curve modelling approach so that it can more effectively fit into the structural equation modelling framework (Equation B.1). The partial derivative information is stored in the matrix Λ such that

$$\Lambda = \begin{bmatrix} \frac{\partial L(\Theta,t_1)}{\partial\theta} & \frac{\partial L(\Theta,t_1)}{\partial\alpha} & \frac{\partial L(\Theta,t_1)}{\partial\beta} & \frac{\partial L(\Theta,t_1)}{\partial\gamma} \\ \frac{\partial L(\Theta,t_2)}{\partial\theta} & \frac{\partial L(\Theta,t_2)}{\partial\alpha} & \frac{\partial L(\Theta,t_2)}{\partial\beta} & \frac{\partial L(\Theta,t_2)}{\partial\gamma} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial L(\Theta,t_n)}{\partial\theta} & \frac{\partial L(\Theta,t_n)}{\partial\alpha} & \frac{\partial L(\Theta,t_n)}{\partial\beta} & \frac{\partial L(\Theta,t_n)}{\partial\gamma} \end{bmatrix}.$$

As in the structural equation modelling framework where each column of Λ specified a basis curve (i.e., a loading of a growth parameter onto all time points), each column of Λ here in the structured latent curve modelling approach contains the loadings of a logistic function parameter onto all the n time points, with the loadings being determined by the partial derivative of logistic function with respect to that parameter. To predict unique curves for each person, each column can be multiplied by a specific weight $\boldsymbol{\iota_p}$ that contains person-specific deviations from each mean estimated parameter value as shown below:

$$\boldsymbol{\iota_p} = \begin{bmatrix} \hat{\theta} - \theta_p \\ \hat{\alpha} - \alpha_p \\ \hat{\beta} - \beta_p \\ \hat{\gamma}_i - \gamma_p \end{bmatrix},$$

where a caret (ˆ) indicates the mean value estimated for a given parameter and a subscript p indicates a parameter value computed for a person. With a matrix Λ containing logistic function parameter loadings and a vector $\boldsymbol{\iota_p}$ containing person-specific weights, the Taylor series of Equation B.6 that predicted a person's scores over time can be rewritten to become the following expression of Equation B.7:

$$\mathbf{y_p} = \mathbf{L}(\Theta, \mathbf{t}) + \Lambda \boldsymbol{\iota_p} + \mathcal{E}_\mathbf{p}. \tag{B.7}$$

Importantly, because of the logistic function ($\mathbf{L}(\Theta, \mathbf{t})$) in the above expression (Equation B.7), the model no longer fits into the general structural equation modelling framework (Equation B.1). To modify Equation B.7 such that it fits into the structural equation modelling framework, the structured latent curve modelling approach recognizes that the logistic function ($\mathbf{L}(\Theta, \mathbf{t})$) is invariant under a scaling constant and uses this property to rewrite $\mathbf{L}(\Theta, \mathbf{t})$ as a weighted sum of the partial derivative loading matrix (Λ; Shapiro & Browne, 1987). Briefly, the logistic function vector $\mathbf{L}(\Theta, \mathbf{t})$ is invariant under a constant scaling property because, given some constant scalar value $k \geq 0$ and a set of parameter values (Θ), there exists another set of parameter values ($\tilde{\Theta}$) that can produce the same values (see Equation B.8 and Example .2 below).

$$k\mathbf{L}(\Theta, \mathbf{t}) = \mathbf{L}(\tilde{\Theta}, \mathbf{t}) \tag{B.8}$$

Example .2. Invariability under a constant scaling factor of logistic function (Equation B.5).

Given $t = [0, 1, 2, 3]$, $\Theta = [\theta = 3.00,\ \alpha = 3.32,\ \beta = 180.00,\ \gamma = 20.00]$, and some constant scaling factor $k = 2.00$, then there exists some set of parameter values $\tilde{\Theta}$ that produces the same values as $kL(\Theta)$. In the current example, $\tilde{\Theta} = [\theta = 6.00,\ \alpha = 6.64,\ \beta = 180.00,\ \gamma = 20.00]$.

$$\mathbf{kL}(\Theta, \mathbf{t}) = \mathbf{L}(\tilde{\Theta}, \mathbf{t})$$

$$2 * [3.00, 3.02, 3.30, 3.32] = [6.00, 6.04, 6.60, 6.64]$$

$$[6.00, 6.04, 6.60, 6.64] = [6.00, 6.04, 6.60, 6.64]1$$

If a function has the property of being invariant under a scaling factor, then it can also be expressed as the following matrix-vector product shown in Equation B.9 (Shapiro & Browne, 1987):

$$\mathbf{L}(\Theta, \mathbf{t}) = \boldsymbol{\Lambda}\tau, \tag{B.9}$$

where Λ contains the partial derivative loadings[26] and τ is a vector whose values are otbained by pre-multiplying the output of the logistic function ($\mathbf{L}(\Theta, \mathbf{t})$) by the inverse of the partial derivative loading matrix $\Lambda\tau^{-1}$. Solving for τ yields a vector whose contents contain the mean values estimated for parameters that enter the logistic function in a linear way and zeroes for parameters that enter the function in a nonlinear way (i.e., parameters that exist within their own partial derivative). Hence, τ is often called a mean vector (Blozis, 2004; K. J. Preacher & Hancock, 2015). In the current example, θ and α enter the logistic function in a linear way and β and γ enter the logistic function in a nonlinear way and so the first two entries of τ contain the values estimated for θ and α (i.e., $\hat{\theta}$ and $\hat{\alpha}$) and the last two entries contain zeroes. Example .3 below shows that the first two values of τ are indeed the values estimated for θ and α and the last two values are zero.

Example .3. Computation of mean vector τ.

Given the parameter estimates of $\hat{\theta} = 3.00$, $\hat{\alpha} = 3.32$, $\hat{\beta} = 180.00$, and $\hat{\gamma} = 20.00$ and $\mathbf{t}$

[26]This is also known as a Jacobian matrix.

$= [0,\ 1,\ 2,\ 3]$, $\tau = [3.00,\ 3.32,\ 0,\ 0]$, then

$$\mathbf{L}(\Theta, \mathbf{t}) = \mathbf{\Lambda}\tau$$

$$[3.00, 3.02, 3.30, 3.32] = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.95 & 0.05 & -0.00 & 0.00 \\ 0.05 & 0.95 & -0.00 & -0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 \end{bmatrix} \tau$$

$$\begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 \\ 0.95 & 0.05 & -0.00 & 0.00 \\ 0.05 & 0.95 & -0.00 & -0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 \end{bmatrix}^{-1} \begin{bmatrix} 3.00 \\ 3.02 \\ 3.30 \\ 3.32 \end{bmatrix} = \mathbf{\Lambda}\tau$$

$$\tau = [3.00, 3.32, 0, 0]$$

With $\mathbf{L}(\Theta, \mathbf{t}) = \Lambda\tau$, Equation B.7 can be rewritten in a linear equation as shown below in Equation B.10:

$$\mathbf{y_p} = \Lambda\tau + \Lambda\iota_\mathbf{p} + \mathcal{E}_\mathbf{p}. \tag{B.10}$$

The mean vector τ and vector of person-specific deviations $\iota_\mathbf{p}$ can be combined into a new vector $\mathbf{s_p}$ that represents the person-specific weights applied to the basis curves in Λ such that

$$\mathbf{s_p} = \tau + \iota_\mathbf{p} = \begin{bmatrix} \hat{\theta} + \hat{\theta} - \theta_p \\ \hat{\alpha} + \hat{\alpha} - \alpha_p \\ 0 + \hat{\beta} - \beta_p \\ 0 + \hat{\gamma} - \gamma_p \end{bmatrix}$$

and

$$\mathbf{y_p} = \Lambda \mathbf{s_p} + \mathcal{E}_\mathbf{p}. \tag{B.11}$$

Because the expected value of the person-specific weights ($\mathbf{s_p}$) is the mean vector (τ; $\mathbb{E}[\mathbf{s_p}] = \tau$, the expected set of scores predicted across all people ($\mathbb{E}[\mathbf{y_p}]$) gives back the original expression for the logistic function matrix-vector product in Equation B.9 as shown below in Equation B.12:

$$\mathbb{E}[\mathbf{y_p}] = \Lambda \tau = \mathbf{L}(\Theta, \mathbf{t}). \tag{B.12}$$

Therefore, the structured latent curve modelling approach successfully reproduces the output of the nonlinear logistic function (Equation B.5) with the linear function of Equation B.11. Note that that no error term exists in Equation B.12 because the expected value of the error values is zero ($\mathbb{E}[\mathcal{E}_\mathbf{p}] = 0$).

B.1.2.4 Estimating Parameters in the Structured Latent Curve Modelling Approach

To estimate parameter values, the full-information maximum likelihood shown in Equation B.13 was computed for each person (i.e., likelihood of observing a p person's data given the estimated parameter values):

$$\mathcal{L}_p = k_p \ln(2\pi) + \ln(|\boldsymbol{\Sigma}_\mathbf{p}| + (\mathbf{y_p} - \boldsymbol{\mu}_\mathbf{p})^\top {\boldsymbol{\Sigma}_\mathbf{p}}^{-1}(\mathbf{y_p} - \boldsymbol{\mu}_\mathbf{p}), \tag{B.13}$$

where k_p is the number of non-missing values for a given p person, $\boldsymbol{\Sigma}_\mathbf{p}$ is the model-implied covariance matrix with rows and columns filtered at time points where person p has missing data, $\mathbf{y_p}$ is a vector containing the data points that were collected for a p person (i.e., filtered data), and $\boldsymbol{\mu}_\mathbf{p}$ is the model-implied mean vector that is filtered at time points where person p has missing data. Note that, because all simulations assumed complete data across all times points, no filtering procedures were executed (for a review of the filtering procedure, see Boker et al., 2020, Chapter 5). Thus, computing the above full-information maximum likelihood in Equation B.13 was equivalent to computing the below likelihood function in Equation B.14:

$$\mathcal{L}_p = k_p \ln(2\pi) + \ln(|\boldsymbol{\Sigma}| + (\mathbf{y_p} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y_p} - \boldsymbol{\mu}), \tag{B.14}$$

where $\boldsymbol{\Sigma}$ is the model-implied covariance matrix, $\mathbf{y_p}$ contains the data collected from a p person, and $\boldsymbol{\mu}$ is the model-implied mean vector. The model-implied covariance matrix $\boldsymbol{\Sigma}$ is computed using Equation B.15 below:

$$\boldsymbol{\Sigma} = \Lambda\Psi\Lambda + \Omega_{\mathcal{E}}, \tag{B.15}$$

where Ψ is the random-effect covariance matrix and $\Omega_{\mathcal{E}}$ contains the error variances at each time point. The mean vector μ was computed using Equation B.16 shown below:

$$\mu = \Lambda\tau. \tag{B.16}$$

Parameter estimation was conducted by finding values for the model-implied covariance matrix $\mathbf{\Sigma}$ and the model-implied mean vector μ that maximized the sum of log-likelihoods across all P people (see Equation B.17 below):

$$\mathcal{L} = \underset{\mathbf{\Sigma},\mu}{\arg\max} \sum_{p=1}^{P} \mathcal{L}_p. \tag{B.17}$$

In OpenMx, the above problem was solved using the sequential least squares quadratic program (for a review, see Kraft, 1994).

Appendix C: Procedure for Generating Measurement Schedules Measurement SchedulesMeasurement Schedules

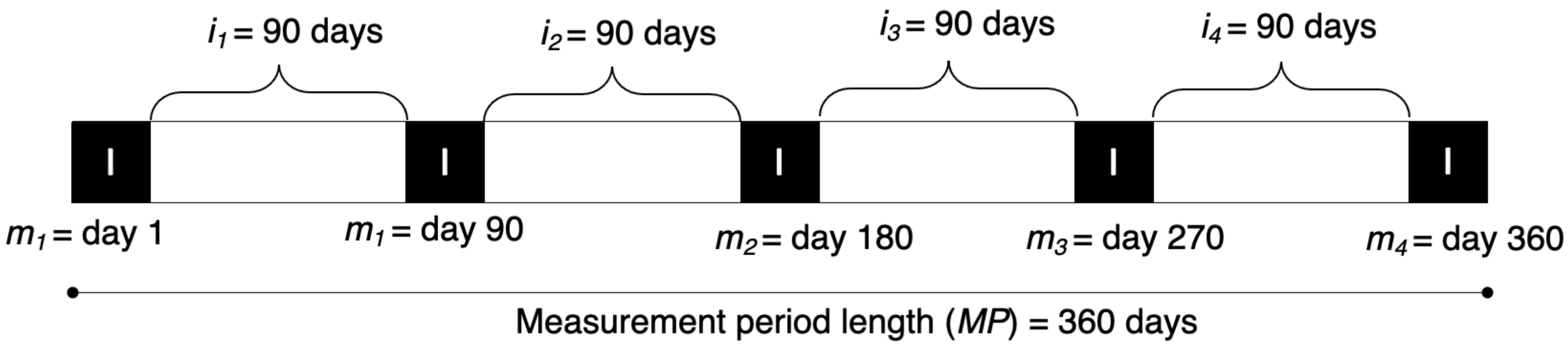
Given that no procedure existed (to my knowledge) for creating measurement schedules, I devised a method for generating measurement schedules for the four spacing conditions (equal, time-interval increasing, time-interval decreasing, and middle-and-extreme spacing). For each measurement spacing conditions across all measurement number levels, a two-step procedure was employed to generate measurement schedules in Experiments 2–3. At a broad level, the first step involved computing setup variables and the second step computed the interval lengths.

