HYBRID LOGISTIC AND CONFINED EXPONENTIAL GROWTH MODELS: ESTIMATION USING SEM SOFTWARE

A Thesis

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

KYLE R RIPLEY

Dr. Phillip Wood, Thesis Supervisor

DECEMBER 2019

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

HYBRID LOGISTIC AND CONFINED EXPONENTIAL GROWTH MODELS: ESTIMATION USING SEM SOFTWARE

presented by Kyle R Ripley,

a candidate for the degree of Master of Arts,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Phillip Wood

Professor Timothy Trull

Professor Don Spiers

DEDICATION

This work is dedicated to all of my family, friends, and colleagues who have helped me through this process – whether it was words of encouragement, ideas to improve my work, or just hanging out and playing games – you all helped me to get this project to this point, and I'm forever in your debt.

In particular, I'd like to dedicate this work to my other half – your love, endless support, unwavering belief in me, and willingness to bake me cookies are why I'm able to keep going.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all of the people who had a hand in the preparation of this document. I can't express how truly grateful I am for your help and support.

First, I'd like to thank my advisor and thesis chair, Dr. Phillip Wood, for his continued guidance, support, encouragement, and input. You've been invaluable and a tremendous help throughout this entire process, and I can't thank you enough.

Second, I'd like to thank my committee members – Dr. Timothy Trull and Dr. Don Spiers – for their feedback, support, and insights into other potential applications of the hybrid model presented.

Finally, I'd like to thank the Department of Psychological Sciences at the University of Missouri for allowing me the opportunity to conduct this work.

TABLE OF CONTENTS

List of Tables

List of Figures

Figure | Page

Abstract

The logistic and confined exponential curves represent growth over time in various contexts such as learning and technology transfer. Logistic growth operates as a contagion process in a population of interest, while the confined exponential curve represents the diffusion of an external process on a system, such as the transfer of information through communication channels. Prior work (e.g., Grimm & Ram, 2009) has shown that such nonlinear curves can be estimated using structural equation modeling (SEM) software, allowing model comparison. As an alternative to binary choice between such models, this paper shows how a hybrid model representing a weighted combination of the two models may be specified. In order to assess whether the hybrid model can be successfully estimated using SEM software and conditions under which it can be successfully differentiated from the stand-alone logistic and confined exponential alternatives, Monte Carlo simulations varying the number of measurement occasions (5, 10, and 15), internal consistency (α = .5, .7, and .8), and sample size (N = 1,000, 500, and 300) were conducted. Convergence failures appeared appreciable only when the estimated hybrid models were the special cases of logistic or confined exponential curves. The hybrid model was successfully preferred over the stand-alone models when 10 or 15 measurement occasions are employed and when internal consistency is moderate (α = .7 or .8) across all sample sizes but not when only five measurement occasions are used or when internal consistency is low (α = .5). Implications for the application of the hybrid model to learning, growth, and psychopathology are discussed.

Hybrid Logistic and Confined Exponential Growth Models:

Estimation using SEM Software

The study of growth over time, or growth curve analysis, has long been an object of scientific inquiry. Early records from the Babylonians suggest a scientific interest in growth curves as early as 2,000 BCE, mostly for calculating interest and payments (Webb, 2000). More recently, growth curve analysis has become popular in a variety of fields, such as models of psychology (e.g., Meredith & Tisak, 1990), technology transfer (e.g., Sharif & Ramanathan, 1981), bioenergetics (e.g., Brody, 1945; Wishart, 1938), and learning (e.g., Thurstone & Ackerson, 1929).

Early mathematical models for growth curve modeling of longitudinal data began in the mid-twentieth century. Bollen (2007) provides an extensive discussion of many of the early models described here. For example, Wishart (1938) measured growth weights of pigs during the first 16 weeks of life and fit the data using orthogonal polynomials. Box (1950) used differences between successive measurement values, a rough measure of instantaneous change, to analyze growth and wear curves. Tucker (1958) and Rao (1958) presented a growth factor model based on analysis of sums of squares and cross-products (SSCP) often referred to as "Tuckerized growth curves." Potthoff and Roy (1964) compared growth curves using multivariate analysis of variance (MANOVA). More recently, growth curve models using an exploratory structural equation modeling approach (ESEM, Grimm, Steele, Ram, & Nesselroade, 2013) and multilevel growth models (Grimm, Ram, & Estabrook, 2016) have been proposed. Within the fields of biometry and bioenergetics, surveys of numerous parametric growth curve models

(Panik, 2014) and growth and diffusion models (Banks, 2013) have also addressed the dynamics of growth and change over time.

Assessment of the correct dimensionality and patterning of growth over time is a central problem in the evaluation and comparison of mathematical models of growth and change. Wood, Steinley, and Jackson (2015) proposed a three-step procedure for the "right-sizing" of structural models for longitudinal data whereby the dimensionality, parsimony, and mean level effects in the data were considered sequentially. Such sequential consideration of models often involves consideration of psychometric and factor analytic models which have been proposed for the characterization of growth and change as discussed below.

Within parametric growth curve models, research has focused on the consideration of the parameterization of an individual model. As discussed later in this paper, however, such discrete choices may not be necessary if a more general model is considered which contains multiple discrete models as special cases. This general idea has been highlighted in some presentations of growth models. Panik (2014), for example, notes that Schnute and Gompertz curves contain other sigmoid growth curves as special cases when selected parameters of the model assume particular canonical values. This paper presents a model in which two commonly used models, the logistic and confined exponential models, are considered as special cases of a more general model in which components of both growth models are combined. Before proceeding, however, a brief introduction of the logistic and confined exponential curves and their applications is necessary.

Logistic Model

The logistic distribution was initially developed by Verhulst (1838) to express the relationship between population growth as a function of the current population size and available resources in Belgium. It is defined by the following equation (which will be explored in further detail in the subsequent sections):

$$N_i = \frac{N_*}{1+ke^{-at}} \quad , \tag{1}$$

where:

$$k = (N_*/N_0) - 1 \quad . \tag{2}$$

N_i above represents the population size at time i, N_* the carrying capacity of the model (a horizontal asymptotic upper value of N_i), and a the growth coefficient which represents how quickly N_i approaches N_*. When the likelihood function is calculated based on this equation, N_0 then represents the expected value of N_i at $t = 0$. Graphically, the logistic curve is characterized by slow initial growth, rapid growth in the middle, and then slower growth eventually reaching a plateau. The inflection point of the logistic curve is exactly in the middle, where the rate of growth changes from increasing to decreasing, or vice versa as shown in Figure 1. The expected value of *N* at the inflection point is defined by the following equation:

$$N_{inflect} = \frac{N_*}{2} \quad . \tag{3}$$

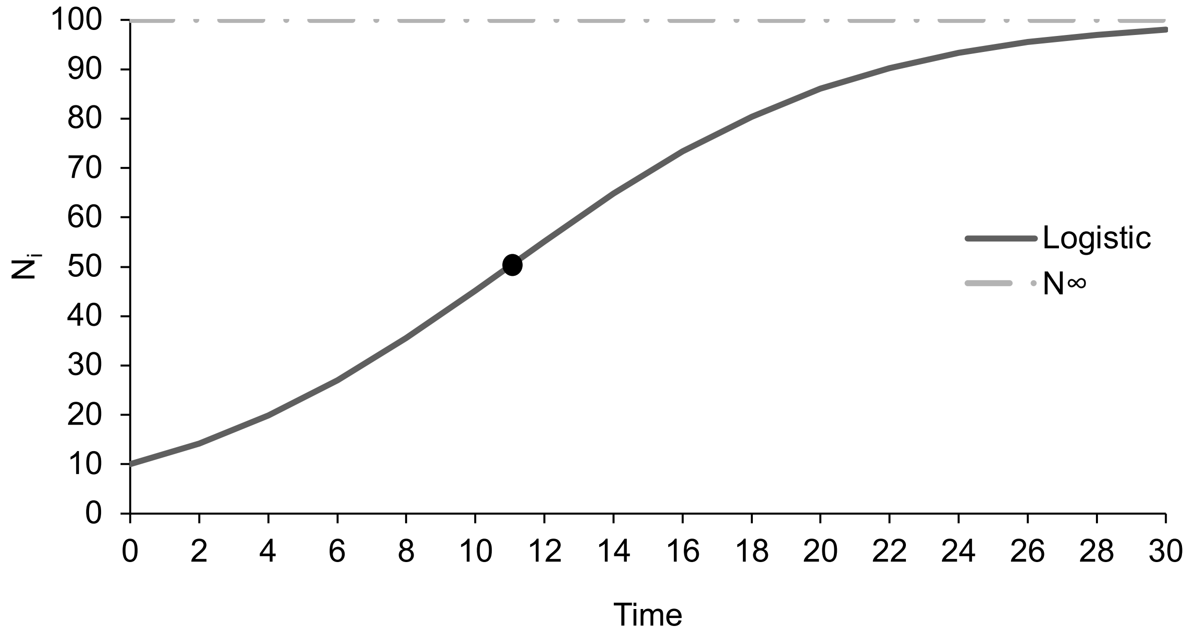

Figure 1. Example of the logistic curve over a thirty-year period. The asymptotic N_* value of 100 displayed as a dashed line. The inflection point of 50 shown as a dot.

