# Intra and Interindividual Variation Modeling: Bayesian Mixed-Effects Nonstationary Latent Differential Equation Model

By

## Mauricio Garnier-Villarreal

David K. Johnson, Chairperson

_____

Pascal R. Deboeck

_____

Committee members

Wei Wu

_____

Amber Watts

_____

William P. Skorupski

Date defended: _____ July 19, 2016

The Dissertation Committee for Mauricio Garnier-Villarreal certifies
that this is the approved version of the following dissertation :

Intra and Interindividual Variation Modeling: Bayesian Mixed-Effects Nonstationary Latent
Differential Equation Model

<hr/>

David K. Johnson, Chairperson

Date approved: _____ July 19, 2016 _____

# Abstract

Longitudinal analysis are powerful methods to estimate change over time. The combination of nomothetic and idiographic approaches within longitudinal analysis would allow to answer questions related to intra and interindividual variability in one integrated method. In order to have lag independent results, longitudinal analysis should be made with a continuous-time method. Continuous-time methods take into account that the phenomena of interest does not stop existing between measurement time points. Differential equations modeling is a method that studies intraindividual variability as a form of continuous-time modeling, which can be implemented as fixed-effects or mixed-effects. Mixed-effects models allows to integrate interindividual variability, and properly estimate non-dependent data. Latent Differential Equation (LDE) model is a method to estimate differential equation models from a familiar framework in psychology (Structural Equation Modeling). This dissertation tend to extend the LDE by adding the mixed-effects, estimating subject and sample parameters, including interindividual variability on parameters of interest. The analyses were implemented from the Bayesian framework, this framework provides several advantages, one of them being that allow us to make direct inference of the estimated parameters. The proposed model (Bayesian Mixed-Effects Nonstationary Latent Differential Equation Model, BMNLDE) was tested in a simulation, and exemplify with a real data describing the sedentary behavior in older adults. The simulation shows that the BMNLDE model estimate parameters with low bias, the 95% Credible Interval coverage is unreliable when the model presents low frequency and high damping. The frequency of the oscillating processes was the main factor that affected bias and CI coverage. The

BMNLDE model showed to be an appropriate method to include intra and interindividual variability. The simulation was capable to demonstrate the conditions in which the model performs as desire, and under which conditions the model does not perform as desire. The real data example shows an application of the BMNLDE model, were the BMNLDE model describes the oscillating behavior of the sedentary behavior. It also shows how it can used to compare parameters of interest between groups.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Longitudinal analysis is a gold standard method in psychology for studying change over time (McArdle & Nesselroade, 2014) because it can combine both nomothetic and idiographic investigation that are fundamental to the behavioral sciences. The nomothetic approach looks for the overall population's true processes, whereas the idiographic approach seeks to understand the unique individual process (Maxwell & Boker, 2007; Molenaar, 2004).

Understanding how an individual changes over time and how these changes are similar or different by group is at the center of many research fields, like developmental, cross-cultural, and clinical trials. The research focus on describing and comparing groups gains information at a macro level in order to generalized to larger groups of populations, while at the same time can loose perspective of how individuals behaviors within these groups. McArdle & Nesselroade (2014) narrowed the term *change* to refer to intraindividual change, focusing on the idiographic approach. The ideographic approach places the focus of psychological research on the individual, psychological research can only be considered complete until it includes the idiographic analysis as well as the nomothetic analysis (Molenaar, 2004).

The most powerful of longitudinal models partition variance associated with the repeated measurement of a single individual (within subject variability) and the variance associated with the group (between-subject variability) to measure the effects of each model component, also know as

mixed-effects modeling. Mixed-effects models yield better estimates than traditional fixed-effects (only) models that assume the same characteristics of change describe all subjects equally well. Fixed effects models assume that every subject follows the same functional form with the same parameter values (Maxwell & Boker, 2007). While a mixed-effects longitudinal model explicitly estimates the interindividual variability, it is assumed that each subject follows the same functional form of change, but there is no assumption that each subject has the same parameter values. The specific nature of change is allowed to differ across subjects (Maxwell & Boker, 2007; Pinheiro & Bates, 2000). Mixed-effects longitudinal models allow researchers to estimate intraindividual change and interindividual variability because they estimate how a longitudinal model fit each participant and to estimate how much participants vary between one another. Mixed-effects models are strong models of behavior because they allow investigators to examine the dynamics of a single individual and to generalize the dynamics to a population of individuals (Maxwell & Boker, 2007).

Although many mixed-effects take into account within and between subject variability, most do not emphasizes the role of time (Voelkle et al., 2012). Time is a crucial factor that leads to mounting change, investigators usually neglect time by failing to treat it as continuous. For instance most repeated measures analyses treat time as discrete (e.g., pre-post treaments) and the continuity of time is overlooked. When researchers treat time as discrete they create lag-dependent results (Oud, 2007)—assuming that time and change stop between measurements. Continuous time represents a framework in which researchers perform longitudinal analyses without the assumptions of discrete time.

There have been recent developments for the practical application of continuous-time models in psychology by applying common calculus (derivatives) to repeated measures data (Boker et al., 2004; Boker, 2007; Voelkle et al., 2012; Oravecz et al., 2011). These differential equation-based continuous time models are commonly implemented with fixed-effects, but there is little work to be found on the application of mixed-effects differential equations in psychology (e.g., Oravecz et al., 2011). The combination of differential equations with mixed-effects models allows us to estimate the intra and interindividual variability, which grants more flexibility and leads to a wider

range of meaningful research questions.

The Latent Differential Equations (LDE) presented by Boker et al. (2004), allows us to approximate derivatives from a familiar framework in psychology, Structural Equation Modeling (SEM). This allows researchers to estimate a variety of differential equation models with the advantages of SEM (Boker et al., 2004; Boker, 2007; Steele & Ferrer, 2011). Boker et al. (2016) developed an extension of the LDE, where some parameters in the model are estimated for each subject but lack the estimated sample averages and the variance of these individual estimates.

In this dissertation, I assume the task of developing an extension of the LDE model to be used in psychological research (Boker et al., 2004; Boker, 2007; Boker et al., 2016) that also accounts for interindividual variability with mixed-effects (intraindividual variability), this model is implemented within the Bayesian framework. The model proposed will be tested with a simulation study to determine its performance. Subsequently, the model will be apply to a real data example to characterize sedentary behavior in older adults.

## 1.1 Continuous Time and Differential Equations

### 1.1.1 Continuous Time

Though we are all embedded in the constant flow of time, behavior scientists can only measure an individual at discrete moments in time. Further the most common repeated measures analyses treat time as discrete (Oud, 2007; McArdle & Nesselroade, 2014). By using discrete-time models, researchers assume that the variables exist only at the measured time points. However, by using continuous-time models, researchers assume that the variable exists in an infinite number of time points from which the measured time points are a finite set. Discrete-time models are an oversimplification and often a distortion of reality (Oud, 2007).

Methods of analysis that treat time as a continuum exist, but these methods are not yet widely used in psychological research (Voelkle et al., 2012; Bertenthal, 2007; Brown, 2007; Oravecz et al., 2011; Boker et al., 2004; Boker, 2007). Moving forward, the field of psychology can learn

from other fields such as pharmacokinetics and pharmacodynamics (PK/PD) (Tornoe et al., 2004; Wang et al., 2008), finance (Ait-Sahalia, 2009; Gander & Stephens, 2007), and environmental studies (Tomassini et al., 2009; Israel & Migon, 2012), in which continuous time is commonly implemented.