C.1 Procedure for Constructing Measurement Schedules With Equal Spacing

Figure @ref{fig:equal_spacing_diagram} shows how the two-step procedure was implemented to construct a measurement schedule with equal spacing and five measurements. In the first step, the number of intervals (NI) was computed by subtracting one from the number of measurements (NM), giving five measurements ($NM = 5$) and four intervals ($NI = 4$). In the second step, interval lengths were calculated by dividing the length of the measurement period (MP) by the number of intervals (NI), yielding an interval length of 90 days ($\frac{MP}{NI} = \frac{360}{4} = 90$).

Figure C.1

Procedure for Generating Equal Spacing Schedules With Equal Spacing



Setup variables (Step 1)

$\sum$ ■ = number of measurements (*NM*) = <u>5</u> measurements

$\sum$ □ = number of intervals (*NI*) = *NM – 1* = 4 intervals of ***x*** length

Interval calculation (Step 2)

Interval length (*I*) = $\frac{MP}{NI} = \frac{360}{4x} =$ <u>90</u> days

C.2 Procedure for Constructing Measurement Schedules With Time-Interval Increasing Spacing

Figure @ref{fig:time_inc_diagram} shows how the two-step procedure was used to calculate the interval lengths for measurement schedules defined by time-interval increasing spacing with five measurements. In the first step, the number of intervals was determined by subtracting one from the number of measurements, yielding a value of four for the number of intervals ($NI = NM - 1 = 5 - 1 = 4$). Importantly, the length of each interval increased by a constant value c over time as shown below in Equation C.1:

$$\text{Interval length} = x + \#IN(c) \tag{C.1}$$

where $\#IN$ represents the interval number in increasing order such that $\#IN \in \{0, ..., \text{number of interva}$ $1\}$. In the second step, the constant value by which interval lengths increased (c) was computed by first subtracting the smallest interval length from each interval (i.e., $x = 36$ days) from the measurement period (MP), yielding 216 remaining days ($N_{remain} = MP - NIx = 360 - 4(36) = 216$). The number of remaining days then had to be divided across the constant interval lengths. Because each interval increased by some constant value (c after each measurement point, the total number of constant-value interval lengths was obtained by computing the following sum in Equation C.2:

$$\text{Number constant intervals} = \sum_{i=0}^{\#IN} i. \tag{C.2}$$

With $\#IN = 3$, the number of constant intervals was 6, and so the constant value was obtained by using Equation C.3 below:

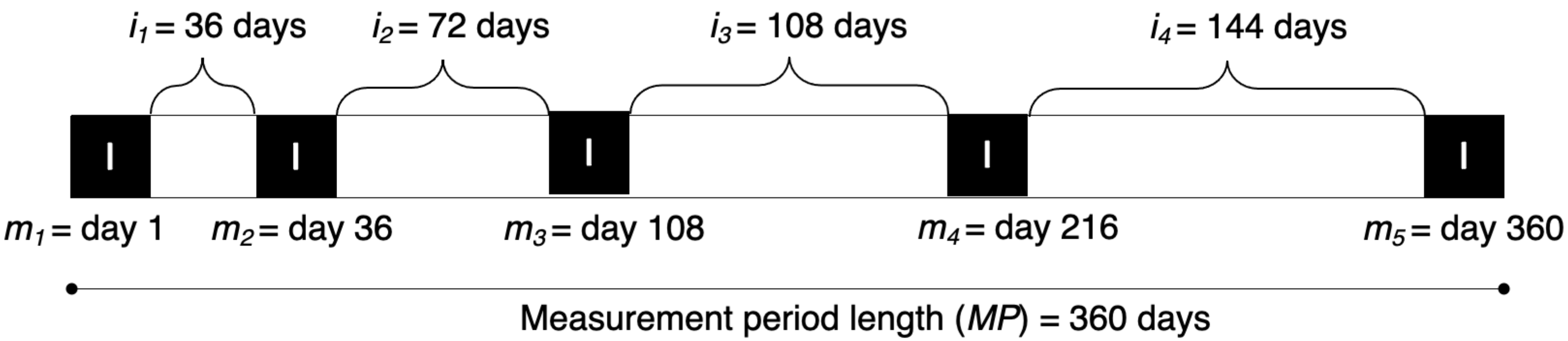
$$\text{Number constant intervals} = \frac{N_{remain}}{\sum_{i=0}^{\#IN} i}, \tag{C.3}$$

giving a length of 36 days for the constant value ($c = \frac{216}{6} = 36$ days). Having computed the value for c, the following interval lengths were obtained:

- $i_1 = x + \#IN(c) = 36 + 0(36) = 36$ days
- $i_2 = x + \#IN(c) = 36 + 1(36) = 72$ days
- $i_3 = x + \#IN(c) = 36 + 2(36) = 108$ days
- $i_4 = x + \#IN(c) = 36 + 3(36) = 144$ days

Figure C.2

Procedure for Generating Equal Spacing Schedules With Time-Interval Increasing Spacing



Setup variables (Step 1)

$\sum$ I = number of measurements (*NM*) = 5 measurements

$\sum$ ☐ = number of intervals (*NI*) = *NM* – *1* = 4 intervals of *x + #IN(c)* length

#IN = interval Number ∈ {0, … , number of intervals (NI) – 1}

c = constant by which interval lengths increase

Interval calculation (Step 2)

Number of remaining days (N_{remain}) = MP – NI(x) = 360 – 4(36) = 216 days

Constant length (*c*) = $\frac{N_{remain}}{\sum_{i=0}^{\#IN} i} = \frac{216}{(0+1+2+3)} =$ 36 days

$i_1 = x + \#IN(c) = 36 + 0(c) =$ **36 days**

$i_2 = x + \#IN(c) = 36 + 1(c) =$ **72 days**

$i_3 = x + \#IN(c) = 36 + 2(c) =$ **108 days**

$i_4 = x + \#IN(c) = 36 + 3(c) =$ **144 days**

C.3 Procedure for Constructing Measurement Schedules With Time-Interval Decreasing Spacing

Figure @ref{fig:time_dec_diagram} shows how the two-step procedure was used to calculate the interval lengths for measurement schedules defined by time-interval decreasing spacing with five measurements. In the first step, the number of intervals was determined by subtracting one from the number of measurements, yielding a value of four for the number of intervals ($NI = NM - 1 = 5 - 1 = 4$). Importantly, the length of each interval decreased by a constant value c over time as shown below in Equation C.4:

$$\text{Interval length} = x + \#INT(c) \tag{C.4}$$

where $\#IN$ represents the interval number in decreasing order such that $\#IN \in$ {number of intervals (N $1, ..., 0\}$. In the second step, the constant value by which interval lengths decreased (c) was computed by first subtracting the smallest interval length from each interval (i.e., $x = 36$ days) from the measurement period (MP), yielding 216 remaining days ($N_{remain} = MP - NIx = 360 - 4(36) = 216$). The number of remaining days then had to be divided across the constant interval lengths. Because each interval decreased by some constant value (c after each measurement point, the total number of constant-value interval lengths was obtained by computing the following sum in Equation C.5:

$$\text{Number constant intervals} = \sum_{\#IN}^{i=0} i. \tag{C.5}$$

With $\#IN = 3$, the number of constant intervals was 6, and so the constant value was obtained by using Equation C.6 below:

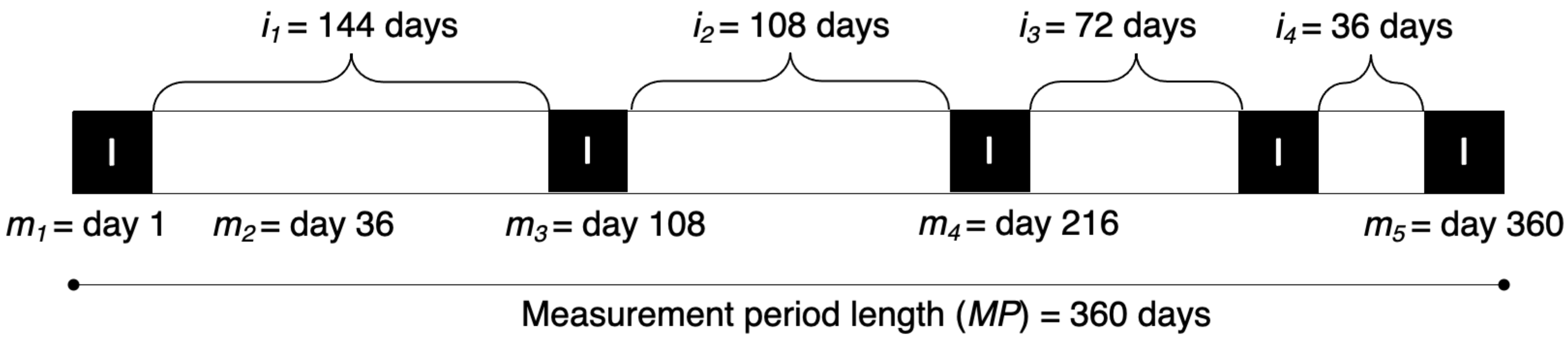
$$\text{Number constant intervals} = \frac{N_{remain}}{\sum_{i=0}^{\#IN} i}, \tag{C.6}$$

giving a length of 36 days for the constant value ($c = \frac{216}{6} = 36$ days). Having computed the value for c, the following interval lengths were obtained:

- $i_1 = x + \#IN(c) = 36 + 3(36) = 144$ days
- $i_4 = x + \#IN(c) = 36 + 0(36) = 36$ days
- $i_3 = x + \#IN(c) = 36 + 1(36) = 72$ days
- $i_2 = x + \#IN(c) = 36 + 2(36) = 108$ days

Figure C.3

Procedure for Generating Equal Spacing Schedules With Time-Interval Decreasing Spacing



Setup variables (Step 1)

$\sum \blacksquare$ = number of measurements (*NM*) = $\underline{5}$ measurements

$\sum \square$ = number of intervals (*NI*) = *NM – 1 =* 4 intervals of *x + #IN(c)* length

#IN = interval Number ∈ {number of intervals (NI) – 1, … ,0}

c = constant by which interval lengths increase

Interval calculation (Step 2)

Number of remaining days (N_{remain}) = MP – NI(x) = 360 – 4(36) = $\underline{216}$ days

Constant length (c) = $\frac{N_{remain}}{\sum_{i=\#IN}^{0} i} = \frac{216}{(3+2+1+0)} = \underline{36}$ days

$i_1 = x + \#IN(c) = 36 + 3(c) =$ **144 days**

$i_2 = x + \#IN(c) = 36 + 2(c) =$ **108 days**

$i_3 = x + \#IN(c) = 36 + 1(c) =$ **72 days**

$i_4 = x + \#IN(c) = 36 + 0(c) =$ **36 days**

C.4 Procedure for Constructing Measurement Schedules With Middle-and-Extreme Spacing

Figure @ref{fig:mid_ext_diagram} shows how the two-step procedure was used to calculate the interval lengths for measurement schedules defined by middle-and-extreme spacing with five measurements. In the first step, the number of intervals was determined by subtracting one from the number of measurements, yielding a value of four for the number of intervals ($NI = NM - 1 = 5 - 1 = 4$)...

Appendix D: Convergence Success Rates

Table D.1
Convergence Success in Experiment 1

		Days to halfway elevation		
Measurement Spacing	Number of Measurements	80	180	280
Equal	5	1.00	0.98	0.95
	7	1.00	1.00	0.99
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00
Time-interval increasing	5	1.00	1.00	1.00
	7	1.00	1.00	1.00
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00
Time-interval decreasing	5	1.00	0.96	0.82
	7	1.00	0.99	0.98
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00
Middle-and-extreme	5	1.00	0.96	0.86
	7	1.00	1.00	1.00
	9	1.00	1.00	1.00
	11	1.00	1.00	1.00

Note. Cells shaded in gray indicate conditions where less than 90% of models converged.

Table D.2
Convergence Success in Experiment 2

		Sample size (*N*)					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000
Equal	5	1.00	1.00	0.99	0.98	0.95	0.92
	7	1.00	1.00	1.00	1.00	0.99	0.98
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time-interval increasing	5	1.00	1.00	1.00	1.00	1.00	1.00
	7	1.00	1.00	1.00	1.00	1.00	1.00
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time-interval decreasing	5	1.00	0.99	0.98	0.95	0.93	0.88
	7	1.00	1.00	0.99	0.99	0.98	0.95
	9	1.00	1.00	1.00	1.00	1.00	0.99
	11	1.00	1.00	1.00	1.00	1.00	1.00
Middle-and-extreme	5	1.00	0.99	0.98	0.96	0.90	0.81
	7	1.00	1.00	1.00	1.00	1.00	1.00
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00

Note. Cells shaded in gray indicate conditions where less than 90% of models converged.

Table D.3
Convergence Success in Experiment 3

		Sample size (*N*)					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000

Time structured	5	1.00	0.99	0.99	0.98	0.96	0.90
	7	1.00	1.00	1.00	1.00	0.99	0.98
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time unstructured (fast response)	5	1.00	1.00	0.98	0.99	0.96	0.90
	7	1.00	1.00	1.00	0.99	0.98	0.99
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00
Time unstructured (slow response)	5	1.00	1.00	0.99	1.00	0.95	0.92
	7	1.00	1.00	1.00	0.99	0.99	0.98
	9	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00	1.00

Note. Cells shaded in gray indicate conditions where less than 90% of models converged.