Applications of the logistic model. The logistic model has seen wide application in psychology and related social sciences, especially within the context of logistic regression, in which dichotomous or ordered categorical dependent variables are modeled as a criterion. Because disease states are often viewed as dichotomies, the logistic distribution has been widely used in areas of epidemiology such as studies of familial disease (Bonney, 1986).

The logistic model has had a wide application to a variety of other contexts as well. In addition to its uses for modeling the growth of human populations (Verhulst, 1838; Pearl & Reed, 1920; Schultz, 1930; and Oliver, 1982), the logistic distribution has been used to model the growth of yeast (Schultz, 1930), as well as agricultural production (Schultz, 1930; Oliver 1964). In item response theory (IRT), the logistic equation is used in various models, such as the Rasch or 1PL (one parameter logistic), 2PL, 3PL, and the less common 4PL. The logistic equation also forms the basis of Luce's (1959) beta response-strength model in which the improvement in performance induced by an event is proportional to the product of the improvement still possible and the amount already achieved. Additionally, the logistic distribution has been used to model the diffusion of new products in markets with risk-sensitive consumers (Oren & Schwartz, 1988).

Confined Exponential Model

The second distribution that will be considered in this paper is the confined exponential. It is defined by the following equation:

$$N_i = N_* - (N_* - N_0)e^{-a_* t} \quad . \tag{4}$$

The confined exponential equation is similar to the logistic equation in that N_* represents the carrying capacity of the model (a horizontal asymptotic upper value of N_i), and N_0 is the expected value of *N* when $t = 0$. In this equation however, a_* represents the growth coefficient. Unlike the logistic curve, the confined exponential curve is characterized by large initial growth followed by progressively smaller growth as it reaches the asymptote. Because of this, the confined exponential curve has no inflection point (see Figure 2).

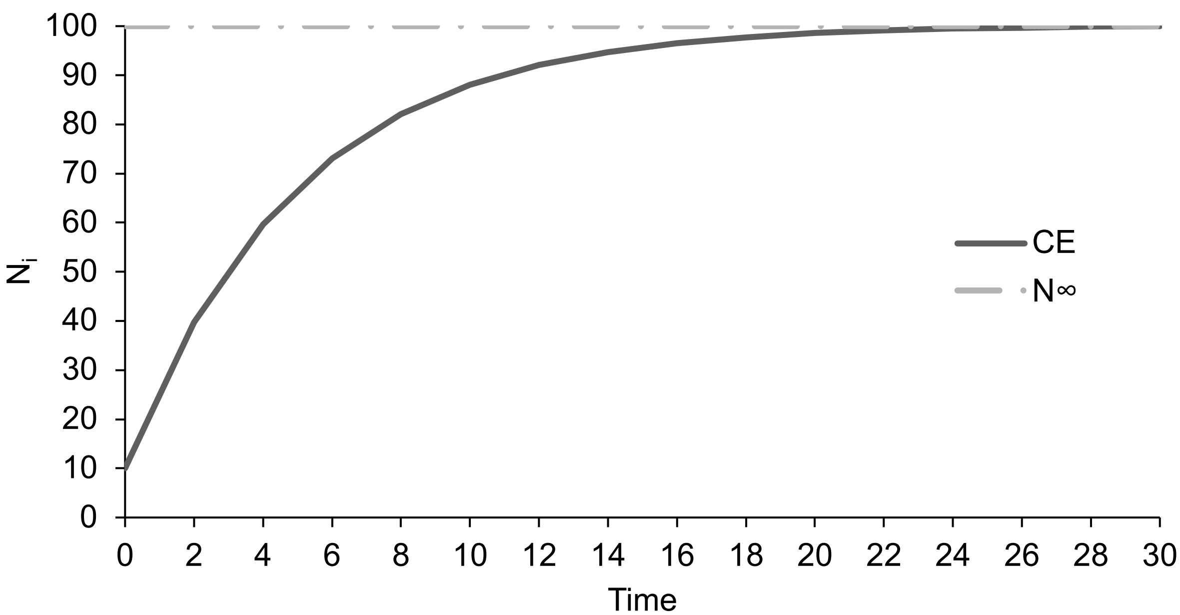

Figure 2. Example of the confined exponential curve over a thirty-year period. The asymptotic N_* value of 100 displayed as dashed line.

Applications of the confined exponential model. As with logistic growth, the confined exponential curve has also seen a wide variety of applications. It has been used to model various phenomena in the social sciences, such as the diffusion of information through different communication channels (e.g., Bartholomew, 1981; Ralston, 1983). The confined exponential distribution has also been useful in mass and heat transfer applications as discussed by Kreith (1958) and Bird, Stewart, and Lightfoot (1961).

In contrast to the logistic model underlying the beta response-strength model proposed by Luce (1959), Bush and Mosteller (1955) proposed a linear-operator model under which improvement in performance induced by an event is proportional to the amount of improvement still possible (i.e., a confined exponential process.). In biology, this distribution has been used to model tree growth (Valentine, 1985) and the effects of fertilizers on crop growth (Batschelet, 1979).

Logistic Model Versus Confined Exponential Model

In addition to there being several applications of these two models, it has often happened that these two particular models have been set in opposition to each other. For example, the earlier mentioned learning models contrast with each other. This relationship also holds in the literature discussing the diffusion of technology and innovation (Bartholomew, 1981; Haynes, Mahajan, & White, 1977; Lekvall & Wahlbin, 1973; Ralston, 1983; Sharif & Ramanathan, 1981) in which the logistic and confined exponential curves are thought to represent fundamentally different underlying transfer mechanisms. The logistic growth curve is thought to represent a dynamic of internal contagion (Lekvall & Wahlbin, 1973) as occurs, for example, when some farmers in a community are early adopters of some technology (such as the use of a new wind turbine) who are then imitated by more risk adverse neighbors. The confined exponential growth curve, on the other hand, is appropriate when change in behavior occurs due to forces external to the population (Lekvall & Wahlbin, 1973) undergoing innovation as would occur in our earlier example if adoption of a new wind turbine were solely a function of the amount of external radio advertising targeted to the farming community (Banks, 2013). The idea of internal contagion versus external influence could apply to psychological phenomenon as well. For example, the progression of psychological disorders could follow a logistic growth pattern if, for example, one psychopathology symptom has a contagious confluence on other symptoms. By contrast, a confined exponential growth pattern would occur if external stressors (an event, substance, etc.) cause a cumulative increase in psychopathology.

To date, comparison of candidate growth models involves the choice between discrete alternatives (e.g., logistic growth vs. confined exponential growth). However, choosing between discrete alternatives may not be ideal if the growth process is actually a combination of the discrete alternatives. Investigation of such a possibility makes a binary choice between models unnecessary and permits the specification of a third alternative in which components of these two processes are weighted. Such an approach has the added benefit of identifying intermediary models that contain aspects of both models as discussed below (see Figure 3).

Hybrid Model

We now present such a hybrid model which contains both the individual logistic and confined exponential curves as special cases. Although informally introduced by Brody (1945), a more formal presentation of the model was reintroduced by Banks (2013). This model will then be used as the basis for a structural model of the hybrid model. The hybrid curve is defined by the following equation:

$$N_i = N_* \frac{1-(ma_*/a)e^{-(a+a_*)t}}{1+me^{-(a+a_*)t}} \tag{5}$$

where:

$$m = \frac{\left(\frac{N_*}{N_0}-1\right)}{\left(\frac{N_* a_*}{N_0\ a}+1\right)} \tag{6}$$

with:

$$w = a/(a + a_*) \quad . \tag{7}$$

The notation for the hybrid model is the same as the previous models such that, N_* represents the carrying capacity of the model, N_0 is the starting value of N_i, and a and a_* are both growth coefficients. In Eq. 7, w is not an estimated parameter but rather a weight between the two growth coefficients which reflects how closely the model is to either the logistic or confined exponential models. This weight allows for the logistic and confined exponential curves to exist as special cases if a or a_* are equal to zero.

This model is particularly interesting to individuals interested in growth models because it includes the logistic and confined exponential growth curves as special cases but also allows assessment of intermediary models in which, for example, the curve can be 50% logistic and 50% confined exponential (see Figure 3, w = 0.5). The added benefit of this model, however, is that it also permits the researcher to entertain the possibility that the growth process under investigation is actually an amalgam of the mathematical forms of the logistic and confined exponential curves.