A common way to handle time as continuous with discrete-time measures is with the use of differential equations (DE). DEs are commonly used in the physical and natural sciences and in finance (Bertenthal, 2007; Ait-Sahalia, 2009; Klim et al., 2009; Andersen & Hojbjerre, 2005). Continuous-time models estimate the data-generating process; for example, this generating process can be seen as the process by which two variables and their relationship evolve over time (Voelkle et al., 2012). Continuous-time models assume that phenomena exist at every infinitesimal point in time, but researchers can only measure it at finite time points. The generating process explains the underlying continuous process that was measured finitely.

Unless the process is discrete in time, there is no reason to prefer discrete-time analysis over continuous time because the continuous time contains the same information as the discrete-time model and more (Voelkle et al., 2012). Knowing the discrete-time parameters tells researchers little about the underlying continuous-time model; the discrete-time model only reflects the parameters in function of the measured-time intervals. Because the discrete-time models are dependent on lag, their estimates are limited to the effect at the measured interval or multiple measured intervals. Discrete-time results are incomparable when the measured intervals differ between studies. Discrete-time models may return seemingly contradictory results because the effect at a specific measured interval can have a different direction than the effect at another measured interval (Oud, 2007; Oud, 2007). By knowing the continuous-time parameters under the correct model (i.e., generation process) researchers can calculate the effect at infinite lags (Oud, 2007). Continuous-time parameters usually have direct and useful interpretations (Boker, 2007).

Continuous-time models present several advantages—one of the greatest being that they are theoretical and can be use to develop theories relating to social phenomena (Oravecz et al., 2011). The specification of a model is limited only by a researcher's theory (Brown, 2007). DEs can

be formulated to describe a great variety of systems. There are theoretical and methodological advantages of modeling; namely, that the psychological traits should be regarded as constructs with properties that continuously evolve over time (Tiberio, 2008). The use of continuous-time models allows researchers to evaluate the internal and external dynamics of the different psychological traits. If researchers accept the statement that psychological constructs change continuously over time, then they should treat time as a continuous variable in the models (Boker, 2007; Tiberio, 2008). A popular kind of system may be observed when the current state of the variables in the systems are dependent, to some degree, on the previous states of the system (Boker, 2007; Steele & Ferrer, 2011).

There are two principal reasons why continuous-time models are not as widely used in the social sciences as in the physical sciences: Theory and Legacy (Brown, 2007). Historically, many advances in social science came from population surveys conducted in the 1950s. The data were analyzed using cross-tabulation and correlation methods. Social scientists have been trained to follow this framework, which resulted in this statistical training becoming rote. Traditionally statistics has focused on the implementation of established structured tests (such as $t-test$ and correlation). Computationally they have the advantage of being design to make analysis simple to test, since originally they had to be performed without computers. But these tests tend to be inflexible and fragile; meaning that they have limited ways to adapt to unique research contexts, and they fail in unpredictable ways when applied to new contexts. These classical tools are not diverse enough to handle research questions that demand greater time resolution, making them inappropriate for more complex research (Brown, 2007; McElreath, 2016).

Continuous-time models have been introduced into psychology from different fields of science (Steele & Ferrer, 2011; Hu et al., 2014; Oud, 2007). It has been attributed it (in part) to a need to ask questions that classical techniques cannot answer and to researchers being exposed to new models that can easily be applied to psychology. One of the reasons for the underuse of continuous-time models in psychology is because they heavily employ mathematical terms—for conceptual foundations as well as analytic techniques. For this reason, many concepts are inaccessible to the

average psychological scientist (Brown, 2007; Bertenthal, 2007; Oud, 2007; Voelkle et al., 2012).

Physical and natural sciences use DEs because they want to approximate the nature of continuous time in a specific study (Brown, 2007; Bertenthal, 2007). Despite the recent attention paid to these models in psychological research (Boker et al., 2004 ; Oud, 2007; Voelkle et al., 2012; Deboeck et al., 2008; Deboeck et al., 2009; Boker et al., 2009; Boker & Nesselroade, 2002; Boker, 2007; Steele & Ferrer, 2011), there are substantial differences in how the continuous-time models are interpreted in the physical sciences as opposed to the social sciences. In the physical sciences, researchers can assume that the system's characteristics can be generalized in relevant ways; they can assume that the same mathematical equation can describe the dynamics of a pendulum without knowing any details of a specific pendulum. But once researchers use continuous-time models to describe human behavior (e.g., insight, emotional states, etc.), they must take into account the variability and differences between individuals (Maxwell & Boker, 2007).

## 1.1.2 Differential Equations

Differential Equations (DE) modeling is a method that studies intraindividual variability as a form of continuous-time modeling (Murray, 2009; Brown, 2007; Deboeck et al., 2008), which can be implemented as fixed effects (Boker et al., 2004) or mixed-effects (Oravecz et al., 2011). Change over time can be modeled with different methods; common methods implemented in psychology are linear or quadratic trends and are usually analyzed with techniques such as latent growth curve modeling (LGC, McArdle & Nesselroade, 2014) and hierarchical linear modeling (HLM, Hox, 2010). With these methods the residual variability is represented as error and treated as independent observations from an assumed distribution (it is expected that time series data will be dependent of each other) because these methods model trajectories of observed scores (Deboeck et al., 2008). Methods that focus on intraindividual variability may allow research to answer questions that can not otherwise be answered (Deboeck et al., 2008; Bertenthal, 2007).

A DE is any equation where the variable is a derivative—a derivative represents the change in a variable with respect to other variables. DEs model the dynamic mechanism in intensively

measured variables. The implementation within the framework of analyzing time series and modeling intraindividual variability, the derivatives will represent change in a variable $x$ with respect to change in time (Deboeck et al., 2008). With DEs researchers can model relationships between a subject's current state and his or her changes over time in a univariate model, including multiple variables DEs can model the change in a variable with respect to the change in other variable.

DE models presented make two assumptions: continuity and differentiability (Boker & Nesselroade, 2002). The process should show continuous change without abrupt change from one value to another. The derivatives should be continuous; that is, the process should change smoothly. Researchers require a sufficient quality of repeated measures within their data, otherwise the estimated derivatives are unreliable (Hu (2013)).

A simpler function of a DE is to describe a continuous-time model with known parameter values. This is also called an initial-value problem, wherein researchers can describe a system just by knowing its initial value. Let us remember that the solution of a DE is a function that describes a generation process. A DE under the initial-value problem has infinite solutions until the initial-value has been stated, the solution is dependent of the giving initial-value (Brown, 2007). A general form of a DE is (Tornoe et al., 2004)

$$\dot{x} = g(x, t, p) \tag{1.1}$$

where $x$ is an $N$-dimensional dependent variable, $g$ is the structural model, $t$ (time) is the independent variable, and $p$ is an $M$-dimensional vector of parameters. Equation 1.1's system would depend on the functional form of $g$. With any form of $g$, Equation 1.1 states that the change in $x$ ($\dot{x}$, first derivative) will depend on $x$ itself, some $p$ parameters, and $t$ (although the dependency of time is not necessary for every DE). This is an example of a first order DE—the order of a DE is define by the order of the highest derivate included in the equation (Brown, 2007). An example of a second order DE would be

$$\ddot{x} = g(x, t, p) \tag{1.2}$$

where $\ddot{x}$ represents the second derivative, which is the acceleration of the DE process.