Appendix E: Complete Versions of Parameter Estimation Plots (Day- and Likert-Unit Parameters)

E.1 Experiment 1

E.1.1 Equal Spacing

Figure E.1

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 1

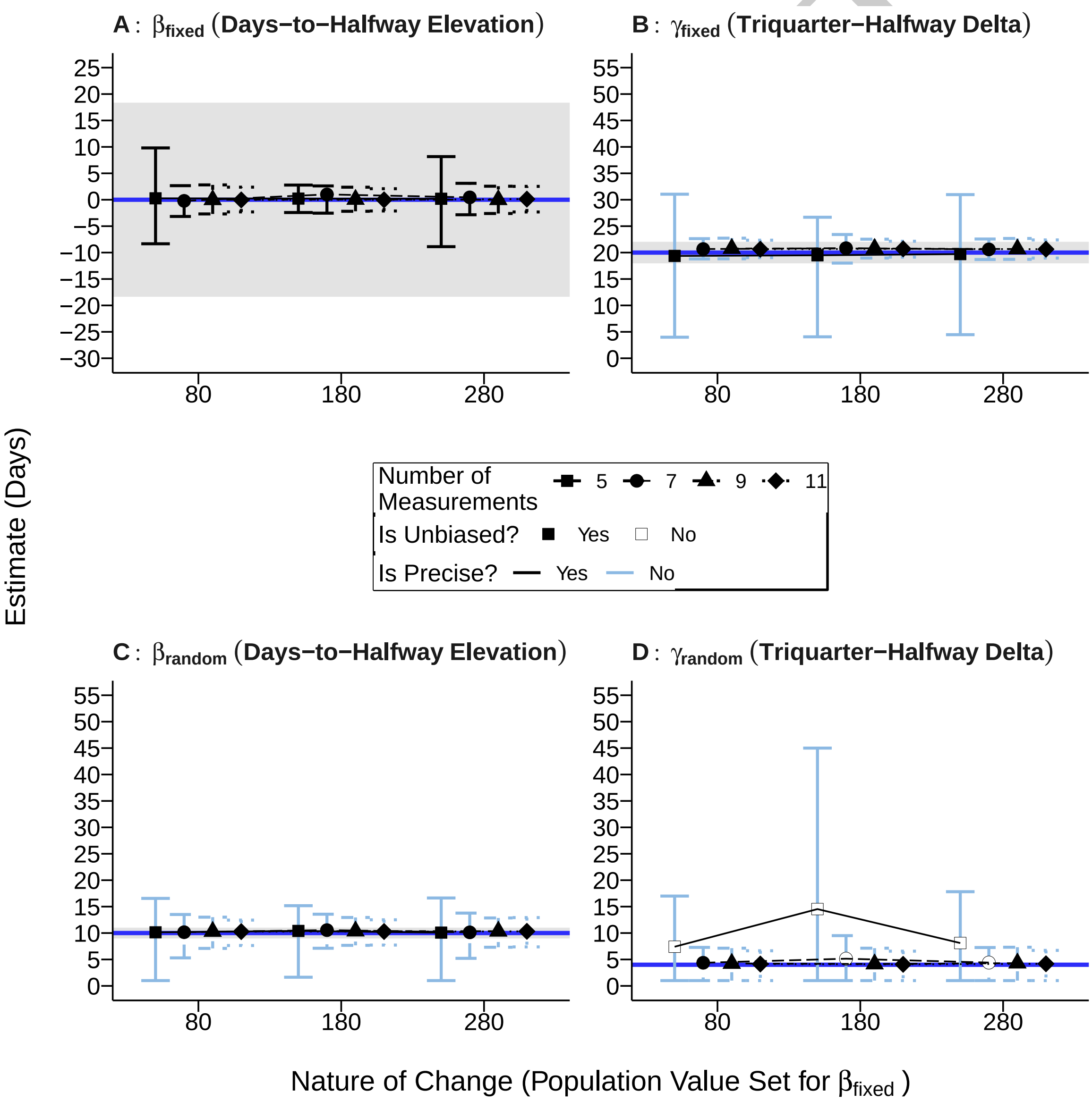
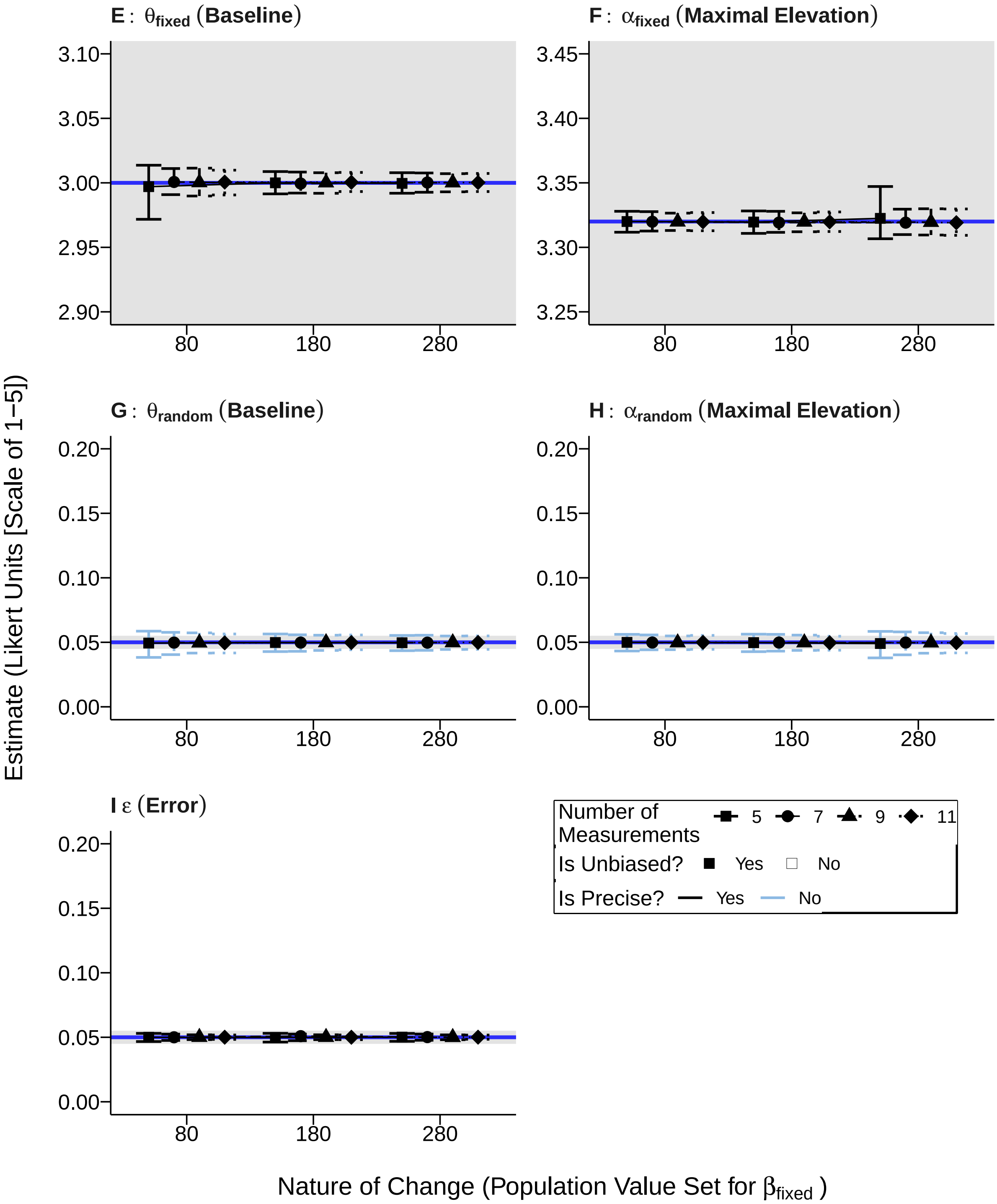


Figure E.1

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 1 (continued)



Note. Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table F.1 for specific values estimated for each parameter.

E.1.2 Time-Interval Increasing Spacing

Figure E.2

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 1

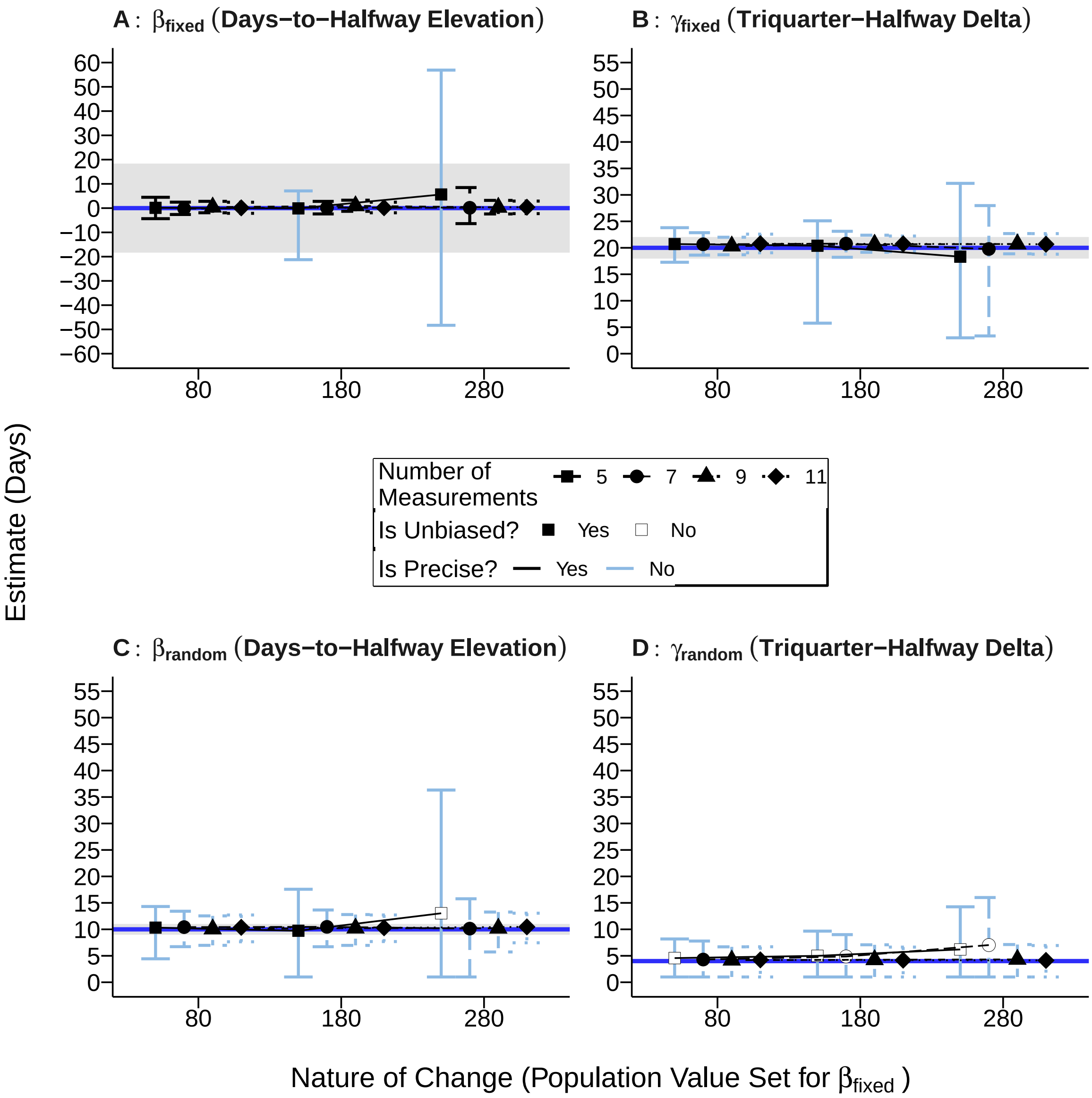
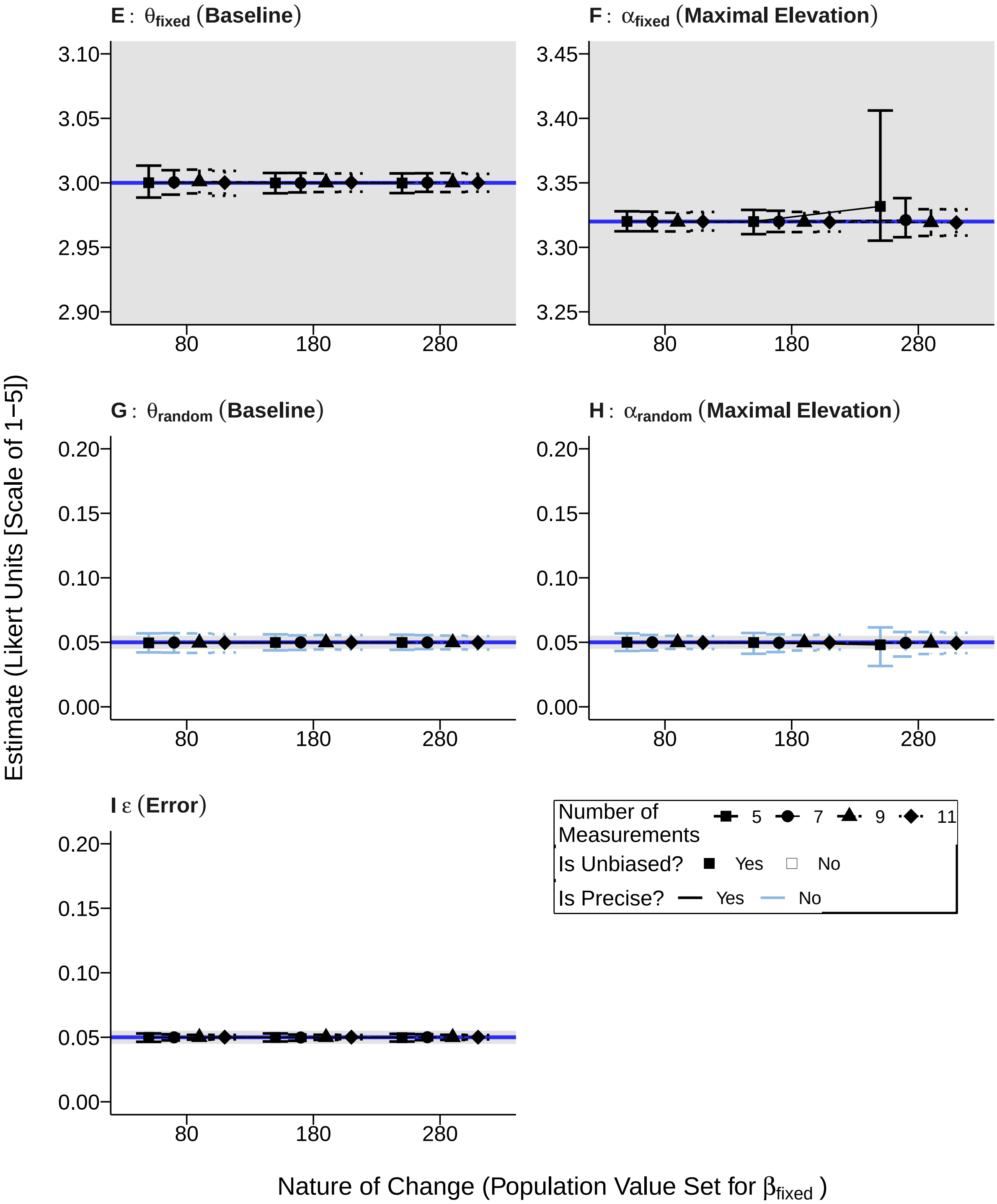