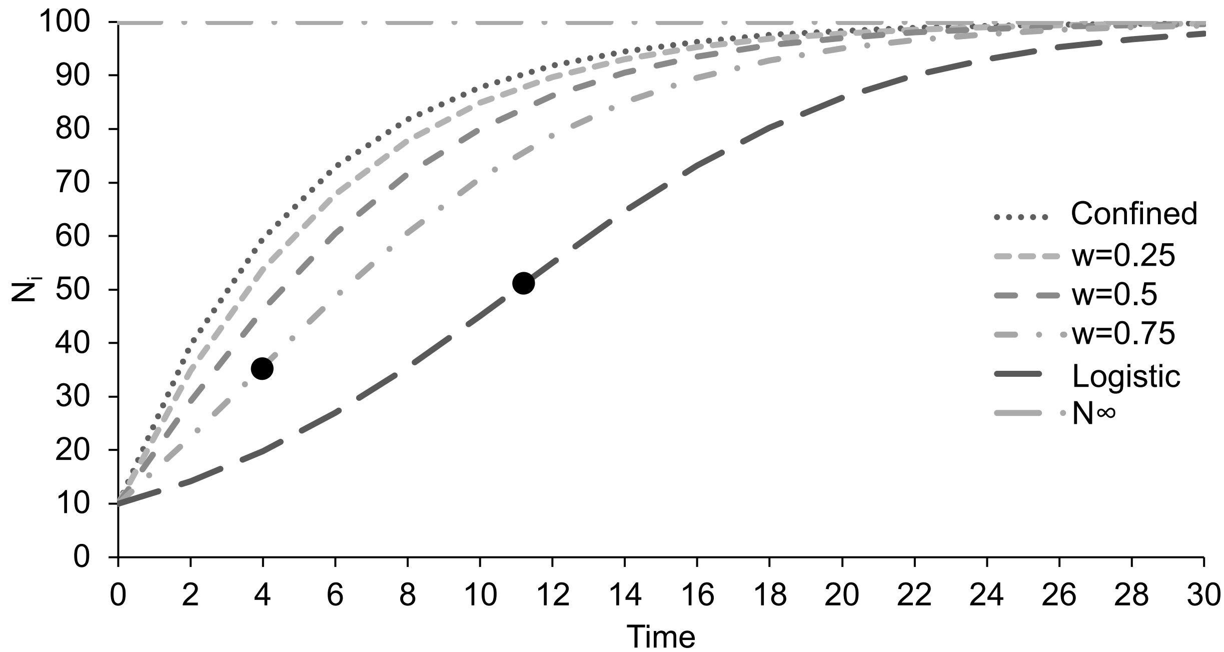

Figure 3. Exemplar curves for the hybrid model over a thirty-year period for incremental values of w between 0 and 1. The asymptotic N_* value of 100 displayed as dashed line. Inflection points are shown as dots on the growth curves.

Applications of the hybrid model. This model has previously been used for various types of diffusion involving external and internal influences, such as innovation diffusion (Haynes, Mahajan, & White, 1977; Lekvall & Wahlbin, 1973; Sharif & Ramanathan, 1981) and the diffusion of news and rumors (Bartholomew, 1981), as well as the spread of epidemics (Bartholomew, 1981). However, this model has not been estimated using SEM (structural equation modeling) software.

Model Equations

In this section, the logistic and confined exponential differential equations will be presented. The hybrid model will then be created by combining those differential equations and integrating the result. An extended discussion of this derivation is presented in Banks (2013).

To begin, the logistic and confined exponential models are presented by first introducing their respective differential (dimension-free) equations. These equations are then integrated to produce the curves as functions of time. In similar fashion, the differential equations of the logistic and confined exponential models are then used to create the differential equation for the hybrid model which is then integrated to produce the hybrid curve as a function of time.

Logistic model. The logistic model has the following differential equation:

$$\frac{dN}{dt} = aN - bN^2 \ , \tag{8}$$

where *a* is the growth coefficient and *b* represents the crowding coefficient. Banks (2013) shows that the solution to this is the following equation,

$$N_i = \left[\frac{1}{N_*} + \left(\frac{1}{N_0} - \frac{1}{N_*}\right) e^{-at}\right]^{-1} \tag{10}$$

which yields the following form of the logistic model:

$$N_i = N_*[1 + \left(\frac{N_*}{N_0} - 1\right) e^{-at}]^{-1} \quad . \tag{11}$$

In this equation, the growth coefficient, *a*, is assumed to be an unknown but positive constant. Accordingly, as time, *t*, increases, the second term in the denominator approaches zero, and the value of N_i approaches N_*. Because of this, N_* is the asymptotic value of N_i for large values of *t*. N_0 is the expected value of *N* when $t = 0$.

To simplify Eq. 11, we will rewrite it as:

$$N_i = \frac{N_*}{1+ke^{-at}} \quad , \tag{12}$$

where:

$$k = (N_*/N_0) - 1 \quad . \tag{13}$$

Confined exponential model. The confined exponential model has the following differential equation:

$$\frac{dN}{dt} = a_*(N_* - N_i) \quad , \tag{14}$$

in which a_* is a growth coefficient and N_* is the carrying capacity. For now, a_* and N_* are considered to be unknown but estimable constants. At the initial condition, $N(0) = N_0$, the solution becomes:

$$N_i = N_* - (N_* - N_0)e^{-a_* t} \quad , \tag{15}$$

which is labeled as the confined exponential equation.

Hybrid model. To create the hybrid model, the differential equations of the confined exponential and logistic are combined as follows:

$$\frac{dN}{dt} = a_*(N_* - N_i) + aN_i\left(1 - \frac{N_i}{N_*}\right) \quad . \tag{16}$$

To write it in dimensionless form, we set $U = N_i/N_*$, $T = at$, and $w = a/(a + a_*)$. If $w = 0$ (i.e., $a = 0$), Eq. 16 reduces to the confined exponential; if $w = 1$ (i.e., $a_* = 0$), it becomes the logistic. Because of this, w acts as a measure of how logistic the model is (with higher values indicating more logistic). Substituting in these values yields:

$$\frac{dU}{dt} = \left(\frac{1-w}{w} + U\right)(1 - U) \quad . \tag{17}$$

This dimensionless equation relates the growth rate, dU/dT, to U for values of w between zero and one.

The case $w = 1/2$ is particularly important to consider, because Eq. 17 becomes:

$$\frac{dU}{dt} = 1 - U^2 \quad . \tag{18}$$

As Eq. 18 indicates, the maximum growth rate is achieved when $U = 0$. Since $w = 1/2$, then $a_* = a$, and we can say that the phenomenon is equally due to an “internal contagion” (logistic) and an “external influence” (confined exponential).

Integrating the differential equation yields the equations for the hybrid model.

$$N_i = N_* \frac{1-(ma_*/a)e^{-(a+a_*)t}}{1+me^{-(a+a_*)t}} \quad , \tag{19}$$

where:

$$m = \frac{\left(\frac{N_*}{N_0}-1\right)}{\left(\frac{N_* a_*}{N_0\, a}+1\right)} \quad . \tag{20}$$

The slope of the hybrid curve is:

$$n_i = \frac{dN}{dt} = N_* ma\left(1 + \frac{a_*}{a}\right)^2 \frac{e^{-(a+a_*)t}}{\left(1+me^{-(a+a_*)t}\right)^2} \quad , \tag{21}$$

and from the derivative, we can obtain the value of t at the inflection point:

$$t_i = \frac{1}{a+a_*} log_e m \tag{22}$$

and the value of N_i at the inflection point (see Table 1):

$$N_{inflect} = \frac{1}{2} N_* \left(1 - \frac{a_*}{a}\right) \quad . \tag{23}$$

Table 1
Numeric Example of Parameter Values for Various w-Values in the Hybrid Model.

w	a	a_*	m	t_i	N_i	Comments
0	0.00	0.20	0	-∞	-∞	Confined Exponential (See Figures 4 and 5)
1/4	0.05	0.15	0.29	-6.18	-100.00	
1/2	0.10	0.10	0.82	-1.00	0.00	(See Figures 6 and 7)
3/4	0.15	0.05	2.08	3.66	33.33	
1	0.20	0.00	9.00	10.99	50.00	Ordinary Logistic (See Figures 8 and 9)

Note. N_* and N_0 are assumed to be 100 and 10, respectively.

Current Study

Although specification of the hybrid model is straightforward, the question remains as to whether researchers can use structural equation modeling software to successfully recover the parameters of the model under various sample scenarios and with the relatively smaller number of measurement occasions typical in psychological research. Further, it is also unclear as to whether structural model fit can successfully discriminate between the estimated logistic, confined exponential, and hybrid models. To answer these questions, Mplus (Muthén & Muthén, 1998-2017) was used to conduct simulations to examine the utility of the hybrid curve, compared to the logistic and confined exponential curves, in modeling growth.