DE can be estimated with mixed-effects, accounting for interindividual variability in the intraindividual DE models. This combination is desirable to look into longitudinal change across multiple subjects or groups.

## 1.2    Mixed-Effects Differential Equations

Mixed-effects models describe relationships between a dependent variable and covariates—where the data are not independent. Common examples of nonindependant data include longitudinal data, multilevel data, and block design (Pinheiro & Bates, 2000). In this paper, I focus on the estimation of mixed effects in longitudinal data with DE.

Several statistical tests (ANOVA, MANOVA, linear regression) assume that observations are independent; data structures such as longitudinal data violate this assumption. Statistical tests that do not account for the dependent nature of the data affect the standard errors (Hox, 2010), and having incorrect standard errors may lead to incorrect inferences. Mixed-effects models correctly estimate models with these data structures.

Mixed-effects models are appropriate to estimate non-independent data, including the estimation of random effects. This framework is capable to include covariates measured at various levels of a hierarchical structure, and estimate their relation with the outcome variable. Mixed-effects models corrects for bias in the parameter estimates and standard error with non-independent data, ignoring this bias the parameter estimates and standard errors. By estimating variances and covariances for the random effects, it is possible to decompose the total variance of the outcome variable into portions associated with different levels of the data structure (Guo & Zhao, 2000).

The mixed-effects models offer the advantage of fitting to individual-state functions for each subject. This allows researchers to take into account for between-subject variability. Researchers expect to find between-subject variability in biological processes (Tornoe et al., 2004; Wang et al., 2008). At the same time, this allows researchers to see the system behavior for each subject and not

be limited by the overall sample system behavior (Maxwell & Boker, 2007). Prague et al. (2013) identified the benefit of using mixed-effects models by gaining information from the between-subject variability, especially when there are few time points of measurement for each participant.

Mixed-effects DE models are commonly apply in fields of biological research such as PK/PD (Tornoe et al., 2004; Wang et al., 2008; Jaeger & Lambert, 2013) and HIV dynamics (Huang et al., 2010; Wang et al., 2014). In the field of PK/PD, researchers seek to construct models to find the concentration of a substance (usually a drug) in blood. When researchers draw blood from subjects, after the substance has been injected, they are interested in models that describe the absorption and elimination rate of that substance (Wang et al., 2008).

Using a different approach, a Bayesian mixed-effects P-spline method is intended to be less computational intensive as previous methods, while stating affine systems of differential equations including information from several subjects. This method is highly dependent on the number of knots and their respective locations (Jaeger & Lambert, 2013). If the function lacks the necessary amount of knots, the function will not capture all of the information from the data. Conversely, if there are too many knots, the function will capture noisy variations in the data. To avoid this problem, they suggested using a large number of knots and controlling the spline coefficient with a DE model base-penalty term (Jaeger & Lambert, 2013). The authors used simulations to compare their Bayesian penalized P-spline DE method with their equivalent frequentist counterpart. The results suggested that the Bayesian method is as efficient or more efficient that other methods. The authors presented an extension to a mixed-effects model specifying interindividual variability for the DE parameters.

There is an initiative to develop programs to estimate mixed-effects DE models (Tornoe et al. (2004); Prague et al., 2013). A group has developed a package (nlmeODE; Tornoe et al., 2004) for the R platform (R Core Team, 2015). Their program merges the nlme (Pinheiro et al., 2015) and deSolve (Soetaert et al., 2010) R packages. This package demonstrates flexibility and accuracy in estimating nonlinear mixed-effects DE models; however, the nlmeODE package has not been updated since 2012.

An independent program NIMROD (Prague et al., 2013), follows a method that approximated the posterior distribution. They assumed a normal approximation of the posterior to compute the maximum of the posterior (MAP). Then they calculated the MAP by computing the likelihood function by integrating over the random effects via the adaptive Gaussian quadrature. This method returns accurate results (even within complex models) without an analytical solution.

Mixed-effects DE models properly estimate continuous-time dynamics while taking into account interindividual variability. Combining nomothetic and idiographic information allows researchers to intensively examine the dynamics of a single individual and to generalize the dynamics of a population of individuals (Maxwell & Boker, 2007).

## 1.3   Estimation Methods

Several papers (Wang et al., 2014; Wang et al., 2008; Huang et al., 2010; Prague et al., 2013; Jaeger & Lambert, 2013) mentioned the difficulty of estimating DE models because they may not have an analytical solution (Brown, 2007). Without an analytical solution, the most common method to solve a DE is with a numerical approximation (Brown, 2007; Murray, 2009). For these reasons, there are papers that present different methods to estimate DE models—some focus on semiparametric methods (Wang et al., 2008; Wang et al., 2014) and others focus on Bayesian methods of estimation (Prague et al., 2013; Huang et al., 2010; Jaeger & Lambert, 2013).

The Bayesian approach is preferred from a theoretical perspective, it allows prior and informative distributions to be incorporated and from a statistical-estimation point of view, a Bayesian approach is preferable because of the difficulties researchers often encounter from a classical approach in which models involve large numbers of parameters and complex, nonlinearity of the subject-specific models (Huang et al., 2010; Jaeger & Lambert, 2013). The Bayesian framework shows great power and flexibility for fitting DE models and is proving to be a popular method to estimate them (Andersen & Hojbjerre, 2005; Stramer & Roberts, 2007; Gander & Stephens, 2007; Duan et al., 2009; Donnet et al., 2010; Oravecz et al., 2011; Parslow et al., 2013; Murray,

2009; Tomassini et al., 2009; Israel & Migon, 2012; Higham, 2011; Huang et al., 2010; Jaeger & Lambert, 2013).

A method consists of generating an *m* number of phantom variables between the discrete-measure time points; this is an Euler–Maruyama numerical approximation (Stramer & Roberts, 2007; Murray, 2009; Donnet et al., 2010; Higham, 2011). In the continuous-time models, researchers assume that there are an infinite number of time points. From this infinite number of time points, researchers are only able to sample a finite or discrete number. Numerical approximation methods, such as Euler–Maruyama, intend to estimate the infinite number of time points between measurements (Murray, 2009; Stramer & Roberts, 2007; Brown, 2007; Higham, 2011).

By approximating a finite *m* number of phantom time points between measures with imputed time points, the model made use of data augmentation in the Bayesian framework to define *m* imputed time points. A minimum of $m \geq 40$ is proposed as enough to approximate proper results. This method of numerical approximation allows researchers to approximate the analytical solution—as *m* approaches $\infty$, the numerical approximation will be equal to the analytical solution (Stramer & Roberts, 2007; Murray, 2009). Euler–Maruyama numerical approximation is useful for models that do not have a tractable form (Donnet et al., 2010). This implementation can be applied for different mixed-effects growth curves. The standard growth functions describe monotone increasing growth but fail to model unexpected changes in growth rates. DE applications of growth functions allow us to model these variations in growth rate.

The Bayesian framework is an attractive tool to integrate mechanistic processes and observations. It provides a consistent, formal, and probabilistic framework that combines uncertainty in model parameters, model processes, and observations (Parslow et al., 2013; Higham, 2011). The main difference between Bayesian and frequentist statistical inference concerns the nature of the unknown parameters. The frequentist framework assumes that unknown parameters have a fixed value that we wish to estimate. Within the Bayesian framework, the unknown parameters are considered unknown and should be viewed as random variables with a probability distribution that reflects the uncertainty about their true value (Kaplan, 2014).