Figure E.2

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 1 (continued)



3257 *Note.* Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation
3258 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and
3259 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
3260 Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table F.1 for specific values estimated for each parameter.

3273 E.1.3 Time-Interval Decreasing Spacing

Figure E.3

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 1

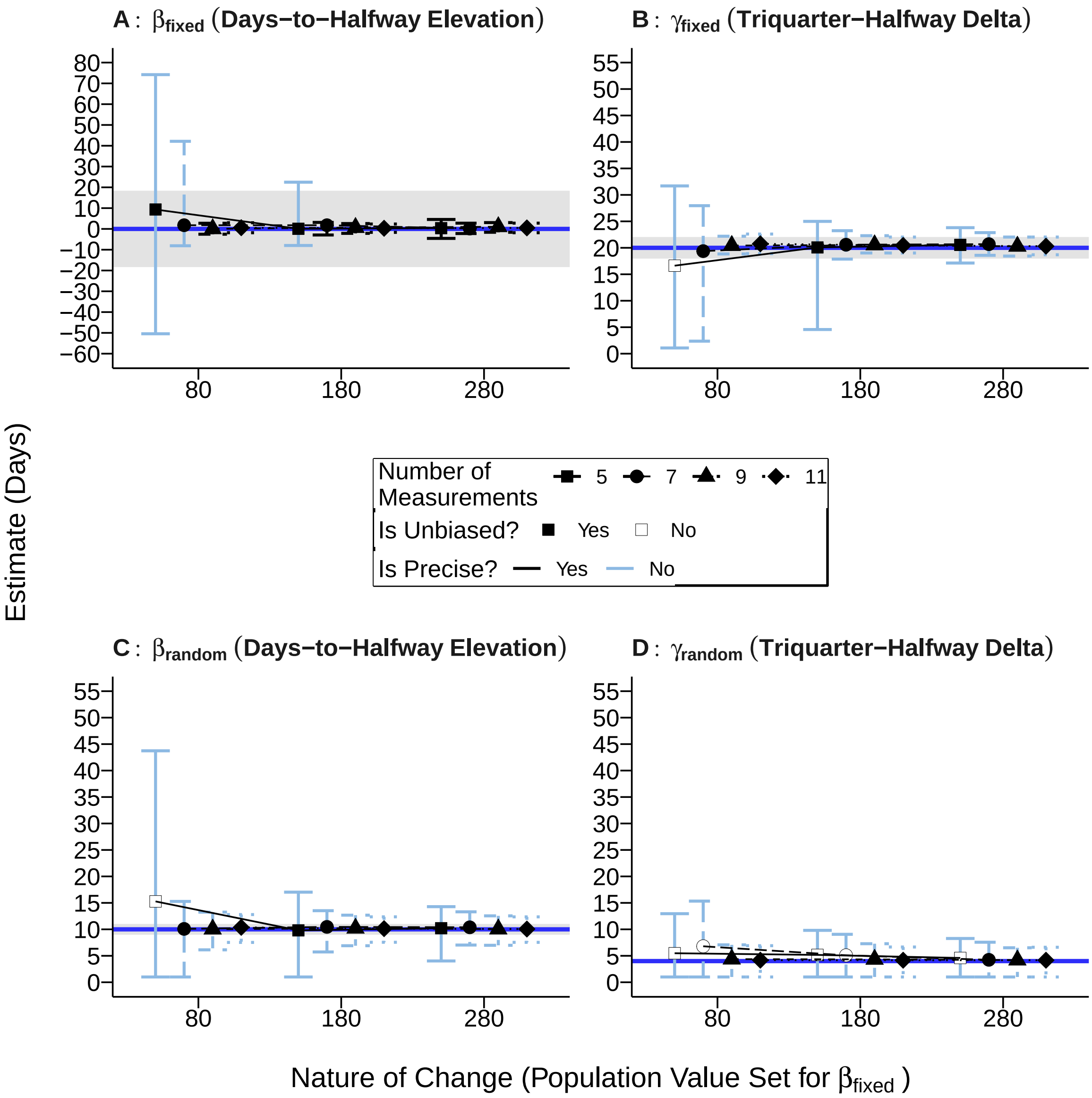
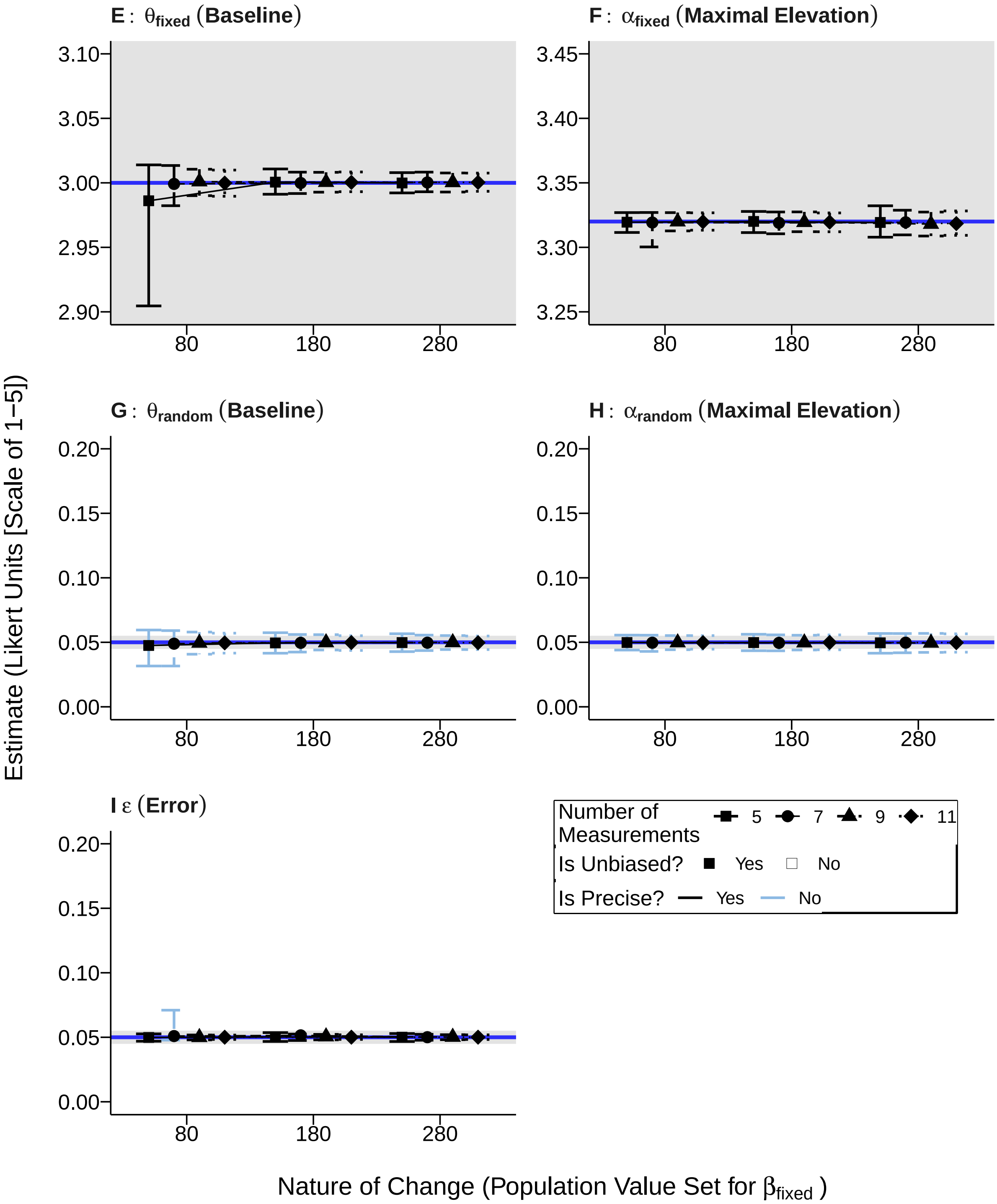


Figure E.3

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 1 (continued)



3274 *Note.* Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation
3275 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and
3276 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
3277 Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table F.1 for specific values estimated for each parameter.