Simulation Study

Monte Carlo simulations were conducted for various values of w (the parameter that defines if the curve is logistic, confined exponential, or a hybrid) ranging from 0 to 1 in increments of 0.1. For each value of w, 1,000 simulations were conducted. The data used in the simulations were generated in Mplus. These simulations included 15 measurement occasions, sample sizes of n = 1,000, and intraclass correlations (ICC) of .75. Bayesian information Criterion (BIC) values were used to determine model preference in the simulations. Additionally, Bayes factor approximations (Jarosz & Wiley, 2014) were computed to aid in the interpretation of results.

Example. For a numerical example as outlined by Banks (2013), the following values were specified in addition to the previously outlined parameters: $N_0 = 10, N_* = 100$, and $(a + a_*) = 0.20$; we let $w = a/(a + a_*) = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$, and 1. Results are displayed in the Table 1 and Figure 3.

Convergence. During the simulations, the logistic model experienced convergence issues when the underlying data was confined exponential. Similarly, the confined exponential model experienced convergence issues when the underlying data was logistic. The hybrid model experienced no convergence issues (Table 2).

Table 2

Convergence Rates for Simulations in the Simulation Study by Model and w-value.

w	Logistic	Confined Exponential	Hybrid
0.0	82.1%	100.0%	100.0%
0.1	76.1%	100.0%	100.0%
0.2	84.9%	100.0%	100.0%
0.3	99.8%	100.0%	100.0%
0.4	100.0%	100.0%	100.0%
0.5	100.0%	100.0%	100.0%
0.6	100.0%	100.0%	100.0%
0.7	100.0%	100.0%	100.0%
0.8	100.0%	100.0%	100.0%
0.9	100.0%	100.0%	100.0%
1.0	100.0%	63.9%	100.0%

Purely confined exponential. As discussed earlier in the paper and shown in Table 1, when $w = 0$ the hybrid equation (Eq. 19) reduces to the confined exponential equation (Eq. 15). In the simulations, comparable growth curves (Figure 4) and growth rates (Figure 5) were found for the hybrid model and the confined exponential model.

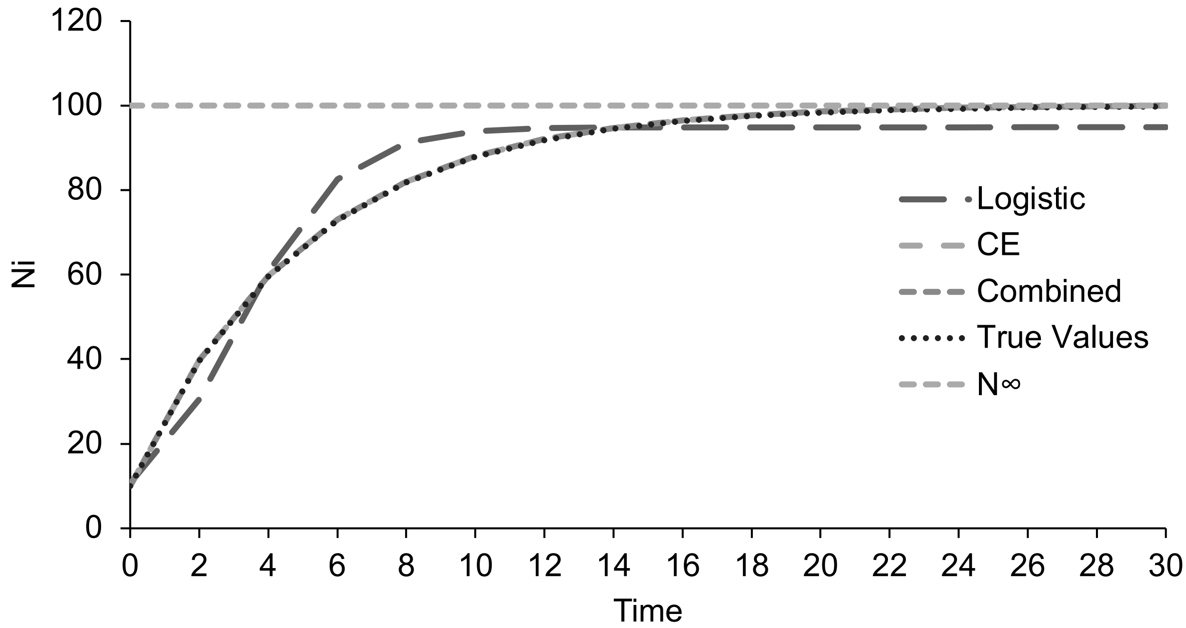

Figure 4. Growth curve for the hybrid model at w = 0.0. For reference, the logistic and confined exponential (CE) curves are also displayed.

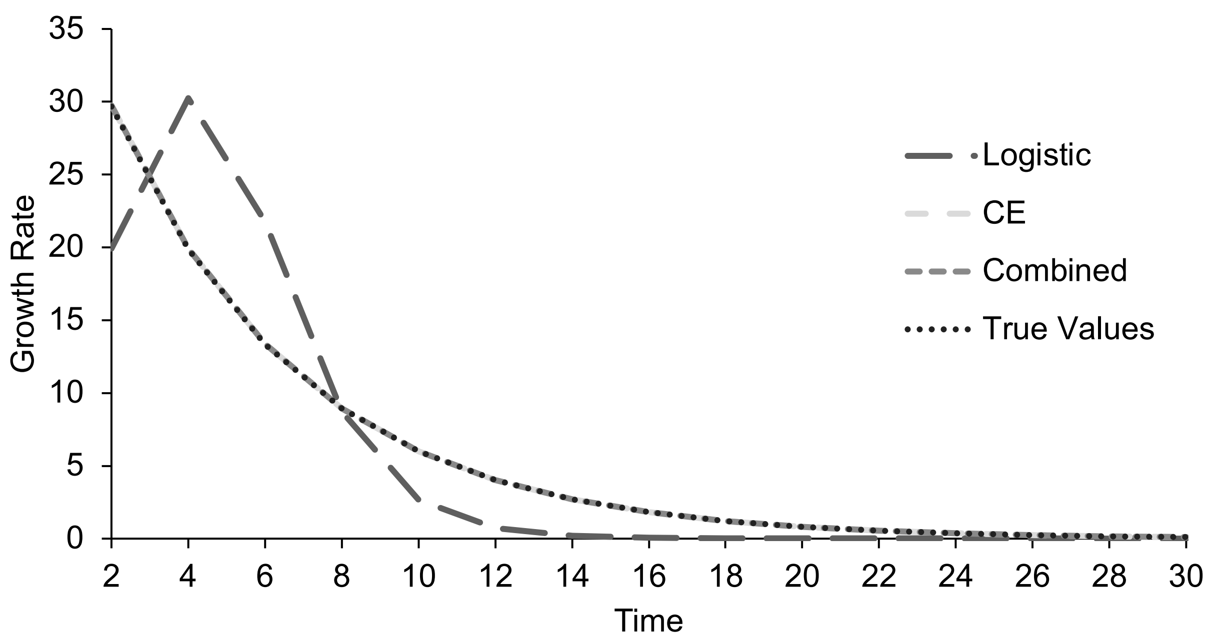

Figure 5. Growth rate for the hybrid model at w = 0.0. For reference, the logistic and confined exponential (CE) growth rates are also displayed.

Equally weighted logistic and confined exponential. When $w = 0.5$, the hybrid equation (Eq. 19) becomes equally logistic and confined exponential (as examined earlier in the paper). As expected, growth curves (Figure 6) and growth rates (Figure 7) for the hybrid model did not accurately reflect either of the other models.