Figure 1.1: Idealized plot of temperature over time for two simple thermostats (Boker & Nessel-roade, 2002)
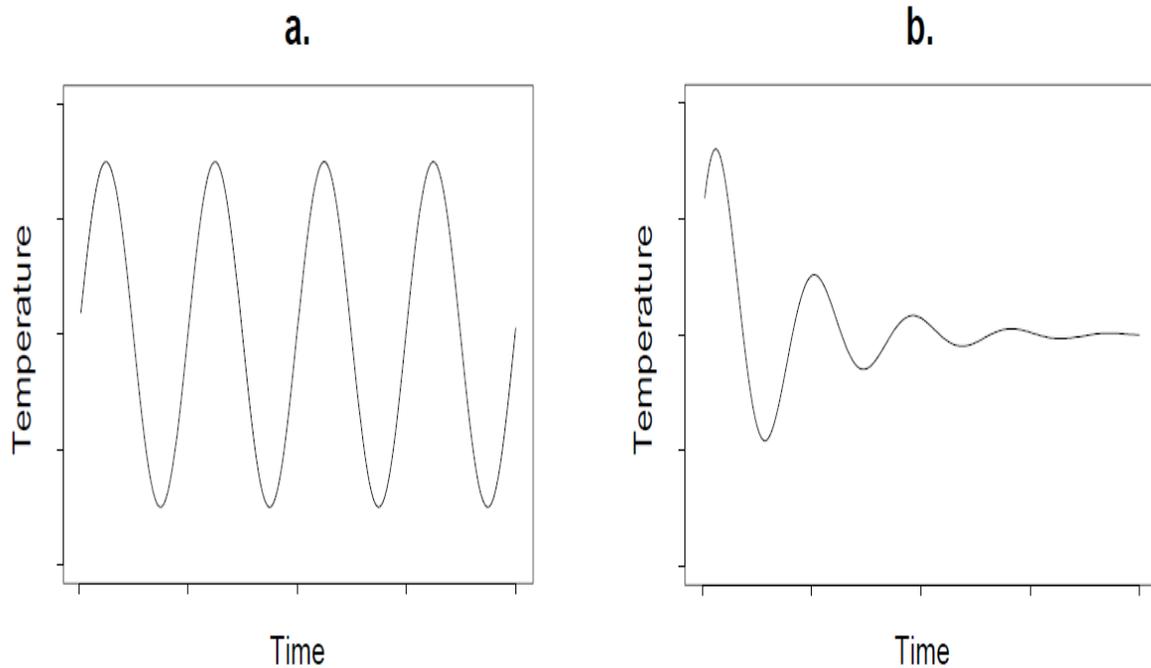


*Note.* (a) A thermostat that only responds to difference from the desire equilibrium temperature. (b) A thermostat that responds to both difference from equilibrium and rate of change in temperature.

### 1.3.1  Differential Equations Modeling in Psychology

There is a way to estimate and model DEs, with a clear intention to make these kinds of models available to psychologists (Boker, 2007). Psychologists face questions similar to the frequent question asked by physicists: How does change in one variable relate to change in another variable?

This has been exemplified as a second order DE model (Boker, 2007), that explains processes that follows an oscillation with damping. This can be exemplify by the oscillation of temperature like the ones presented in Figure 1.1; where (a) presents an oscillation without damping, like a thermostat that only adjusts to differences from the desire equilibrium temperature, while (b) presents an oscillation with damping, like a thermostat that adjusts to both differences from the equilibrium and the rate of change, which slows down the oscillation until the oscillation stops and keep constant at the equilibrium. This process is describe by Equation 1.3.

$$\ddot{x} = \eta x + \zeta \dot{x} + e \tag{1.3}$$

where the first ($\dot{x}$) and second ($\ddot{x}$) derivatives of $x$—with respect to time at a particular time, $x$—are the displacement (zeroth derivative) of the variable $x$ from its equilibrium value at a particular time, and $e$ is normally distributed independent residual. The first derivative represents the velocity of change (i.e., the amount of change per unit of time). The second derivative represents the acceleration of change (i.e., the rate at which the variable changes per unit of time; Hu et al., 2014).

Parameter $\eta$ is related to the frequency where $\eta < 0$ and $\eta + \zeta^2/4 < 0$. The period of the oscillation will be $\lambda = 2\pi/\sqrt{-(\eta + \zeta^2/4)}$, where $\lambda$ represents the number of observations that complete an oscillation. An approximation of the frequency, when $\zeta$ is equal to zero, will be $\omega = \frac{1}{2\pi}\sqrt{-\eta}$. $\zeta$ is the damping of an oscillating system, the sign of $\zeta$ indicates the damping behavior of the system, negative values indicate a tendency of oscillation to diminish to a baseline (damping), and positive values represent a tendency to amplify the oscillation (Steele & Ferrer, 2011; Hu et al., 2014). The model described by Equation 1.3 is a damped linear oscillator (DLO). It is a simple model to account for self-regulating systems that have a stationary equilibrium point. I will focus on the implementation of the DLO, but the same methods can be applied to other differential equations, namely, other relations between derivatives.

The DLO model presents the derivatives of a time series in a familiar way, and there are two common ways to run this model. The first one is in a two-stage process, where the derivatives are estimated and then included in a model as variables (Boker, 2007; Butner et al., 2005). The second method simultaneously approximates empirical derivatives of a time series and estimates a structural model to the covariance matrix between the derivatives (Boker et al., 2004; Boker, 2007; Steele & Ferrer, 2011; Hu et al., 2014).

### 1.3.1.1 Two-stage method

There are different ways to estimate the derivatives of a time series. One is the Local Linear Approximation (LLA), which adequately maintains the properties of the oscillator (Boker & Nesselroade, 2002). Any linear or nonlinear trend must be removed from the time series before the application of LLA, where LLA is used as the method to estimate derivatives from the (detrended) time series. DLO requires the time series without a trend because it estimates the oscillation around a stationary equilibrium point. The displacement ($x(t)$) can be estimated by a moving average or the residuals from a linear regression. The first and second derivatives are calculated as follows,

$$\dot{x}_{(t_i+\tau)} = \frac{x_{(t_i+2\tau)} - x_{(t_i)}}{2\tau} \tag{1.4}$$

$$\ddot{x}_{(t_i+\tau)} = \frac{x_{(t_i+2\tau)} - 2x_{(t_i+\tau)} + x_{(t_i)}}{\tau^2} \tag{1.5}$$

where $\tau$ is the lag between measured time points. After calculating the derivatives, researchers use them as variables in any regression program. The two-stage method can implement a mixed-effects coupled DLO model—a coupled DLO occurs when there are two or more cyclical dynamics that interact with each other. This model can showed great flexibility (Butner et al., 2005); by including nonlinear dynamics (in the parameters) in the model. They also included two escapement parameters, which allow energy to enter the system and affect the damping. Two parameters are also included that affect the frequency of the cyclical process,

$$\ddot{x} = \eta x + \zeta \dot{x} - \rho [\dot{x}]^3 - v [\dot{x}] x^2 - \delta [x]^3 - \mu [\dot{x}]^2 x + e \tag{1.6}$$

where $\rho$ and $v$ represent the escapement terms (Rayleigh, and van der Pol, respectively). $\delta$ and $\mu$ are the Duffing and the $\pi$-mix odd-series terms that influence the frequency. This shows the flexibility of working directly on the estimated derivatives. Now researchers can directly model relationships between those that have a different effect on the cyclical process. Researchers can also have the derivatives from different variables interact with each other.