E.1.4 Middle-and-Extreme Spacing

Figure E.4

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 1

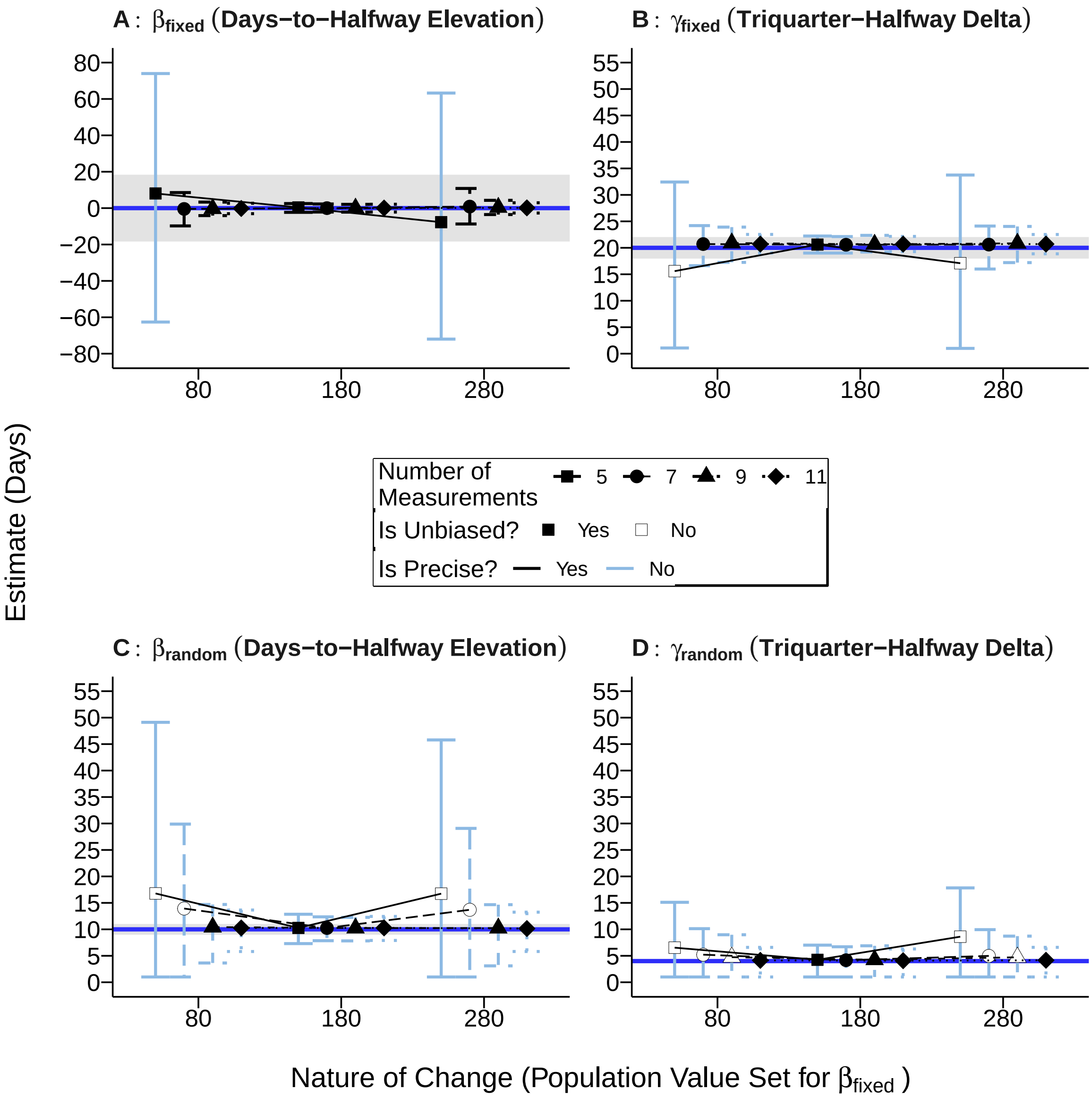
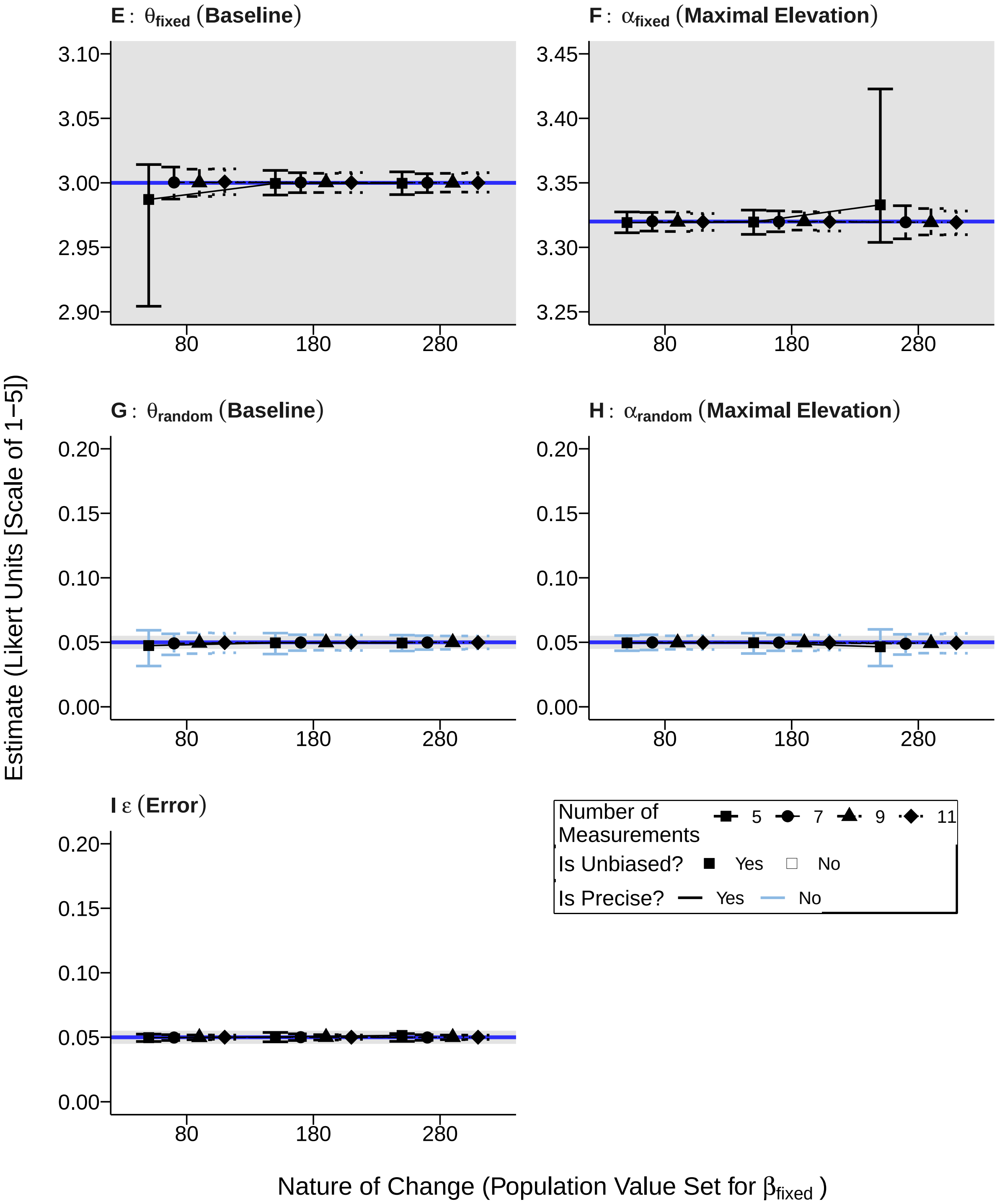


Figure E.4

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 1 (continued)



3291 *Note.* Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation
3292 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and
3293 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
3294 Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. Importantly, across all nature-of-change values (i.e., population values used for β_{fixed}), the acceptable amount of bias and precision was based on a population value of 180. See Table F.1 for specific values estimated for each parameter.

E.2 Experiment 1

E.2.5 Equal Spacing

Figure E.5

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 2

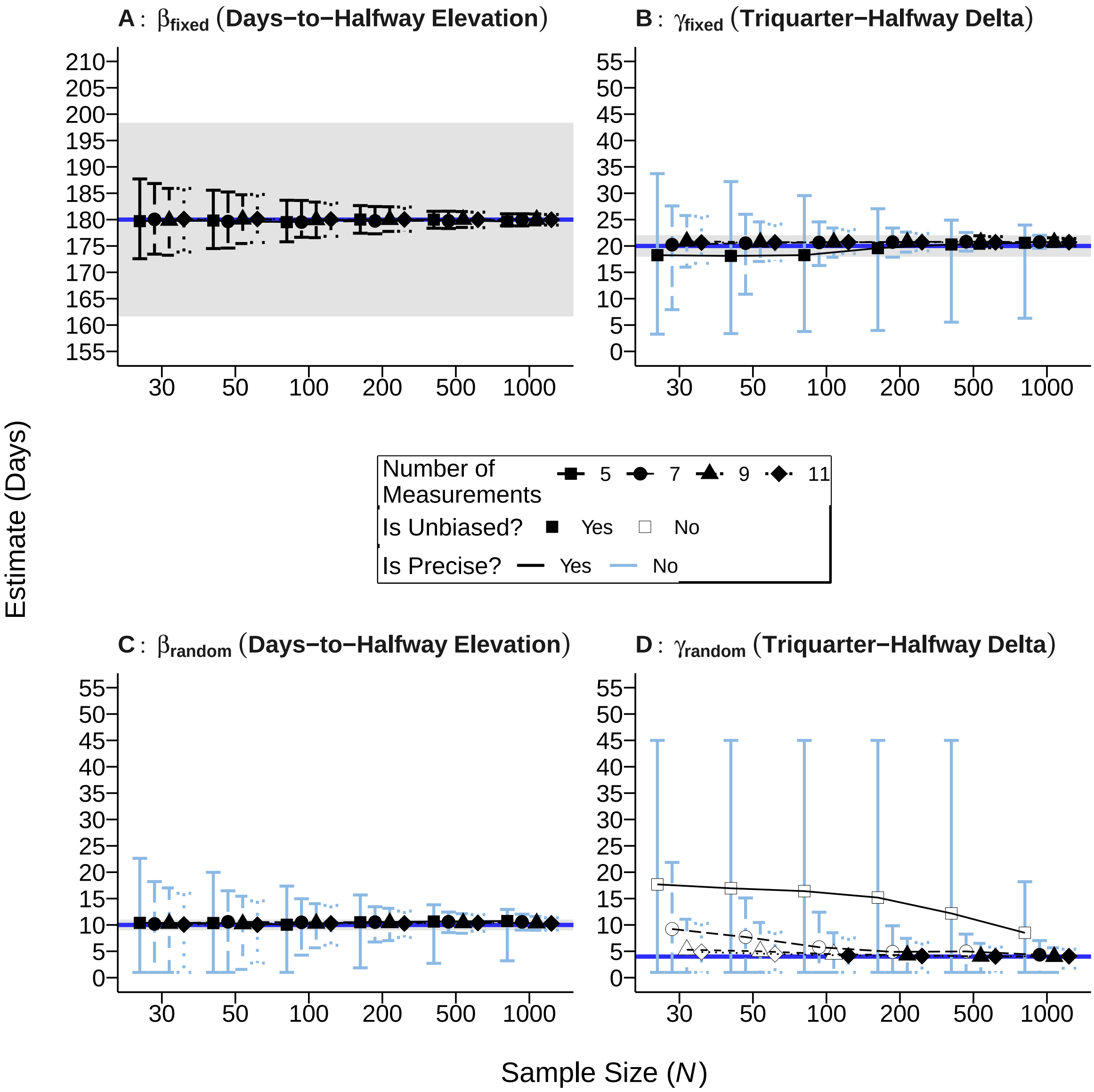
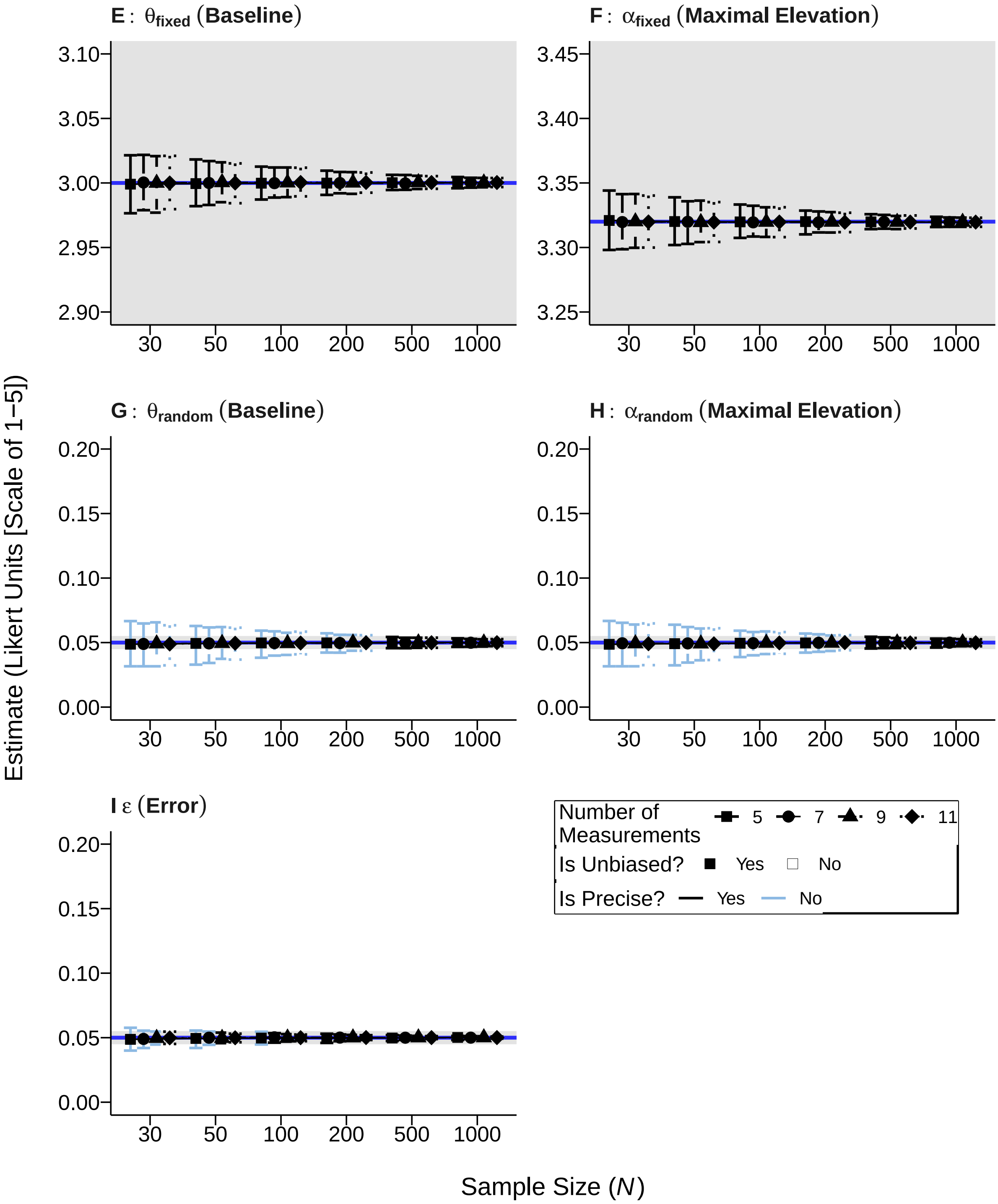


Figure E.5

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Equal Spacing in Experiment 2 (continued)



3309 *Note.* Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation
3310 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and
3311 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
3312 Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table F.2 for specific values estimated for each parameter.