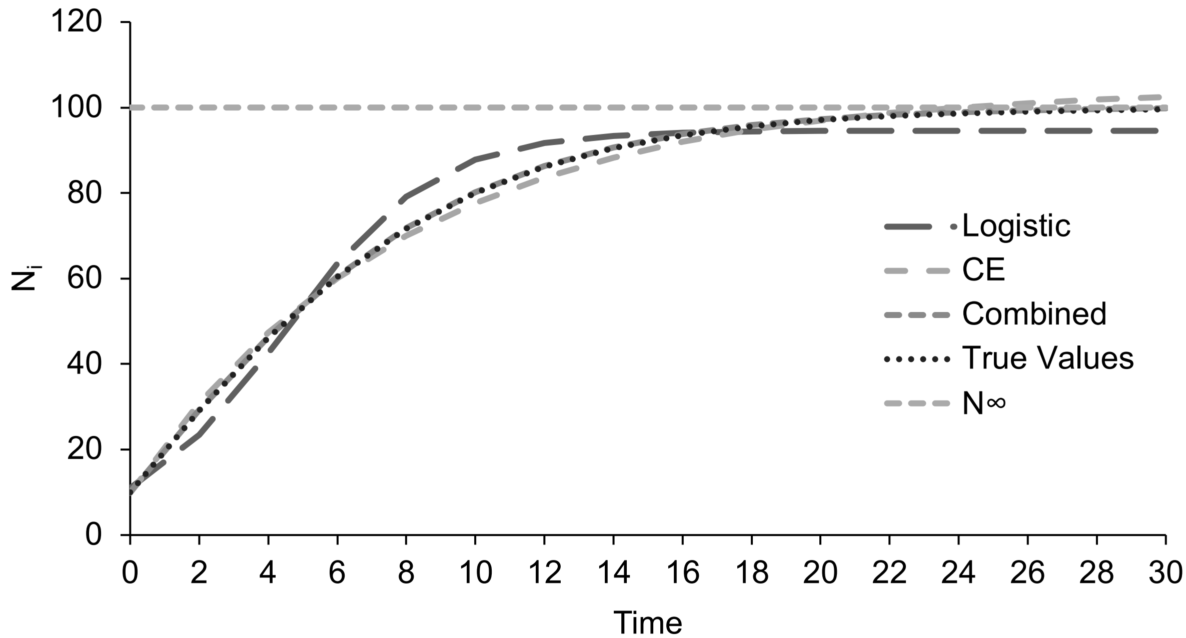

Figure 6. Growth curve for the hybrid model at w = 0.5. For reference, the logistic and confined exponential (CE) curves are also displayed.

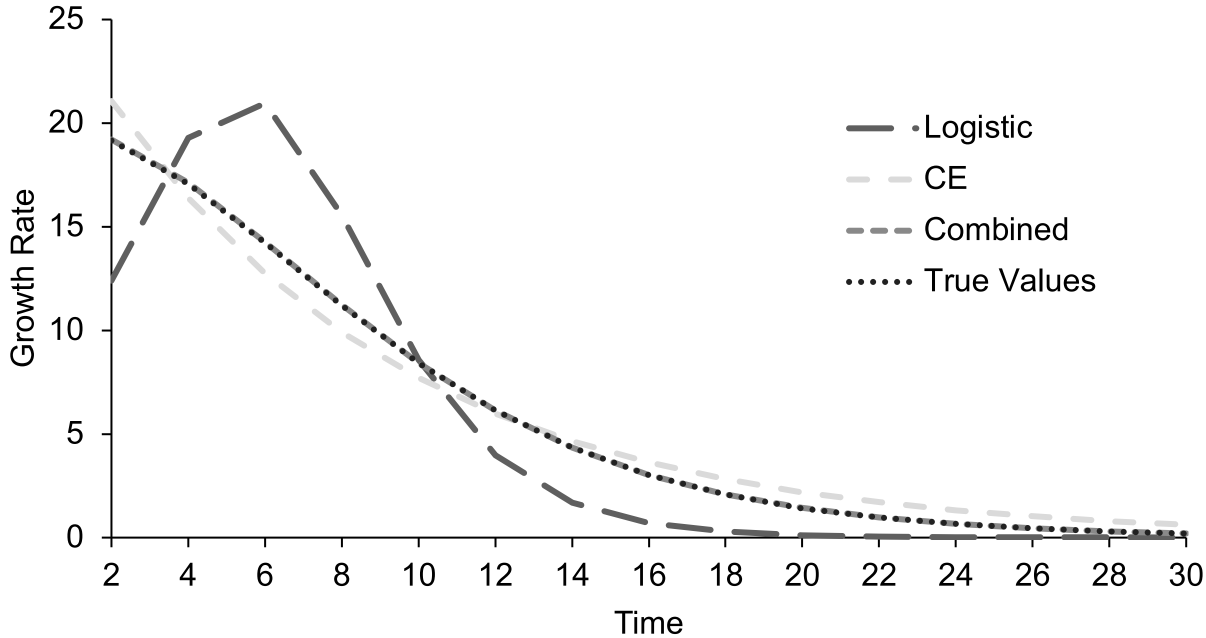

Figure 7. Growth rate for the hybrid model at w = 0.5. For reference, the logistic and confined exponential (CE) growth rates are also displayed.

Purely logistic. When $w = 1$, the hybrid equation (Eq. 19) reduces to the logistic equation (Eq. 12). Similar to what we saw with $w = 0$, in the simulations for $w = 1$ we see comparable growth curves (Figure 8) and growth rates (Figure 9) between the hybrid model and the logistic model.

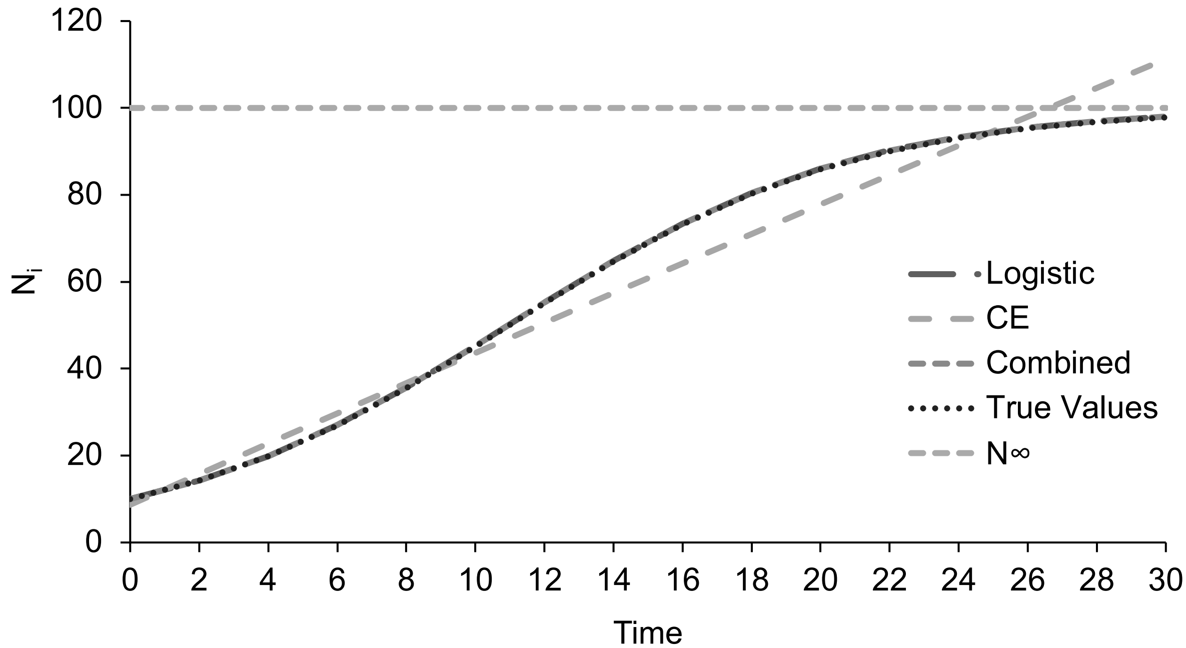

Figure 8. Growth curve for the hybrid model at w = 1.0. For reference, the logistic and confined exponential (CE) curves are also displayed.

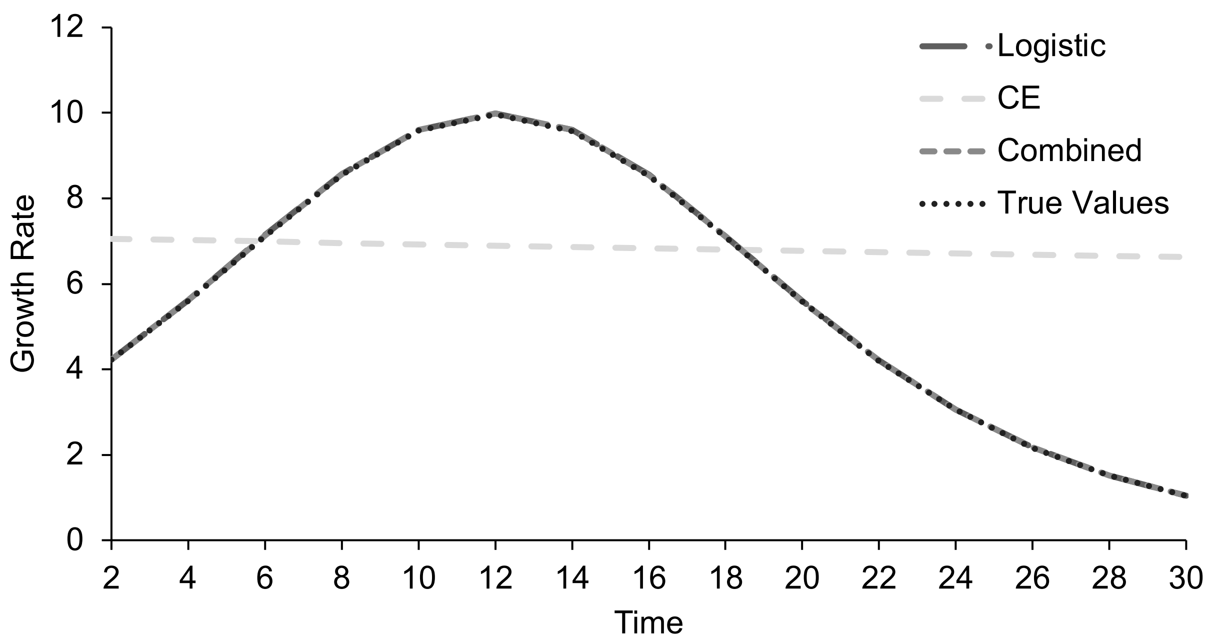

Figure 9. Growth rate for the hybrid model at w = 1.0. For reference, the logistic and confined exponential (CE) growth rates are also displayed.