This model can be simplified by estimating nonlinear terms ($\rho$, $v$, $\delta$, $\mu$) as fixed effects due to a lack of variability in the sample. These kinds of models are very useful in psychology because they describe change similar to how psychologists describe development, affect, and other individual and social phenomena (Butner et al., 2005).

### 1.3.1.2  Latent Differential Equations

LDE is the other method to estimate a DE model such as the DLO (Boker et al., 2004; Steele & Ferrer, 2011; Hu et al., 2014; Boker, 2007). A LDE approximates the derivatives as latent constructs in the framework of an SEM, and it follows the same structure as a latent growth model. The main characteristic is that all the factor loadings are fixed with values as

$$L = \begin{bmatrix} 1 & -2\Delta\tau & (-2\Delta\tau)^2/2 \\ 1 & -1\Delta\tau & (-1\Delta\tau)^2/2 \\ 1 & 0 & 0 \\ 1 & 1\Delta\tau & (1\Delta\tau)^2/2 \\ 1 & 2\Delta\tau & (2\Delta\tau)^2/2 \end{bmatrix} \tag{1.7}$$

where $\Delta\tau$ is the elapsed time between adjacent lagged columns in the time-delay embedded matrix. The first column represents the factor loadings that define $x$, the second column defines $\dot{x}$, and the third column defines $\ddot{x}$. The $x$ is defined by a factor loading equal to 1. $\dot{x}$ is defined by values fixed to the unit basis for a slope with an intercept at the middle indicator. The values increase by the elapsed time between the successive column of the time-delay embedded matrix. $\ddot{x}$ is define by factor loadings as the indefinite integral of the second column. Higher order derivatives can be calculated the same way (Boker, 2007; Steele & Ferrer, 2011).

This leads to another difference in how researchers recommend implementing an LDE model—the use of a time-delay embedded matrix (von Oertzen & Boker, 2010; Boker, 2007). The use of time-delay embedded matrix increases the precision of intraindividual variability (von Oertzen & Boker, 2010; Steele & Ferrer, 2011). If the interest is on modeling intraindividual variability, it is

better to collect longer time series rather than larger samples (von Oertzen & Boker, 2010).

This is a method to construct a set of short time-ordered lapses from a longer time series. If the process that generated the long time series is stationary, every short-term lapse is considered a representative sample (von Oertzen & Boker, 2010; Boker, 2007). This method constructs a matrix for each subject time series. The embedding dimension represents the length of the short-term time lapse, and the loading matrix, presented in Equation 1.7, is an example of an embedding dimension $d = 5$. For example, $\tau = 1$ indicates that researchers select every observation, and $\tau = 2$ indicates that they take every other observation. For example, let us consider a time series with 10 observations ($ts = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$); an embedded matrix with $d = 5$ and $\tau = 1$ results in

$$L = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 8 & 9 \\ 6 & 7 & 8 & 9 & 10 \end{bmatrix} \tag{1.8}$$

and an embedded matrix with $d = 5$ and $\tau = 2$ results in

$$L = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 \\ 2 & 4 & 6 & 8 & 10 \end{bmatrix} \tag{1.9}$$

As shown in the previous examples, the selection of embedding dimensions will affect the structure. For an LDE model, $d = 4$ is the minimum number to identify a second-order differential equation model. There are different recommendations on how to select the embedding dimension ($d$); a general recommendation is to use $d = 5$ because it has proved to be more stable (Boker, 2007). The use of higher embedding dimensions must be implemented carefully for models that include an oscillation because the time from the first to the last column in the embedded matrix should not exceed one half the elapsed time between peaks in the oscillation (Boker, 2007). The

selection of the parameters $\tau$ and $d$ should be based on characteristics of the data, such as sampling frequency and the precision of the measurement (Steele & Ferrer, 2011). The embedding dimension is both data and model dependent (Boker et al., 2016).

With a different recommendation, a suggestion is to plot the estimated frequency ($\eta$) as a function of $d$; the optimal $d$ will occur just after $\eta$ becomes stable. This requires researchers to run the LDE model with several values of $d$ to find the optimal value for $d$ (Hu et al., 2014). It is generally recommended to center each time series around each subject's equilibrium (if known). If the equilibrium is unknown, use the individual's mean as the equilibrium. Centering each time series is important because the interpretation of the derivatives is a function of their movement around the equilibrium (Hu et al., 2014). Not centering the data could bias the results (especially when there are large individual differences in equilibrium). This is related to the two assumptions stated about the DLO (Boker & Nesselroade, 2002): The oscillator is a homogeneous process in a population, and the process is stationary over time. The homogeneous assumption means that researchers assume the process is representative and can be generalize. The stationary assumption states that the process does not change over time.

The use of a time-delay embedded matrix violates the assumption that each row is independent, which may result in $\alpha$ inflation. This happens because the same data gets repeated in different rows, and SEM models assume that the rows are independent (Steele & Ferrer, 2011; Hu et al., 2014). Counterintuitively, it was demonstrated by mathematical derivation and simulation that instead of alpha inflation, using overlapping samples in this way produce a smaller actual standard error in parameter estimation (von Oertzen & Boker, 2010). By violating this assumption, researchers bias the standard errors (Hu et al., 2014).

Factor-loading matrices, as seen in Equation 1.7, work when researchers built a time-delay embedded matrix from a single variable. One of the greatest advantages of the SEM framework is the use of multiple variables that measure the same construct. By measuring the same construct with multiple indicators, the method is able to correct for measurement error and can estimate the construct more accurately (Little, 2013). With multiple variables, the construct is defined as the

shared variance between the multiple variables, and the specific variance is excluded. This way the construct does not include measurement error or the specific variance for each variable (Little, 2013).

The implementation of an LDE with multiple variables is not as straightforward because it needs to built the time-delay embedded matrix for each variable. The multivariate LDE is estimated with a time-delay embedded matrix for each variable. With multiple time-delay matrices, the factor loadings that define the latent derivatives need to be adjusted (Boker, 2007; von Oertzen & Boker, 2010; Tiberio, 2008).

The multivariate LDE estimates both within-time factor structure of the latent factor $F$ as well as the time-delayed structure as derivatives. Now the loading matrix will not be entirely constrained to fixed values (Boker, 2007; von Oertzen & Boker, 2010). Consider a case where factor $F$ is measured with three variables; the factor loadings for the first variable are the same as in Equation 1.7 and for the second variable. The factor loadings are scaled by a free coefficient ($a$ and $b$, respectively). The resulting factor-loading matrix is

$$
L = \begin{bmatrix}
1 & -2\Delta\tau & (-2\Delta\tau)^2/2 \\
1 & -1\Delta\tau & (-1\Delta\tau)^2/2 \\
1 & 0 & 0 \\
1 & 1\Delta\tau & (1\Delta\tau)^2/2 \\
1 & 2\Delta\tau & (2\Delta\tau)^2/2 \\
a & -2a\Delta\tau & a(-2\Delta\tau)^2/2 \\
a & -1a\Delta\tau & a(-1\Delta\tau)^2/2 \\
a & 0 & 0 \\
a & 1a\Delta\tau & a(1\Delta\tau)^2/2 \\
a & 2a\Delta\tau & a(2\Delta\tau)^2/2 \\
b & -2b\Delta\tau & b(-2\Delta\tau)^2/2 \\
b & -1b\Delta\tau & b(-1\Delta\tau)^2/2 \\
b & 0 & 0 \\
b & 1b\Delta\tau & b(1\Delta\tau)^2/2 \\
b & 2b\Delta\tau & b(2\Delta\tau)^2/2
\end{bmatrix}
\tag{1.10}
$$

The factor-loading matrix $L$ in Equation 1.10 assumes that the factor structure is invariant across time and across the derivatives of factor $F$ (Boker, 2007; von Oertzen & Boker, 2010). This is the assumption of differential factor invariance; this assumption is expected to be tenable when the differential equation model is linear in its coefficients and variables. This is an assumption that can be tested by allowing $a$ and $b$ to be different across columns of $L$ in Equation 1.10 (Boker, 2007).