E.2.6 Time-Interval Increasing Spacing

Figure E.6

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 2

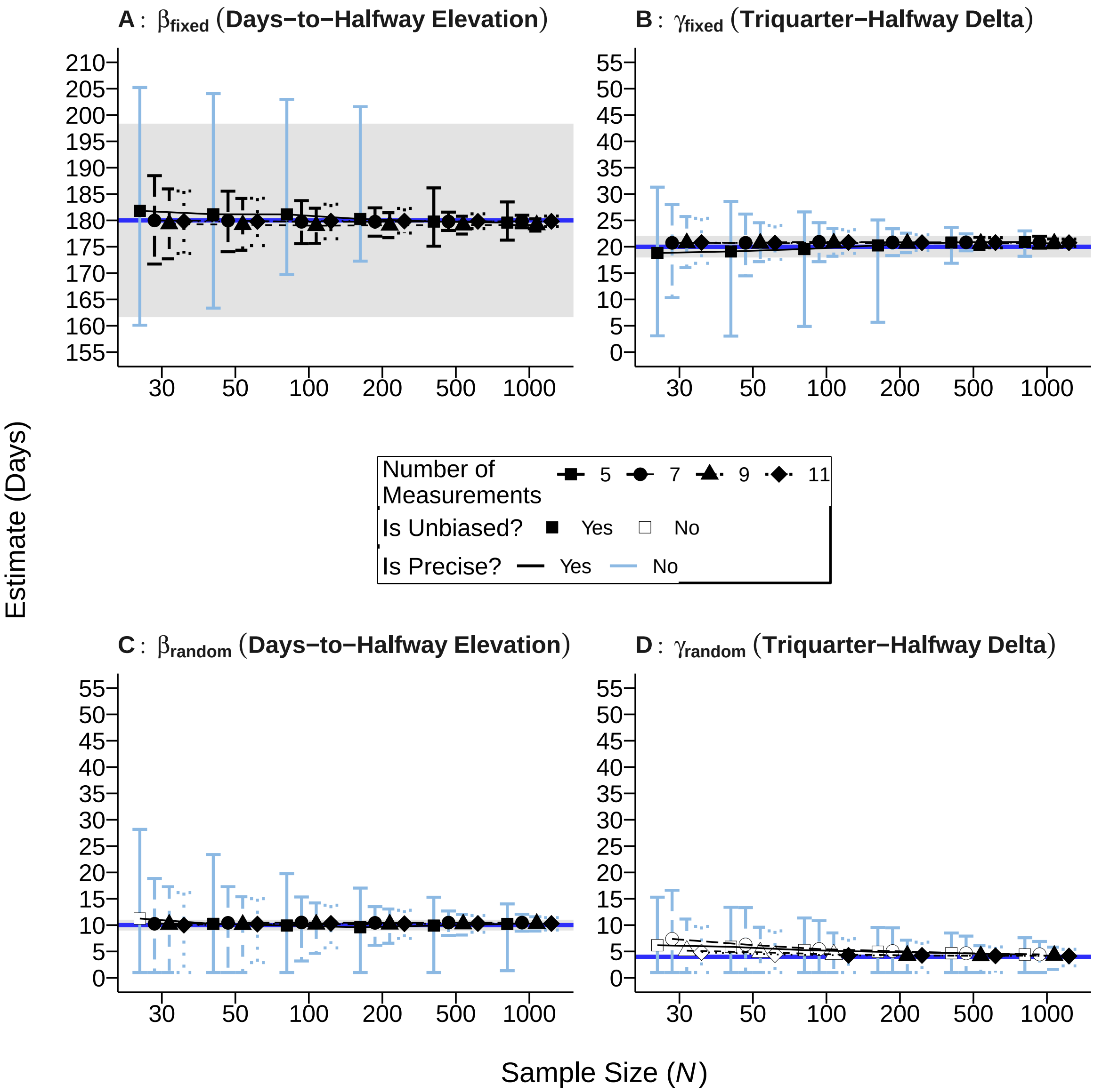
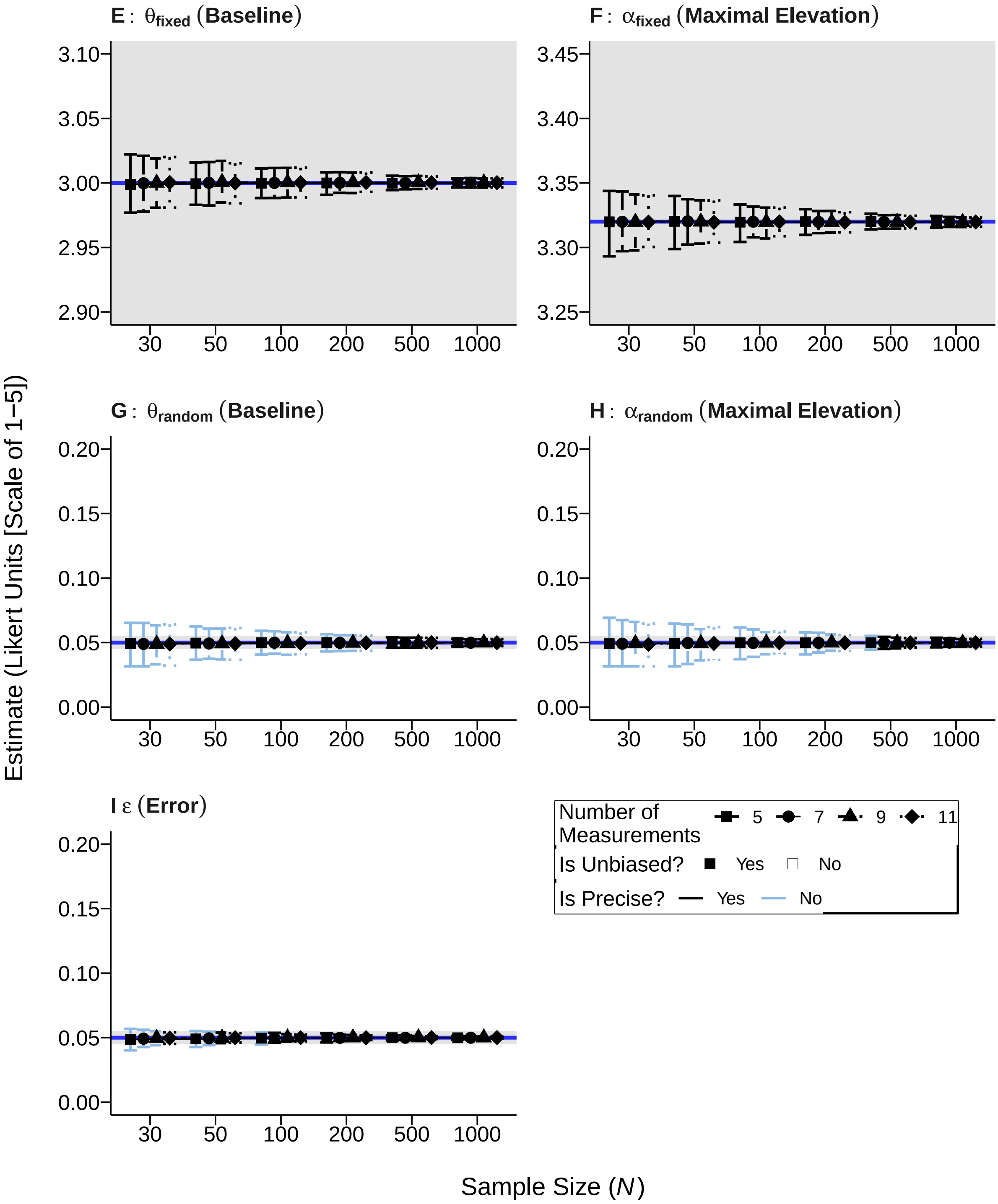


Figure E.6

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Increasing Spacing in Experiment 2 (continued)



Note. Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table F.2 for specific values estimated for each parameter.

E.2.7 Time-Interval Decreasing Spacing

Figure E.7

[]Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 2

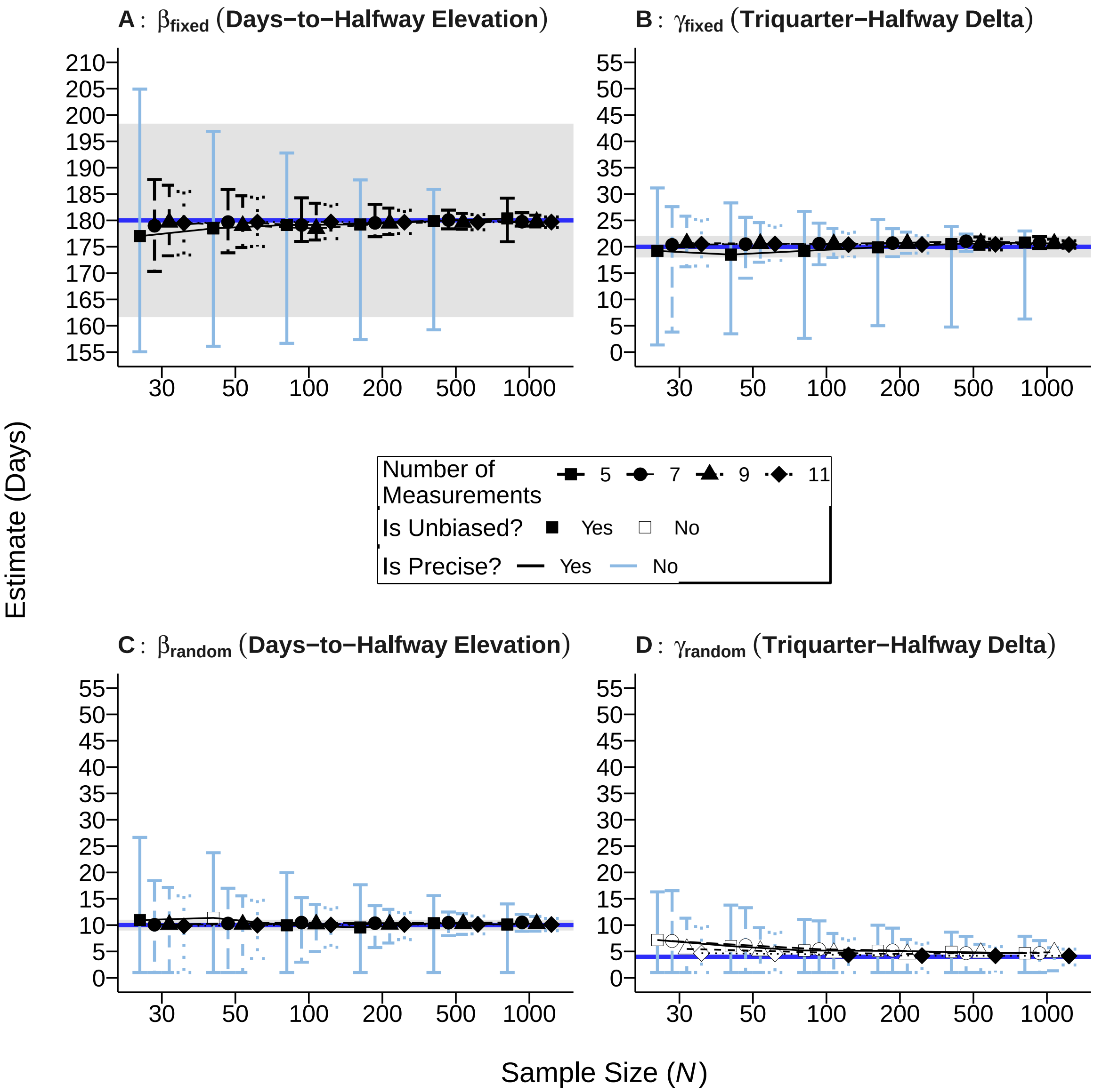
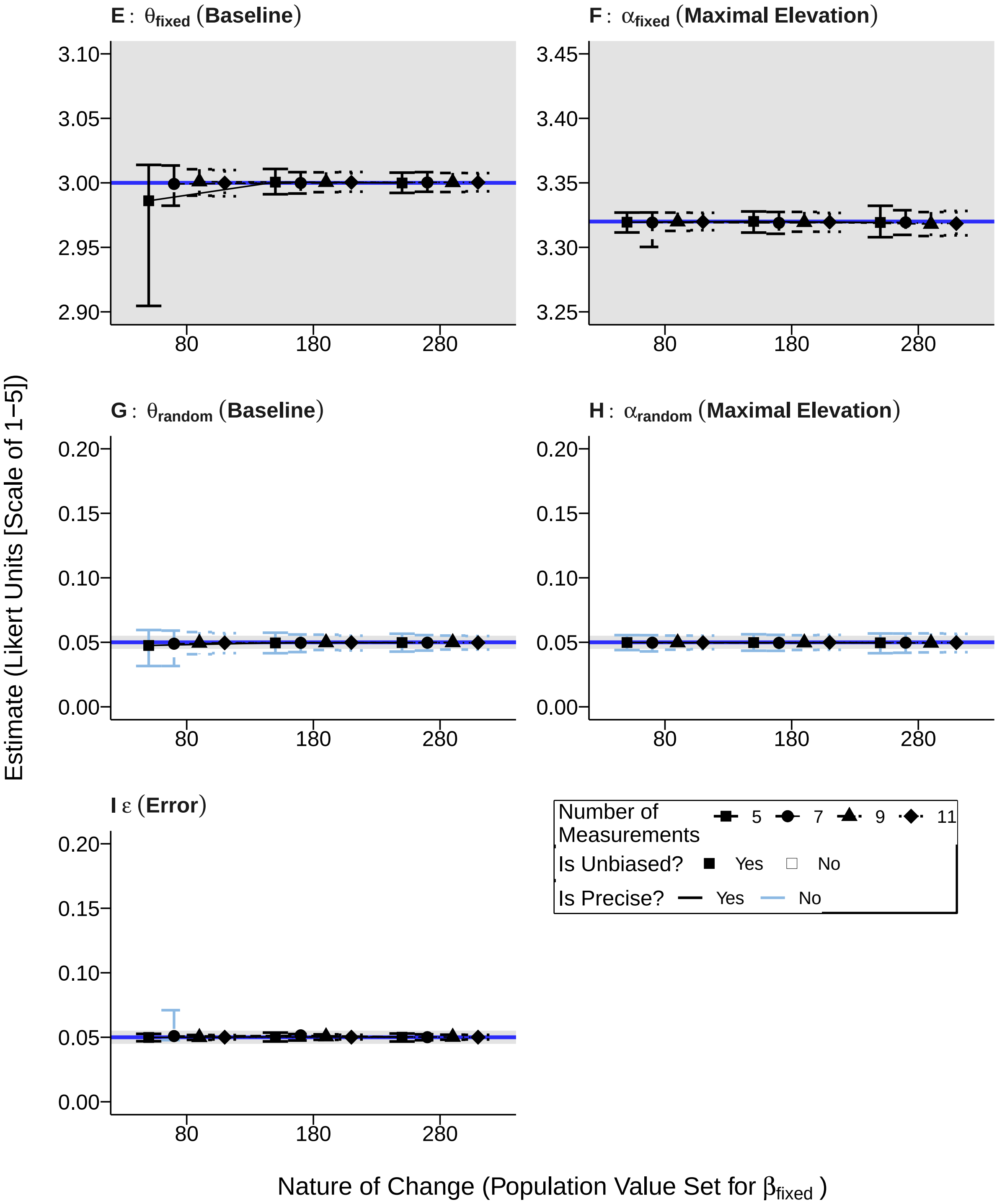