Model comparison. To determine which models most accurately represented the data, BIC values were compared for each model across each increment of w. The confined exponential model was most preferred for values of w between 0 and 0.2, the hybrid model was preferred for values of w between 0.3 and 0.9, and the logistic model was preferred when $w = 1$ (Figure 10).

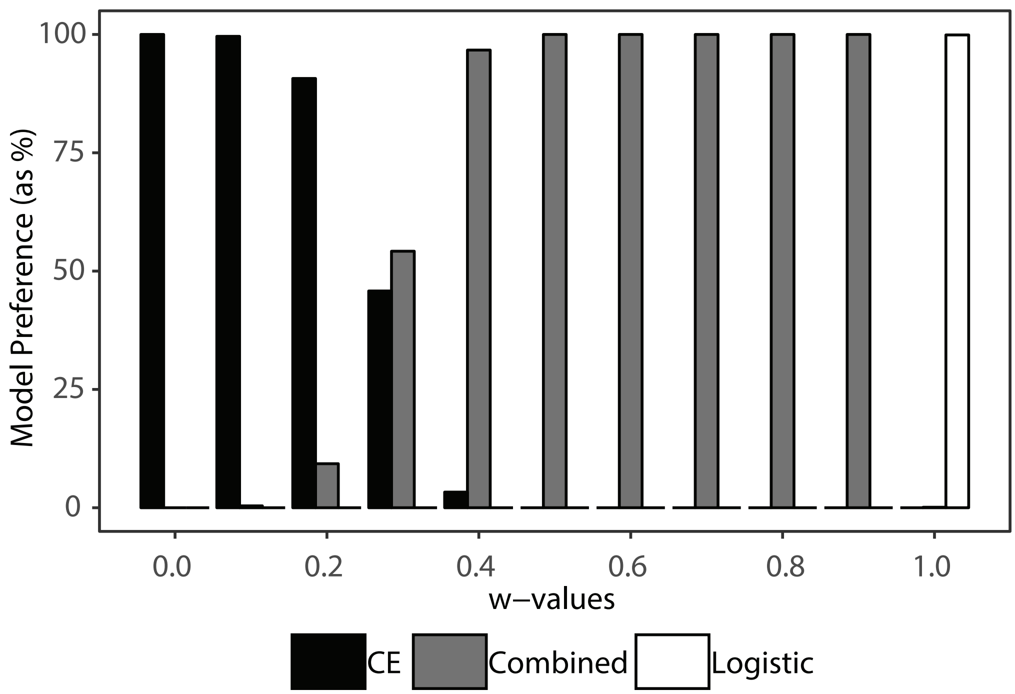

Figure 10. Percent of simulations that preferred the logistic, confined exponential, or the hybrid model as determined by BIC values across increments of w-values.

The BIC values were then used to calculate Bayes factor approximations (Jarosz & Wiley, 2014; BFAs). The BFAs indicated a similar pattern of model preference as the BICs. Examining the distributions of BFA values shows a large range of variability in the strength/weakness of the model preferences (Table 3). For example, the BFA values for the $w = 0.5$ simulations range from 4.64 to 8.69×10^{22}.

Table 3
Descriptive Statistics for the Distribution of Bayes Factor Approximations by Increments of w-Values.

w	Median	Mean	10% Quantile	90% Quantile	Minimum	Maximum
Hybrid Model Compared to the Confined Exponential Model						
0.0	1.20e-3	2.38e-3	9.98e-4	3.99e-3	5.19e-9	1.47e-1
0.1	2.95e-3	2.87e-2	1.11e-3	3.26e-2	1.96e-7	6.79e0
0.2	2.11e-2	4.64e0	2.16e-3	8.61e-1	1.00e-3	2.84e3
0.3	1.47e0	3.47e4	2.55e-2	3.74e2	1.19e-3	3.01e7
0.4	4.04e3	4.28e10	7.98e0	8.69e6	8.65e-3	4.08e13
0.5	3.91e9	8.86e19	6.39e5	1.07e14	4.64e0	8.69e22
0.6	5.66e19	1.49e34	4.21e14	5.00e25	2.52e6	1.48e37
0.7	5.13e36	1.14e56	4.09e29	4.16e44	3.60e17	1.12e59
Hybrid Model Compared to the Logistic Model						
0.8	1.53e55	1.60e74	7.19e45	1.92e65	2.34e36	1.02e77
0.9	8.81e21	3.24e35	1.06e16	2.83e28	2.06e9	2.99e38
1.0	2.07e-3	9.41e-3	1.12e-3	1.05e-2	9.98e-4	1.06e0

Note. BFAs > 1 indicate preference for the hybrid model. BFAs < 1 indicate preference for the alternative models.
Note. The hybrid model was compared to either the confined exponential or logistic based upon which was most preferred by the BICs.

Overall, the simulations provide strong support for the hybrid model, particularly in the midrange w-values. However, the generous sample size used leaves open the question of whether researchers can successfully estimate and compare models when fewer measurement occasions, smaller sample sizes, substantial missing data rates, or measures with different internal consistency (i.e., intra-class correlations (ICC) multiplied by the manifest variable variance) are used.

Method

Data Simulation

M*plus* simulations were conducted for the logistic, confined exponential, and hybrid models with variations of sample sizes (1,000, 500, 300, 100, 50), measurement occasions (15, 10, 5), intraclass correlation values (0.8, 0.7, 0.5), and w-values (ranging from 0 to 1 in 0.1 increments). All simulations had 1,000 replications and included 30% missing data. In total, 1,485 simulations were conducted.

Model Comparison

Bayesian information Criterion (BIC) values were used to determine model preference in the simulations. Additionally, Bayes factor approximations (BFAs) were computed to aid in the interpretation of results.

Results

Convergence Issues

Convergence issues occurred for a number of the model simulations. In particular, the simulations for the confined exponential model with a w-value of 1 (purely logistic) failed to complete after running for a month. The confined exponential model did not display any other convergence issues.

On the other hand, the logistic model simulations with a w-value of 0 (purely confined exponential) completed relatively quickly; however, the convergence rates were very poor (near 10%). Additionally, consistent convergence issues (rates of 80 – 90%) occurred for the logistic model when the w-value was near 0 (e.g., 0.1 – 0.3).

Convergence issues also occurred for the hybrid model. Issues typically occurred during simulations that included lower values of simulation parameters (sample size, measurement occasions, ICC), particularly for w-values near, but not including, 0. For example, the convergence rate was 74.8% for the simulations with n = 100, t = 5, ICC = .50, and w = 0.1. Interestingly, as the simulation parameters began to include combinations of lower parameter values, the convergence issues spread to higher w-values. For example, the convergence rate for n = 100, t = 5, ICC = .50, and w = 0.5 was 87.1%.

Parameter Recovery

Based on these simulations, it seems possible to recover parameters of the hybrid model in a variety of realistic real-world scenarios. This is somewhat tempered by the convergence of the models at the end points of the w-values, however. As a result, researchers may wish to consider separately estimating the logistic or confined exponential models explicitly and not relying solely on the hybrid model approach.

Model Preference

BIC values were used to determine preference between the three models for the various combinations of parameters used in the simulations. Visually examining the patterns that emerge from graphing the proportions of model preference leads to the realization that decreases in ICC and measurement occasions, individually, cause a decrease in the preference for the hybrid model. Additionally, simultaneous decreases in these parameters cause even more pronounced decreases in hybrid model preference. However, with a sample size of 1,000, the hybrid model is still largely preferred when there is adequate measurement reliability and enough measurement occasions (see Figure 11). As would be expected, the confined exponential and logistic models are consistently preferred for w-values of 0 and 1, respectively.