**Substantive applications of LDE** . LDEs have been apply to identify the self-regulatory and coregulatory dynamics of affect between romantic couples (Steele & Ferrer, 2011). The authors started with univariate LDE damping linear oscillator models, and they found that there is almost no damping effect, so they fixed $\zeta$ to zero. They intended to run a mixed-effects model (with random $\eta$ and $\zeta$), but it proved impossible within the SEM framework. Instead, they proceeded to use a two-stage approach to estimate the mixed-effects model. By exporting the factor scores

19

from the LDE model, they found that only $\eta$ had evidence of variability across subjects. This exemplifies the strengths and weakness of each approach.

The coupled LDE model shows the power and flexibility of estimating both systems simultaneously and possibly adds a relationship between them (Steele & Ferrer, 2011). This framework allows us to compare the effects from the derivatives from one system (positive affect) to another (negative affect) and test several models, such whether or not the effect from the $\eta$ of one system is equal to the $\ddot{x}$ of the other system. This shows an example of oscillating processes in psychology and how the LDE method helps to model complex relationships between systems (Steele & Ferrer, 2011).

This has been applied to related ovarian hormone cycles to emotional eating. The authors also made an interesting methodological addition to the model; they showed how to specify a moderating LDE with the following equation (Hu et al., 2014),

$$\ddot{x}_{ij} = (\eta_0 + \eta_1 z_i)x_{ij} + (\zeta_0 + \zeta_1 z_i)\dot{x}_{ij} + e_{ij} \tag{1.11}$$

where $i$ represents each subject, $j$ represents each time point, and $z_i$ is the moderator. Note that the moderator does not vary across time. For example, for subject $i$ with moderator $z_i$ its individual frequency is $\eta_0 + \eta_1 z_i$, and its individual damping parameter is $\zeta_0 + \zeta_1 z_i$.

Emotional eating and negative affect are independent systems; they don't have significant coupling parameters. Also, it was identified a coupling relationship between emotional eating and ovarian hormone cycles. These two systems are coupled by their first derivatives. There was no coupling relationships between emotional eating and negative affect. Instead, they found that the average negative affect is a moderator of the LDE system of emotional eating. Higher average levels of negative affect are associated with faster frequency of emotional eating oscillation. Average negative affect also moderates the coupling relationships between emotional eating and ovarian hormone cycles, concluding that a person with high negative affect has a stronger coupled relationship between emotional eating and hormone cycles, and that a person with low negative affect has an emotional eating system less affected by the hormone cycles (Hu et al., 2014).

A coupled LDE provides information about the within-individual self-regulation and the strength of coupling between systems. Coupled systems of differential equations have great potential as a method to understand the complex ways in which humans interact and mutually regulate social and family systems (Boker, 2007).

### 1.3.1.3   Nonstationary and individual parameters in LDE

For a process to be stationary it means that the estimated parameters may not change over time (Boker et al., 2004; Boker et al., 2016). Nonstationarity can be present in any of the parameters of the model. In Equation 1.3, the model assumes that both $\eta$ and $\zeta$ do not change. Nonstationarity could show as a change in $\zeta$, describing a process that would damp to its equilibrium and later start cycles again. Equation 1.3 assumes a stable equilibrium state—nonstationarity would show if the equilibrium increased or decreased over time (slope). Boker et al. (2016)'s method extends the DLO (Equation 1.3) to handle nonstationarity in the equilibrium of the process by estimating a linear slope of the changing equilibrium.

Boker et al. (2016) already started working on some of these needed developments of the LDE. Their first extension to the LDE model allows for individual differences in equilibrium. The recommended and usual practice has been to center each time series around its equilibrium (mean), so that the model can be interpreted as a function of the whole sample dynamics around their equilibrium (Hu et al., 2014; Boker & Nesselroade, 2002; Boker, 2007). The extended method allows to estimate the equilibrium for each subject (Boker et al., 2016). This is accomplished by estimating a latent intercept in the same model. The derivatives that are approximated with the $L$ matrix (Equation 1.7) model the intraindividual variability, while the new latent intercept models the means of the time series. This latent intercept estimates individual equilibrium for each subject, it does not estimate the average sample equilibrium and the variance of the equilibrium in the sample.

Another extension allows to account for the change of equilibrium across time (nonstationarity), that is, how to estimate equilibrium trends. Regular recommendations and common practice

to have unbiased LDE parameters was to detrend each time series (Hu et al., 2014; Boker & Nesselroade, 2002; Boker, 2007). The quilibrium trends are treated as the individual equilibrium points are treated, which allowed them to estimate individual linear (or other polynomial) trends like estimating a latent slope. Just like the individual equilibrium, this method does not estimate the sample average and sample variance of the equilibrium trends (Boker et al., 2016). To express these models in equations, I will first present the LDE in a matrix equation format,

$$X = GL + U \qquad (1.12)$$

where $X$ is the embedded matrix, $G$ is a matrix of unobserved latent derivative scores, $L$ is the factor loading matrix, as in Equation 1.7, and $U$ is a matrix of unobserved unique scores. When individual estimates for the equilibrium are included the model is written as

$$X_i^D = M_i K + GL + U \qquad (1.13)$$

the new matrix $M_i$ is the $i$th row of an $N$ x 1 matrix of means for the $N$ persons in the sample, $K$ is a 1 x $D$ matrix with fixed values of 1 in each cell (as a regular intercept in a latent growth model). Equation 1.13 estimates the individual intercept, requiring us to take into account the nested structure of the embedded matrix. This extended method is implemented using OpenMx (Neale et al., 2016).

This model follows the same matrix form as Equation 1.13. Now $M_i$ is the $i$th row of an $N$ x 2 matrix of means for the $N$ persons in the sample containing the individual equilibrium and equilibrium slope. The matrix $K$ is a 2 x $D$ matrix with fixed values. Because the delay embedded matrix $X_i^D$ has many rows for each subject, the matrix $K$ will have fix values for each row. The values must be calculated for each occasion $j$ within each individuals $i$'s block f data in $X_i^D$ (Boker et al., 2016).

$K$ will now be calculated as

$$K = JH + C \tag{1.14}$$

with $\Delta\tau$ being the interval of time between successive measurements in a delay-embedding matrix of $D = 5$, $C$ and $H$ will be

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2\Delta\tau & -\Delta\tau & 0 & \Delta\tau & -2\Delta\tau \end{bmatrix} \tag{1.15}$$

$$H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \Delta\tau & \Delta\tau & \Delta\tau & \Delta\tau & \Delta\tau \end{bmatrix} \tag{1.16}$$

For the $j$th occasion of measurement within individual $i$'s data, the matrix $J$ must be created in the following form

$$J = \begin{bmatrix} 0 & 0 \\ 0 & j \end{bmatrix} \tag{1.17}$$

These row-specific calculations are needed to estimate the model properly given the delay-embedded data. So far, this can only be performed using OpenMx (Neale et al., 2016) because no other standard SEM software allows for these row-specific formulations. This extension of the LDE becomes a hybrid between LDE and LGC (Boker et al., 2016).