Figure E.7

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Time-Interval Decreasing Spacing in Experiment 2 (continued)



3341 *Note.* Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation
3342 parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and
3343 random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F:
3344 Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table F.2 for specific values estimated for each parameter.

3356 E.2.8 Middle-and-Extreme Spacing

Figure E.8

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 2

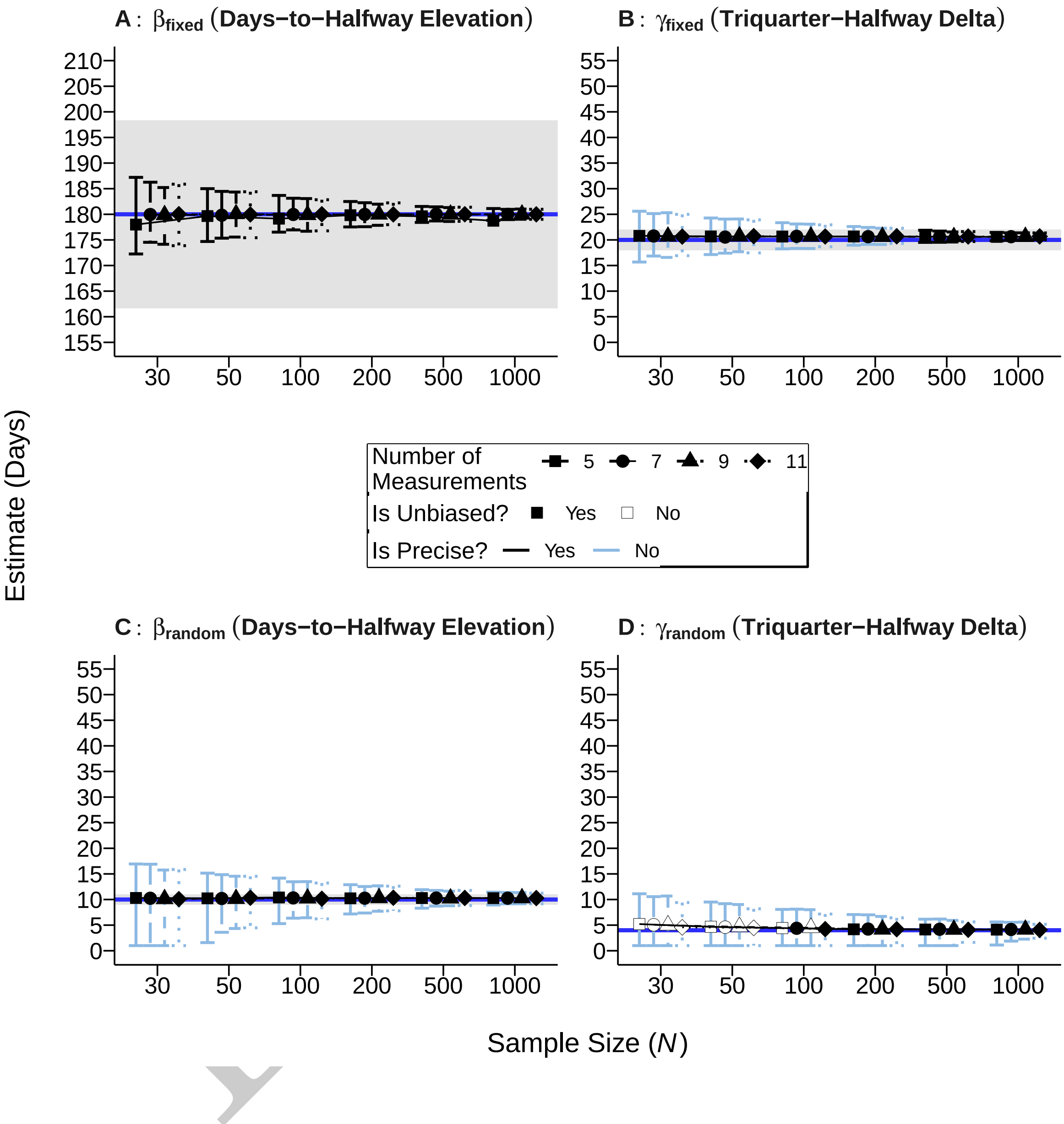
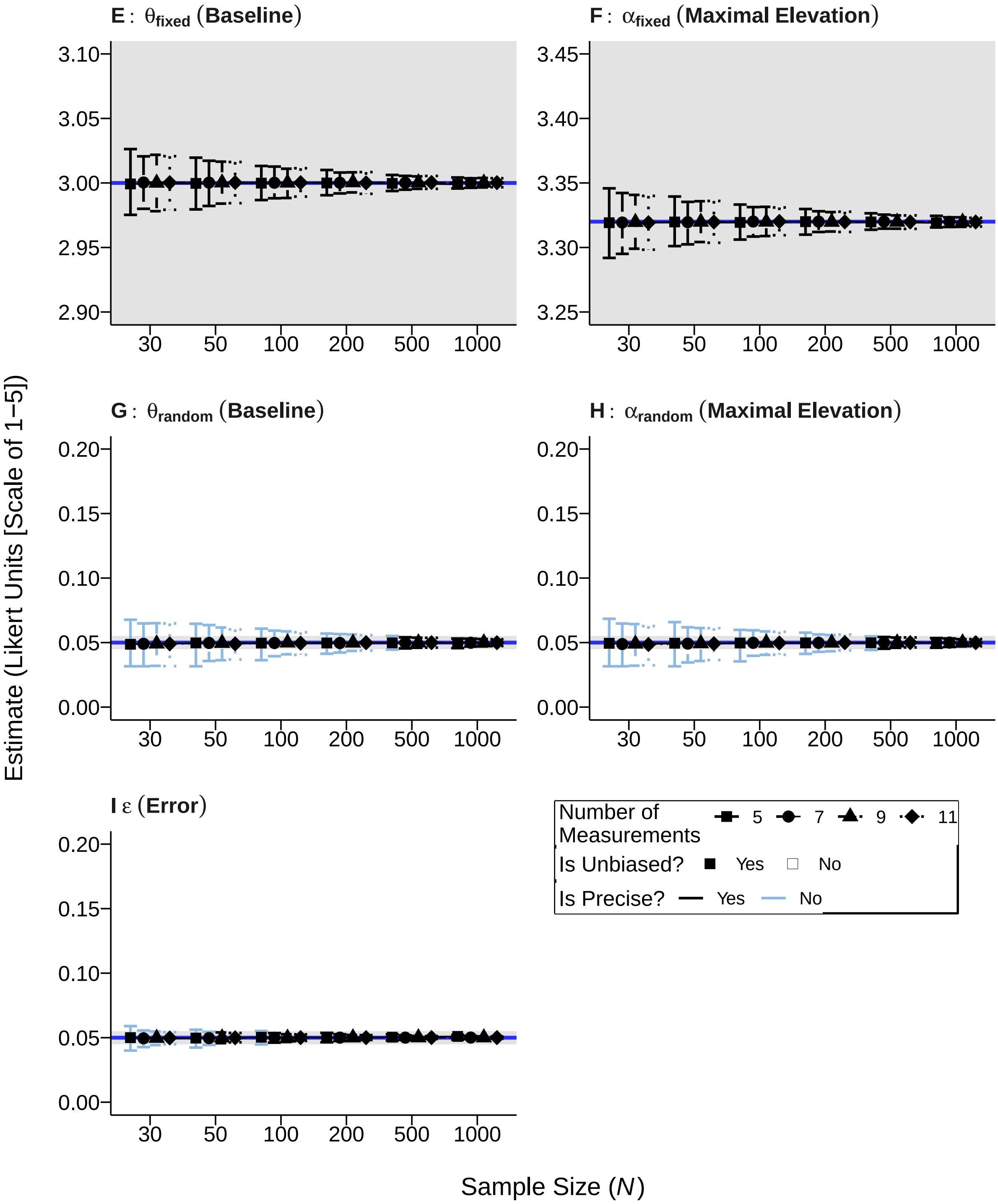


Figure E.8

Parameter Estimation Plots for Day- and Likert-Unit Parameters With Middle-and-Extreme Spacing in Experiment 2 (continued)



Note. Panels A–B: Parameter estimation plots for the fixed- and random-effect days-to-halfway elevation parameters, respectively (β_{fixed} and β_{random}). Panels C–D: Parameter estimation plots for the fixed- and random-effect triquarter-halfway elevation parameters, respectively (γ_{fixed} and γ_{random}). Panels E–F: Parameter estimation plots for the fixed- and random-effect baseline parameters, respectively (θ_{fixed} and

θ_{random}). Panels G–H: Parameter estimation plots for the fixed- and random-effect maximal elevation parameters, respectively (α_{fixed} and α_{random}). Blue horizontal lines in each panel represent the population value for each parameter. Population values for each day-unit parameter are as follows: $\beta_{fixed} \in$ 80, 180, 280, β_{random} = 10.00, γ_{fixed} = 20.00, γ_{random} = 4.00, θ_{fixed} = 3.00, θ_{random} = 0.05, α_{fixed} = 3.32, α_{random} = 0.05, ϵ = 0.05. Gray bands indicate the $\pm 10\%$ margin of error for each parameter and unfilled dots indicate cells with average parameter estimates outside of the margin or biased estimates. Error bars represent the middle 95% of estimated values, with light blue error bars indicating imprecise estimation. I considered dots that fell outside the gray bands as biased and error bar lengths with at least one whisker length exceeding the 10% cutoff (i.e., or longer than the portion of the gray band underlying the whisker) as imprecise. Note that random-effect parameter units are in standard deviation units. See Table F.2 for specific values estimated for each parameter.