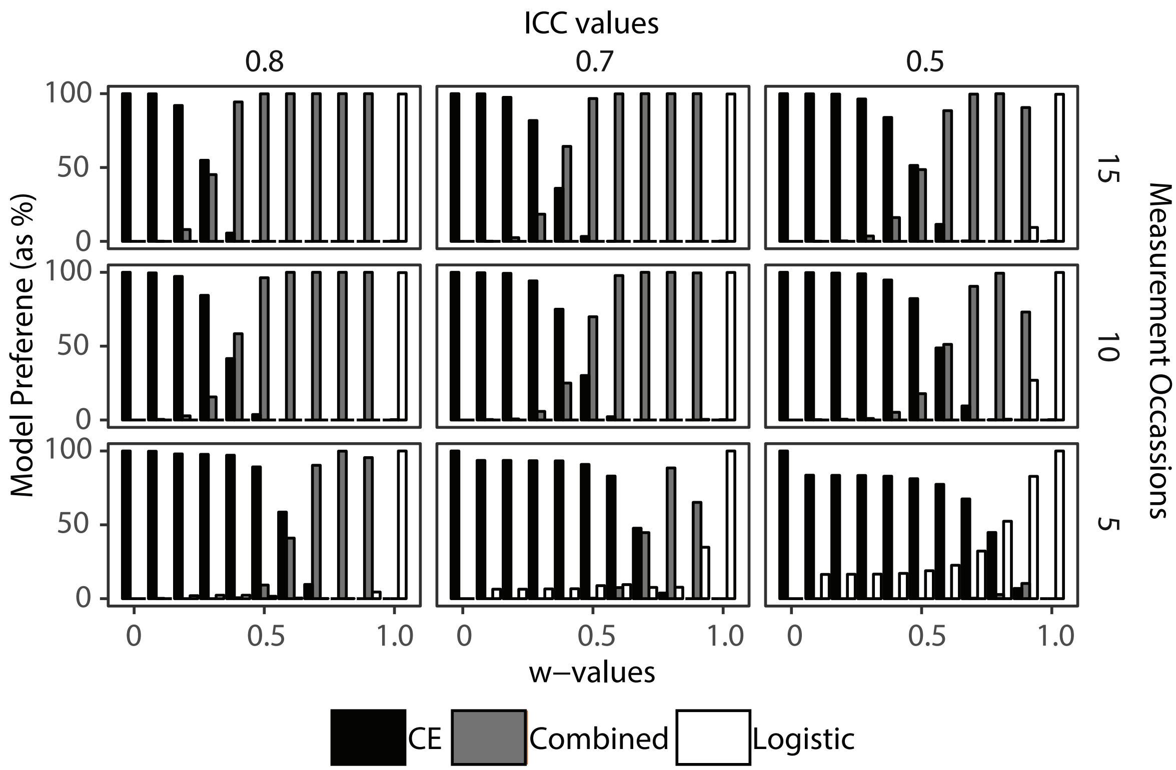

Figure 11. Preference between models, as determined by BIC values, across all combinations of measurement occasions, ICC values, and w-values for a sample size of 1,000.

A similar pattern (decreasing preference for the hybrid model as the parameters decrease in value) emerges for other sample sizes. For a sample size of 500, the hybrid model is still adequately preferred when the other parameter values are high (See Figure 12), but the hybrid model is never the most preferred model for low-value combinations of the parameters. Comparison of Figures 11 and 12 shows that preference for the hybrid model also decreases as a function of sample size.

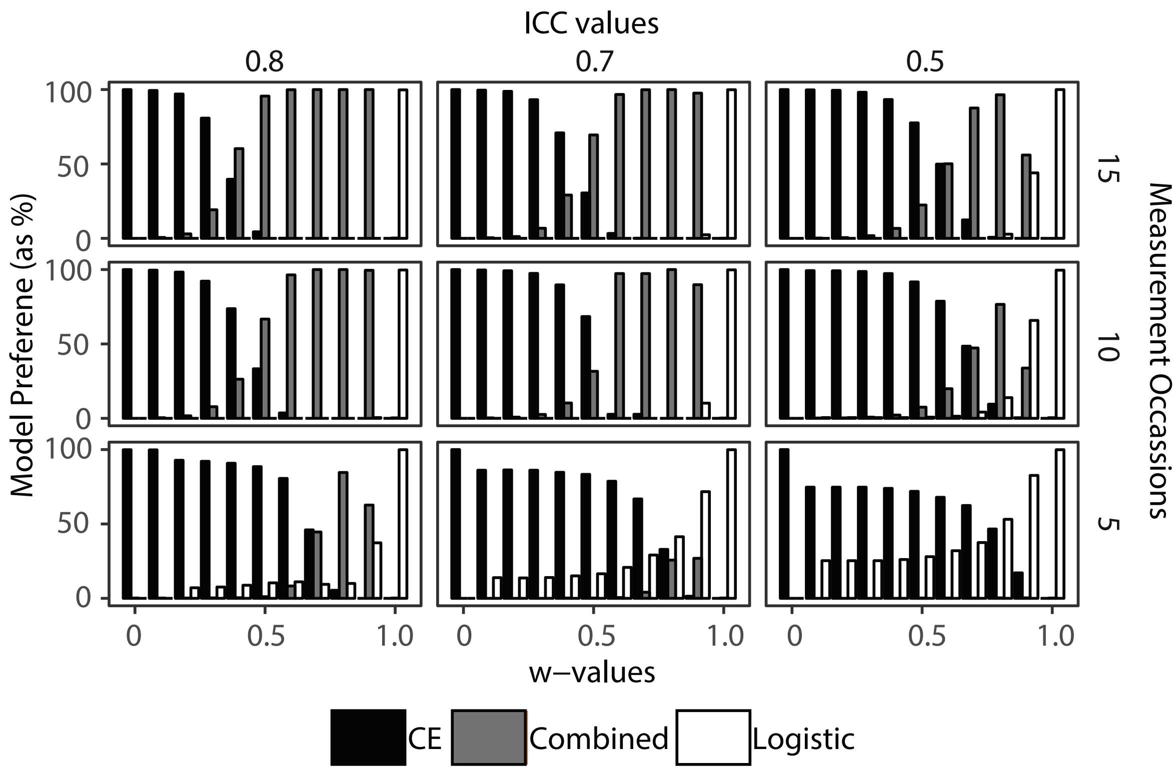

Figure 12. Preference between models, as determined by BIC values, across all combinations of measurement occasions, ICC values, and w-values for a sample size of 500.

For a sample size of 300 (see Figure 13), the hybrid model is still preferred for many w-values when paired with high values of ICC and enough measurement occasions. However, when the sample size decreases further (100 and 50, in this study), the hybrid model is rarely the preferred model even when paired with an ICC of 0.8 and 15 measurement occasions.

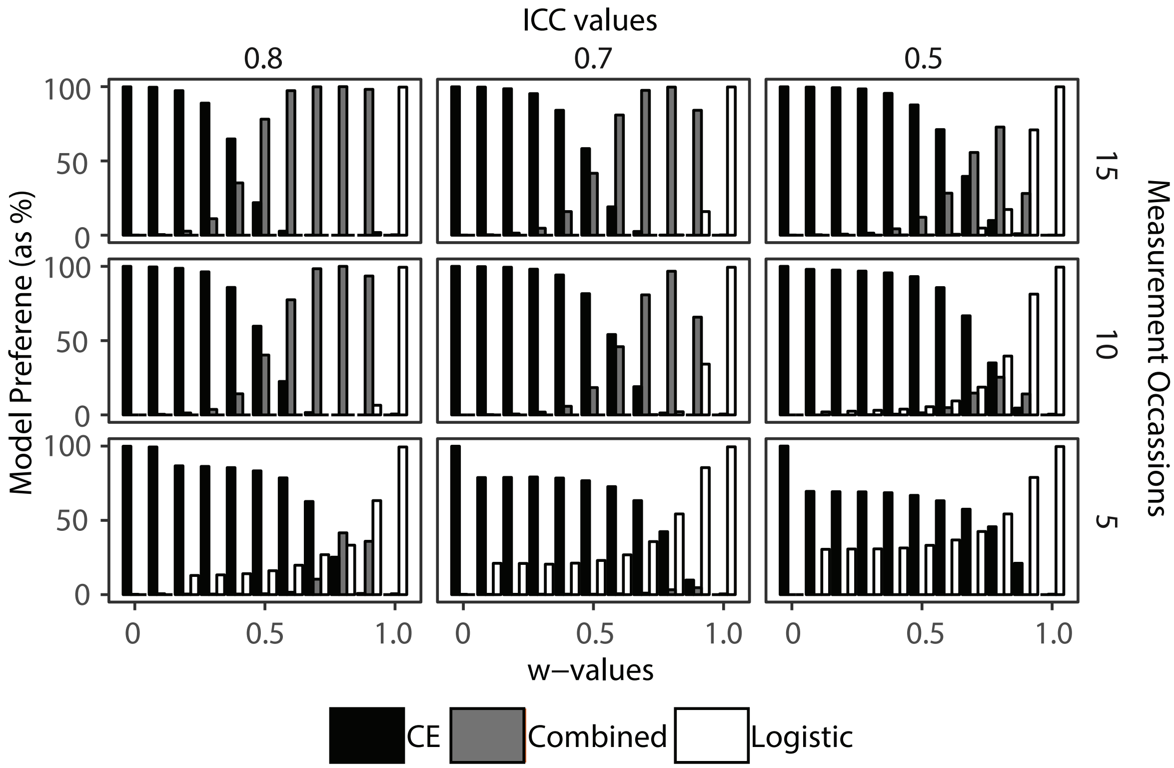

Figure 13. Preference between models, as determined by BIC values, across all combinations of measurement occasions, ICC values, and w-values for a sample size of 300.