One last extension to the LDE model is to estimate individual differences in dynamics. This extension allows researchers to relax the assumption of a homogeneous process. As with the individual equilibrium and equilibrium trend, these differences in the dynamics are model by estimating individual $\eta$ and $\zeta$, but it does not estimate the average parameters and the variance of the sample of these parameters (Boker et al., 2016). This model is written as

$$\ddot{x}_{ij} = \eta_i x_{ij} + \zeta_i \dot{x}_{ij} + e_{ij} \tag{1.18}$$

This is the DLO equation (Equation 1.3), where the parameters $\eta$ and $\zeta$ are estimated for each

subject $i$ at each $j$ row ($\eta_i$, $\zeta_i$) for the time-delay embedded matrix of the time series of variable $x$. In matrix formulation it is specified as

$$X_{ij}^D = M_i K + G_i L + U \qquad (1.19)$$

where $X_{ij}^D$ is the time delay embedded matrix of embedding dimension $d = D$ for each subject $i$ at each row $j$. $M_i$ represents the latent intercept and slope for each subject $i$, and $K$ is the matrix of fixed values from Equation 1.14. $G_i$ represents the latent derivatives for subject $i$ from Equation 1.19, with the matrix of fixed values $L$ as seen in Equation 1.7.

This extended LDE model has certain advantages such as subject parameter estimates for $\eta$, $\zeta$, equilibrium intercept, and slope. Adding the equilibrium slope is estimating a nonstationary process by allowing it to change over time. It still lacks the addition of sample mean and variance of the subject parameters (Boker et al., 2016).

## 1.4 Bayesian Inference

The Bayesian framework is powerful and flexible, capable of improving the estimation and inference of DE model. Psychological research has been mainly approached from the frequentist framework, but the Bayesian framework is becoming more popular (Hong et al., 2013; Wagenmakers et al., 2008). With respect to DEs applied in psychological research, there is little existing literature. But there is the work by Oravecz et al. (2011) who estimated a mixed-effects stochastic DE model. Bayesian estimation is more appropriate for several reasons. A frequentist estimation of the model would be difficult because such a model would require high-dimensional integration over numerous random effect distributions. Furthermore, the method that would allow the researchers to estimate the model become computationally prohibitive (Oravecz et al., 2011). In the Bayesian paradigm, the researchers could avoid the explicit integration over the random effects because the inference is based on the full joint posterior distribution of the parameters. Also, the frequentist method commonly assume that all parameters follow a normal distribution. Bayesian

estimation allowed them to relax this assumption. In this section I will expand on the advantages of Bayesian inference (Oravecz et al., 2011).

Bayesian statistical inference about parameters $\theta$ are made within the function of probability statements; these are conditional on the observed values of $y$. This conditioning on observed $y$ is the most basic difference between Bayesian statistics and the frequentist approach, which is based on a retrospective evaluation of the procedure to estimate $\theta$ over the distribution of possible values of $y$, conditional on the true unknown value of $\theta$ (Gelman et al., 2013). Even when both frameworks may lead to similar conclusions, in many simple analyses, analysis from the Bayesian approach can be extended to more complex problems (Gelman et al., 2013).

### 1.4.1  Baye's rule

The probability statements about $\theta$ given $y$ is a model defining the joint probability distribution for $\theta$ and $y$. This joint probability density is expressed as the product of two densities, known as the prior distribution ($p(\theta)$) and the sampling distribution ($p(y|\theta)$), written as (Gelman et al., 2013; Gelman & Hill, 2007; Wagenmakers et al., 2008)

$$p(\theta, y) = p(\theta)p(y|\theta) \tag{1.20}$$

The posterior distribution is estimated by conditioning on the observed data $y$, implementing the property of conditional probability known as Bayes's rule (Gelman et al., 2013):

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \tag{1.21}$$

where $\theta$ represents the parameters and $y$ is the data. The likelihood is $p(y|\theta)$ and prior is $p(\theta)$ (Gelman et al., 2013; Jackman, 2009). The behavior of $\theta$ under the given data $y$ is fully described by the conditional distribution of $\theta$ given $y$, this conditional distribution is called the posterior distribution of $\theta$ (Gelman et al., 2013; Song & Lee, 2012). The posterior distribution is only affected by the data by $p(y|\theta)$, usually implemented as a function $\theta$ for fixed $y$ (Gelman et al.,

2013). The priors represent previous information about the model, since no study is conducted in the complete absence of previous research there is always some information that can be included in the model. The information is usually incorporated in the choices of designs, variables chosen, or conceptual diagrams. Bayesian statistics requires to make explicit the previous knowledge, but the moderates this knowledge by the likelihood (Kaplan, 2014).

The Baye's rule follows the likelihood principle, stating that for a given sample of data any two probability models, $p(y|\theta)$, that have the same likelihood function will lead to the same inference of $\theta$ (Gelman et al., 2013). Further, the likelihood principle is reasonable within the framework of the family of models adopted for a particular analysis. From the Bayesian framework, inferences are summarized by random draws from the posterior distribution $p(\theta|y)$ (Gelman & Hill, 2007; Gelman et al., 2013).

## 1.4.2   Disadvantages of the frequentist framework

It is important to understand the limitations of the frequentist framework. The main limitation is that it does not condition on the data (Wagenmakers et al., 2008), as clearly expressed by the probability estimated $p(y|\theta)$. In the frequentist paradigm, the accuracy of the evidence is identified with the long-run behavior of the procedure and the $p$-value is defined in function of repeating the procedure (Kruschke, 2010). Because the method is not conditional on the data and dependent on the long-run behavior, a procedure that is expected to work properly could misbehave for a particular data $y$ (Wagenmakers et al., 2008).

The $p$-value is the probability under the null hypothesis of observing data at least as extreme as the observed data, meaning that the $p$-value is in part determined by data that were never observed (Wagenmakers et al., 2008). The use of the $p$-value means that a hypothesis that may be true could be rejected because it has not predicted observable results that have not occurred. This is also related to a common problem: The frequentist inference is easily (and commonly) misinterpreted (Matthews, 2001). With the $p$-value, it is common to fall into the transposition of conditioning fallacy, where $p(A|B)$ is mistaken for $p(B|A)$. Researchers mistakenly interpreted the $p$-value as

an answer to $p(H_0|y)$, while their $p$-value actually answered $p(\geq y|H_0)$ (Matthews, 2001).

In frequentist statistics the inference is dependent on the intention of the study, this is reflected by the adjustment of $\alpha$ in the presence of multiple tests (Wagenmakers et al., 2008). Using the same data, we could get different conclusions given the intention of doing multiple tests. In the case where the $p$-value = .03, if the nominal $\alpha = .05$ is used, we would reject the null hypothesis. In the case of doing multiple groups comparisons on the same variable, it would be recommended to adjust the $\alpha$, and by adjusting it to an $\alpha = .01$ we now fail to reject the null hypothesis. The use of $\alpha$ is another problem for the frequentist inference because choosing a certain cutoff for $\alpha$ is arbitrary (Matthews, 2001). The common use of the nominal $\alpha = .05$ is due to convenience and tradition rather than philosophical or probabilistic reasoning.