Appendix F: Parameter Estimate Tables

F.1 Experiment 1

Table F.1

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 1

Measurement Spacing	Number of Measurements	β_{fixed} (Days to halfway elevation)			β_{random} (Days to halfway elevation) Pop value = 10.00			γ_{fixed} (Triquarter-halfway delta) Pop value = 20.00			γ_{random} (Triquarter-halfway delta) Pop value = 4.00		
		80	180	280	80	180	280	80	180	280	80	180	280
Equal spacing	5	79.73	179.78	279.81□	10.14	10.40	10.08	19.37	19.49	19.71	7.41□	14.53□	8.11□
	7	80.21	178.99	279.55□	10.16	10.55	10.13	20.67	20.83	20.60	4.37	5.14□	4.41□
	9	80.00	179.94	279.99□	10.29	10.37	10.34	20.77	20.76	20.67	4.24	4.14	4.30
	11	80.03	180.01	279.88□	10.27	10.29	10.32	20.64	20.70	20.64	4.13	4.08	4.18
Time-interval increasing	5	79.88	180.10	274.37□	10.32	9.73	13.04□	20.71	20.39	18.32	4.57□	4.99□	6.20□
	7	80.19	179.82	279.86□	10.42	10.47	10.14	20.66	20.79	19.78	4.29	4.87□	7.03□
	9	79.59	179.06	279.70□	10.07	10.22	10.20	20.33	20.66	20.72	4.17	4.25	4.32
	11	79.89	179.84	279.62□	10.38	10.30	10.47	20.78	20.75	20.68	4.23	4.18	4.13
Time-interval decreasing	5	70.67	179.92	279.63□	15.28□	9.80	10.22	16.63	20.07	20.55	5.48□	5.17□	4.59□
	7	78.23	178.22	279.84□	10.08	10.46	10.39	19.38	20.59	20.69	6.80□	5.09□	4.24
	9	79.95	179.34	278.98□	10.03	10.20	10.05	20.42	20.54	20.28	4.37	4.32	4.19
	11	79.42	179.70	279.52□	10.38	10.13	10.06	20.75	20.45	20.31	4.17	4.16	4.17
Middle-and-extreme spacing	5	71.95	179.61	287.73□	16.78□	10.26	16.74□	15.59	20.61	17.09	6.54□	4.24	8.61□
	7	80.45	180.00	279.15□	13.93□	10.25	13.69□	20.71	20.58	20.61	5.21□	4.16	4.98□
	9	80.28	180.05	279.63□	10.42	10.24	10.24	20.91	20.65	20.85	4.74□	4.26	4.72□
	11	80.19	179.96	279.86□	10.27	10.28	10.15	20.71	20.70	20.71	4.14	4.08	4.16

Table F.1

Parameter Values Estimated for Day- and Likert-Unit Parameters in Experiment 1 (continued)

		θ_{fixed} (Baseline) Pop value = 3.00			θ_{random} (Baseline) Pop value = 0.05			α_{fixed} (Maximal elevation) Pop value = 3.32			α_{random} (Maximal elevation) Pop value = 0.05			ϵ(error) Pop value = 0.03		
Measurement Spacing	Number of Measurements	80	180	280	80	180	280	80	180	280	80	180	280	80	180	280
Equal spacing	5	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.33	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	2.99	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	2.99	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.33	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	0.05	0.05	0.05	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05

Note. Cells shaded in light blue indicate cells where estimation is imprecise (i.e., lower and/or upper whisker lengths exceeding 10% of the parameter's population value. Empty superscript squares ($^{\square}$) indicate biased estimates (i.e., bias exceeding 10% of parameter's population value). Importantly, bias and precision cutoff values for the days-to-halfway elevation parameter (β_{fixed}) are based on a value of 180.00.

Table F.2

Parameter Values Estimated in Experiment 2

Measurement Spacing	Number of Measurements	β_{fixed} (Days to halfway elevation) Pop value = 180.00						β_{random} (Days to halfway elevation) Pop value = 10.00					
		30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	179.71	179.82	179.53	180.00	179.99	179.64	10.40	10.36	10.04	10.51	10.65	10.74
	7	180.05	179.65	179.53	179.75	179.76	179.99	10.18	10.59	10.49	10.54	10.60	10.58
	9	179.84	180.07	179.94	180.00	180.02	180.03	10.28	10.20	10.30	10.40	10.39	10.36
	11	180.11	180.11	180.01	180.03	179.98	179.98	10.08	10.04	10.28	10.29	10.38	10.29
Time-interval increasing	5	181.81	181.16	181.14	180.27	179.78	179.57	11.24□	10.24	9.93	9.59	9.91	10.22
	7	179.99	179.96	179.73	179.77	179.79	179.83	10.26	10.43	10.50	10.43	10.47	10.47
	9	179.33	179.18	178.99	179.07	179.11	179.13	10.15	10.10	10.17	10.18	10.21	10.29
	11	179.81	179.79	179.86	179.88	179.81	179.82	9.99	10.19	10.32	10.27	10.30	10.30
Time-interval decreasing	5	177.01	178.48	179.13	179.23	179.86	180.37	10.95	11.38□	9.97	9.55	10.36	10.11
	7	178.98	179.68	179.12	179.53	180.07	179.75	10.07	10.31	10.48	10.37	10.46	10.51
	9	179.65	179.01	178.46	179.47	179.64	179.75	10.11	10.16	10.20	10.17	10.28	10.26
	11	179.48	179.68	179.70	179.65	179.64	179.68	9.85	9.98	10.03	10.12	10.13	10.11
Middle-and-extreme spacing	5	177.99	179.65	179.15	179.83	179.61	178.74	10.30	10.24	10.40	10.24	10.28	10.26
	7	179.96	179.82	179.97	179.98	180.02	179.98	10.25	10.20	10.32	10.26	10.29	10.27
	9	179.88	180.07	179.89	179.98	179.98	179.99	10.12	10.16	10.24	10.30	10.24	10.29
	11	180.02	179.96	180.01	179.98	180.01	179.99	10.08	10.35	10.15	10.35	10.30	10.28

Table F.2

Parameter Values Estimated in Experiment 2 (continued)

		γ_{fixed} (Triquarter-halfway delta) Pop value = 20.00						γ_{random} (Triquarter-halfway delta) Pop value = 4.00					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	18.25	18.11	18.27	19.59	20.27	20.60	17.69□	16.95□	16.41□	15.19□	12.19□	8.51□
	7	20.25	20.53	20.66	20.75	20.81	20.74	9.22□	7.70□	5.77□	4.89□	4.98□	4.34
	9	20.88	20.72	20.73	20.76	20.75	20.73	5.30□	4.99□	4.44□	4.27	4.03	4.00
	11	20.65	20.66	20.73	20.70	20.69	20.71	4.86□	4.49□	4.20	4.10	4.02	4.07
Time-interval increasing	5	18.81	19.11	19.56	20.25	20.80	20.92	6.18□	5.88□	5.25□	4.94□	4.68□	4.42□
	7	20.74	20.74	20.94	20.83	20.83	20.82	7.38□	6.31□	5.45□	5.06□	4.66□	4.45□
	9	20.72	20.65	20.69	20.65	20.63	20.65	5.15□	4.83□	4.44□	4.26	4.16	4.23
	11	20.80	20.69	20.84	20.76	20.78	20.76	4.84□	4.43□	4.25	4.26	4.17	4.14
Time-interval decreasing	5	19.21	18.50	19.21	19.90	20.50	20.79	7.17□	6.01□	5.18□	5.12□	4.91□	4.66□
	7	20.36	20.49	20.57	20.69	21.03	20.76	6.98□	6.18□	5.43□	5.20□	4.67□	4.68□
	9	20.69	20.60	20.55	20.62	20.70	20.63	5.48□	5.12□	4.72□	4.52□	4.72□	4.83□
	11	20.49	20.53	20.38	20.41	20.47	20.41	4.66□	4.57□	4.34	4.20	4.18	4.17
Middle-and-extreme spacing	5	20.80	20.69	20.65	20.67	20.64	20.59	5.21□	4.68□	4.43□	4.18	4.15	4.11
	7	20.76	20.55	20.70	20.63	20.60	20.63	5.07□	4.60□	4.39	4.23	4.19	4.15
	9	20.68	20.71	20.67	20.63	20.58	20.63	4.99□	4.67□	4.49□	4.17	4.13	4.15
	11	20.64	20.74	20.67	20.70	20.66	20.68	4.57□	4.47□	4.22	4.19	4.09	4.07

Table F.2

Parameter Values Estimated in Experiment 2 (continued)

		θ_{fixed} (Baseline) Pop value = 3.00						θ_{random} (Baseline) Pop value = 0.05					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05

Table F.2

Parameter Values Estimated in Experiment 2 (continued)

		α_{fixed} (Maximal elevation) Pop value = 3.32						α_{random} (Maximal elevation) Pop value = 0.05					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Equal spacing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05

Table F.2

Parameter Values Estimated in Experiment 2 (continued)

		ϵ(error) Pop value = 0.03					
Measurement Spacing	Number of Measurements	30	50	100	200	500	1000
Equal spacing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval increasing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time-interval decreasing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Middle-and-extreme spacing	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05

Note. Cells shaded in light blue indicate cells where estimation is imprecise (i.e., lower and/or upper whisker lengths exceeding 10% of the parameter's population value. Empty superscript squares ($^{\square}$) indicate biased estimates (i.e., bias exceeding 10% of parameter's population value).

Table F.3

Parameter Values Estimated in Experiment 3

Time Structuredness	Number of Measurements	β_{fixed} (Days to halfway elevation) Pop value = 180.00						β_{random} (Days to halfway elevation) Pop value = 10.00					
		30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	179.71	179.67	179.75	179.98	180.00	179.66	10.40	10.27	10.37	10.56	10.73	10.69
	7	180.05	179.59	179.02	179.66	180.03	179.63	10.18	10.42	10.65	10.52	10.76	10.60
	9	179.84	180.01	180.01	179.97	180.01	180.00	10.28	10.28	10.37	10.46	10.42	10.41
	11	180.11	179.91	179.94	180.00	180.00	180.00	10.08	10.32	10.21	10.29	10.36	10.31
Time unstructured (fast response)	5	177.48	177.24	176.74	177.50	177.42	177.06	10.65	10.36	10.38	10.65	10.85	10.96
	7	176.89	177.03	176.37	175.92	177.20	176.95	10.53	10.60	10.88	10.83	10.84	10.84
	9	177.54	177.28	177.27	177.31	177.34	177.33	10.66	10.43	10.44	10.61	10.65	10.59
	11	177.25	177.35	177.27	177.37	177.35	177.30	10.41	10.37	10.37	10.45	10.52	10.51
Time unstructured (slow response)	5	174.13	174.02	173.65	173.85	173.41	173.63	11.23$^{\square}$	10.93	11.22$^{\square}$	11.80$^{\square}$	12.10$^{\square}$	12.07$^{\square}$
	7	173.31	173.63	173.01	173.06	173.55	173.55	11.71$^{\square}$	11.67$^{\square}$	11.88$^{\square}$	11.97$^{\square}$	11.91$^{\square}$	11.94$^{\square}$
	9	173.37	173.37	173.54	173.52	173.50	173.49	11.26$^{\square}$	11.38$^{\square}$	11.42$^{\square}$	11.40$^{\square}$	11.47$^{\square}$	11.46$^{\square}$
	11	173.58	173.56	173.50	173.51	173.49	173.47	10.87	10.98	11.12$^{\square}$	11.18$^{\square}$	11.14$^{\square}$	11.16$^{\square}$

Table F.3

Parameter Values Estimated in Experiment 3 (continued)

		β_{fixed} (Days to halfway elevation) Pop value = 180.00						β_{random} (Days to halfway elevation) Pop value = 10.00					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	18.25	18.11	18.46	19.67	20.55	20.65	17.69□	17.05□	16.38□	15.03□	11.63□	9.02□
	7	20.25	20.79	20.67	20.77	20.98	20.93	9.22□	7.32□	6.12□	4.99□	4.45□	4.69□
	9	20.88	20.79	20.84	20.69	20.74	20.71	5.30□	4.95□	4.34	4.13	4.05	3.96
	11	20.65	20.74	20.73	20.69	20.71	20.67	4.86□	4.41□	4.17	4.13	4.09	4.03
Time unstructured (fast response)	5	18.57	18.16	18.59	19.45	20.15	20.58	16.85□	16.21□	14.96□	13.48□	9.94□	7.72□
	7	20.39	20.44	20.67	20.73	20.77	20.77	9.65□	7.07□	6.25□	5.47□	4.61□	4.34
	9	20.54	20.66	20.75	20.71	20.72	20.74	5.27□	4.68□	4.59□	4.08	4.06	4.05
	11	20.77	20.70	20.72	20.70	20.71	20.73	4.85□	4.68□	4.29	4.14	4.16	4.14
Time unstructured (slow response)	5	18.66	17.88	18.34	19.83	20.57	20.67	14.54□	13.26□	11.51□	10.05□	7.89□	6.65□
	7	20.51	20.73	20.75	20.89	20.89	20.86	7.62□	6.65□	5.61□	5.21□	4.83□	4.67□
	9	20.91	20.82	20.82	20.89	20.94	20.89	6.00□	5.32□	4.97□	4.67□	4.74□	4.70□
	11	20.98	20.85	20.90	20.92	20.90	20.90	5.26□	4.92□	4.83□	4.69□	4.75□	4.71□

Table F.3

Parameter Values Estimated in Experiment 3 (continued)

Time Structuredness	Number of Measurements	θ_{fixed} (Baseline) Pop value = 3.00						θ_{random} (Baseline) Pop value = 0.05					
		30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (fast response)	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response)	5	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.00	3.00	3.00	3.00	3.00	3.00	0.05	0.05	0.05	0.05	0.05	0.05

Table F.3

Parameter Values Estimated in Experiment 3 (continued)

Time Structuredness	Number of Measurements	α_{fixed} (Maximal elevation) Pop value = 3.32						α_{random} (Maximal elevation) Pop value = 0.05					
		30	50	100	200	500	1000	30	50	100	200	500	1000
Time structured	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (fast response)	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response)	5	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	7	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	9	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05
	11	3.32	3.32	3.32	3.32	3.32	3.32	0.05	0.05	0.05	0.05	0.05	0.05

Table F.3

Parameter Values Estimated in Experiment 3 (continued)

		ϵ(error) Pop value = 0.03					
Time Structuredness	Number of Measurements	30	50	100	200	500	1000
Time structured	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (fast response)	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05
Time unstructured (slow response)	5	0.05	0.05	0.05	0.05	0.05	0.05
	7	0.05	0.05	0.05	0.05	0.05	0.05
	9	0.05	0.05	0.05	0.05	0.05	0.05
	11	0.05	0.05	0.05	0.05	0.05	0.05

Note. Cells shaded in light blue indicate cells where estimation is imprecise (i.e., lower and/or upper whisker lengths exceeding 10% of the parameter's population value. Empty superscript squares ($^{\square}$) indicate biased estimates (i.e., bias exceeding 10% of parameter's population value).