Discussion

The results of this study demonstrate the ability to create and implement hybrid models in SEM software, as well as the potential utility for a model that combines the logistic and confined exponential models. Given how widely used these models are, the hybrid model could potentially act as a diagnostic tool for determining if the previously mentioned models are actually the best choices in their current applications.

Potential Applications

Practically speaking, use of a hybrid model has many applications. For example, many categorical models assume that the link function is logistic in nature. Use of a hybrid model provides researchers with a way of testing whether this assumption is, in fact, reasonable.

Psychopathology and learning models similarly assume a single parametric form for growth. It may be, for example, that logistic growth, which assumes an internal, self-propagating process, may be inaccurate and that forces external to the individual, to some extent, drive the longitudinal course of pathology. If this is the case, then neither logistic nor confined exponential growth would accurately reflect the true growth pattern of pathology. However, the hybrid model would allow researchers to entertain intermediary models that can have varying degrees of being logistic or of being confined exponential.

Similarly, the hybrid model would allow researchers to examine the relationship between the beta response-strength model of learning (Luce, 1959) and the linear-operator model (Bush & Mosteller, 1955) by simply allowing them, like with the previous example, to commingle. This could be used, for example, to suggest pools of items to include in tests that reflect both the logistic and confined exponential models of learning.

Limitations and Future Directions

The most notable limitation of this study is that the extent of missingness in the data was consistent (30%) throughout the simulations and was therefore assumed to be ignorable, which it likely is not. To that end, it would be worthwhile to consider the extent of missingness in your data when using a hybrid model. It is also important to mention that researchers wishing to use this technique should investigate whether influential observations are present in their data which could unduly influence the parameters of the hybrid model and could affect the adjudication between the hybrid, logistic, and confined exponential models.

Another avenue of future research involves the use of base models other than the logistic and confined exponential. This study has demonstrated that a combination of these two base models could be useful in various applications. However, there are many circumstances where different base models could be thought to represent related mechanisms of growth (e.g., the logistic and confined exponential representing two different transfer mechanisms). Thus, it may be worthwhile to create different hybrid models that are applicable in these varying circumstances.

Conclusions

The creation and implementation of hybrids of commonly used models offers a potential addition to traditional model comparison approaches. This approach would allow for the possibility of intermediary models in addition to the original discrete models.

References

Banks, R. B. (2013). *Growth and diffusion phenomena: Mathematical frameworks and applications* (Vol. 14). Berlin: Springer Science & Business Media. doi: 10.1007/978-3-662-03052-3

Bartholomew, D. J. (1981). *Mathematical methods in social science.* New York: Wiley.

Batschelet, E. (1979). *Introduction to mathematics for life scientists (3rd ed.)*. Berlin – Heidelberg: Springer Science & Business Media. doi: 10.1007/978-3-642-96080-2

Bird, R. B., Stewart, W. E., Lightfoot, E. N., & Meredith, R. E. (1961). Transport phenomena. *Journal of the Electrochemical Society*, *108*(3), 78C-79C. doi: 10.1149/1.2428074

Bollen, K. J. (2007), On the origins of latent curve models. In R. Cudeck and R. C. MacCallum (Eds.) *Factor analysis at 100* (pp. 79-97). Mahwah, NJ: Erlbaum. doi: 10.4324/9780203936764

Bonney, G. E. (1986). Regressive logistic models for familial disease and other binary traits. *Biometrics, 42*(3), 611-625. doi: 10.2307/2531211

Box, G. E. P. (1950). Problems in the analysis of growth and wear curves. *Biometrics, 6*(4), 362-389. doi: 10.2307/3001781

Brody, S. (1945).*Bioenergetics and growth*. New York: Reinhold Publishing.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning.* Oxford, England: Wiley. doi: 10.1037/14496-000

Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling, 16*(4), 676-701. doi: 10.1080/10705510903206055

Grimm, K. J., Ram, N., & Estabrook, R. (2016). *Growth modeling: Structural equation and multilevel modeling approaches*. New York: Guilford.

Grimm, K. J., Steele, J. S., Ram, N., & Nesselroade, J. R. (2013). Exploratory latent growth models in the structural equation modeling framework. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(4), 568-591. doi: 10.1080/10705511.2013.824775

Haynes, K. E., Mahajan, V., & White, G. M. (1977). Innovation diffusion: A deterministic model of space-time integration with physical analog. *Socio-Economic Planning Sciences, 11*(1), 25-29. doi: 10.1016/0038-0121(77)90043-x

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*(1), 2-9. doi: 10.7771/1932-6246.1167

Kreith, F. (1958). *Principles of heat transfer.* Scranton, PA: International Textbook Co.

Lekvall, P., & Wahlbin, C. (1973). A study of some assumptions underlying innovation diffusion functions. *The Swedish Journal of Economics, 75*(4), 362-377. doi: 10.2307/3439146

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* New York: Wiley. doi: 10.1037/14396-000

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

Muthén, L. K.., & Muthén, B. O. (1998-2017). *Mplus User's Guide (Eighth Edition).* Los Angeles, CA: Muthén & Muthén.

Oliver, F. R. (1964). Methods of estimating the logistic growth function. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *13*(2), 57-66. doi: 10.2307/2985696

Oliver, F. R. (1982). Notes on the logistic curve for human populations. *Journal of the Royal Statistical Society. Series A (General), 145*(3)*,* 359-363. doi: 10.2307/2981868

Oren, S. S., & Schwartz, R. G. (1988). Diffusion of new products in risk-sensitive markets. *Journal of Forecasting, 7*(4), 273-287. doi: 10.1002/for.3980070407

Panik, M. J. (2014). *Growth curve modeling: Theory and applications.* Hoboken, NJ: Wiley. doi: 10.1002/9781118763971

Pearl, R., & Reed, L. J. (1920). On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences, 6*(6), 275-288. doi: 10.1073/pnas.6.6.275

Potthoff, R. F., & Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika, 51*(3-4), 313-326. doi: 10.1093/biomet/51.3-4.313

Ralston, B. (1983). The dynamics of communication. In Griffith, D.A. & Lea, A.C. (Eds.), *Evolving geographical structures: Mathematical models and theories for space-time processes* (pp. 130-167). Dordrecht, Netherlands: Springer. doi: 10.1007/978-94-009-6893-6_7

Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics, 14*(1), 1-17. doi: 10.2307/2527726

Schultz, H. (1930). The standard error of a forecast from a curve. *Journal of the American Statistical Association, 25*(170), 139-185. doi: 10.1080/01621459.1930.10503117

Sharif, M. N., & Ramanathan, K. (1981). Binomial innovation diffusion models with dynamic potential adopter population. *Technological Forecasting and Social Change, 20*(1), 63-87. doi: 10.1016/0040-1625(81)90041-x

Thurstone, L. L., & Ackerson, L. (1929). The mental growth curve for the Binet tests. *Journal of Educational Psychology, 20*(8), 569. doi: 10.1037/h0070160

Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, *23*(1), 19-23. doi: 10.1007/bf02288975

Valentine, H. T. (1985). Tree-growth models: Derivations employing the pipe-model theory. *Journal of Theoretical Biology*, *117*(4), 579-585. doi: 10.1016/s0022-5193(85)80239-3

Verhulst, P. F. (1838). Notice sur la loi que la population suit dans son accroissement. *Correspondence in Mathematical Physics, 10*, 113-126.

Wishart, J. (1938). Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. *Biometrika, 30*(1/2), 16-28. doi: 10.2307/2332221

Wood, P. K., Steinley, D., & Jackson, K. M. (2015). Right-sizing statistical models for longitudinal data. *Psychological Methods*, *20*(4), 470-488. doi: 10.1037/met0000037