The frequentist inference overestimate significance (Matthews, 2001). Although the $p$-value gives us $p(\geq y|H_0)$, what we really want to estimate is $p(H_0|y)$. This can be estimated using the Baye's theorem. If we write it in function of the null hypothesis ($H_0$), we get

$$p(H_0|y) = (1 + \frac{1 - p(H_0)}{p(H_0)BF})^{-1} \tag{1.22}$$

where $p(H_0)$ is the prior probability that $H_0$ explains the data and $BF$ is the Bayes factor that represents the weight of evidence provided by the data in favor of $H_0$. The lower bound of the $BF$ is calculate by

$$BF \geq zexp[(1 - z^2)/2], (z > 1) \tag{1.23}$$

where $z$ is the normal test statistic. From the $z$ distribution an $\alpha = .05$ corresponds to $z = 1.96$. By calculating the $BF$ for the cutoff, we get $BF \geq 0.47$. For Equation 1.22 to argue that a $p$-value of .05 is equivalent to $p(H_0|y) = .05$, the researcher must implement a prior probability for the null hypothesis no higher than 0.1, meaning that he or she would need to previously believe with 90% certainty that the null hypothesis is wrong to be justified in using the $p = .05$ (Matthews, 2001). For this condition to hold, the researcher would need to justify his or her choice of prior belief

that there is a 90% chance that the null hypothesis is false. From a different perspective, let us say that a researcher is not confident enough to use such an informative prior for $H_0$. Instead, he or she decides to use $p(H_0) = .5$, thereby cautiously saying that there is 50% chance. Under this new (less informative) prior, a researcher states that $p(H_0|y) = .22$, which is far from the nominal $\alpha = .05$ (Matthews, 2001).

The problems and inconsistencies of the frequentist inference are not new. They are well known and documented (Wagenmakers et al., 2008), and the problems listed in this section are not exhaustive. Despite these clear criticisms, why is it that the frequentist approach is still predominant in the field of psychology?, the reason relates to a combination of several factors: They are simple to implement, most practitioners do not have an appealing alternative or lack statistical knowledge altogether, there is a systemic problem in statistics of faulty and one-sided teaching (even at the university level), and the strong historical precedent. With this comparison between the two approaches, I intent to demonstrate the practical and philosophical reasons to chose the Bayesian framework over the frequentist one (Wagenmakers et al., 2008).

### 1.4.3   Advantages of the Bayesian framework

Bayesian inferences treat the parameters $\theta$ as random parameters. By implementing Equation 1.21, it estimates the posterior distribution for $\theta$. The posterior distribution is conditional on observed data $y$, so data that could have been observed does not affect the Bayesian inference (Wagenmakers et al., 2008). Because the Bayesian framework makes inferences from the posterior distribution $p(\theta|y)$, it allows researchers to make direct inferences about $\theta$ given the data that they have observed. On the other hand, from the frequentist framework, the inference must be made in function of the null hypothesis instead of $\theta$ (Wagenmakers et al., 2008). Let us use the confidence interval as a comparison between these inferences: In the frequentist framework, a 95% confidence interval for $\mu$ of [-0.5, 1.0] leads to a counterintuitive interpretation that when the procedure is applied a large number times, the different intervals cover the true $\mu$ 95% of the time. However, if we were to have the same interval as a Bayesian credible interval [-0.5, 1.0], the interpretation will

be consistent with our intuition and consistent with what interests researchers. This interval tell us that there is a .95 probability that $\mu$ lies between [-0.5, 1.0].

The $p$-value can only lead us to either reject or fail to reject the null hypothesis. This is because the frequentist framework cannot collect evidence in favor of the null hypothesis. In the Bayesian framework, the null hypothesis does not have a special place in its inference (i.e., it is not conditional upon it), so the null hypothesis can be tested and can collect evidence as any other hypothesis (Wagenmakers et al., 2008). The Bayesian framework does not suffer from the same problems as the frequentist one. Like in the case of the $p$-value, there is no version of it, and for the same reason, there is no need to adjust the $\alpha$. Further, Bayesian inference can implement cumulative scientific progress with the inclusion of previous knowledge into the specification of the prior uncertainty (Kruschke, 2010).

The Bayesian framework is flexible enough to allow the implementation of complicated statistical techniques, like mixed-effects models that can estimate both subject and sample parameters (Wagenmakers et al., 2008). Most frequentist methods are constrained by many computational assumptions, but the Bayesian approach lacks these restrictive assumptions because it requires the posterior distribution to be defined, and this can be done in a flexible manner. Bayesian methods are implemented with descriptive models that can be easily customized to different situations without the computational restrictions of the frequentist framework (Kruschke, 2010; Hong et al., 2013).

The most common critique of the Bayesian framework is that it is subjective, and science must be objective. Since the ground of objectivity has been seized by the frequentists (Efron, 1986). On the other hand, there is no such thing as strict objectivity because reasoning under uncertainty is going to be relative to some background knowledge (Wagenmakers et al., 2008). From this point of view, strict objectivity is a goal that cannot be achieve. Every decision made by a statistician (frequentist or Bayesian) is guided by knowledge that will affect the inference, and as explained in the previous section, the intention of the analyst can affect the frequentist inference.

The critic of subjectivity in Bayesian inference is due to the inclusion of priors in the analysis.

The prior distribution is a model of uncertainty and represents the relative credibility of the parameter values in the absence of new data. There are two major answers to this critic (Kruschke, 2010): (a) Objective priors are implemented and are usually noninformative with the intention that they have as little effect on the result as possible, and (b) objective or subjective priors that are included can be tested for robustness and sensibility with respect to certain priors. Objectivity is a goal that cannot be achieve, and both frequentist and Bayesian inference have subjective elements; the difference is that the Bayesian approach is open to inspection (Wagenmakers et al., 2008; Kruschke, 2010).

Even though critics of Bayesian inference see the inclusion of priors as a disadvantage, they are a major strength of Bayesian inference. Indeed, the priors allow us to include prior knowledge in the analysis. The priors are also necessary for a rational inference. In order to make direct inferences in the function of $\theta$, we need the priors to estimate the posterior distribution as shown in Equation 1.21. This means that the inclusion of priors is a practical and philosophical advantage (Wagenmakers et al., 2008; Matthews, 2001; Kruschke, 2010).

### 1.4.4 Bayesian Structural Equation Modeling (BSEM)

In this section I present the characteristics of BSEM because the implemented method is the LDE model. As presented in a previous section, LDE approximate derivatives from the framework of SEM, similar to the specification of a LGC (Boker, 2007). SEMs are specified in terms of the population covariance matrix $\Sigma$ (Kaplan, 2014; Little, 2013; McArdle & Nesselroade, 2014). For a confirmatory factor analysis it would be

$$\Sigma = \Lambda \Psi \Lambda' + \Theta \tag{1.24}$$

where $\Sigma$ is the sample covariance matrix between the observed variabes, $\Lambda$ is the factor loading matrix, $\Psi$ is the covariance matrix between latent variables, and $\Theta$ is the covariance matrix of uniqueness. The matrix formulation (Equation 1.24) is an approximation the factor model equation

$$y_{ij} = \alpha_j + \lambda_j \eta_{ik} + \varepsilon_j \tag{1.25}$$

where $y_{ij}$ is the observed variable $j$ for subject $i$, $\alpha_j$ is the intercept for variable $j$, $\lambda_j$ is the factor loading for variable $j$, $\eta_{ij}$ is the score in the latent variable $k$ for subject $i$, and $\varepsilon_j$ is the uniqueness for variable $j$. In frequentist SEM, the model is estimated using Equation 1.24 instead of Equation 1.25 because in this method, $\eta$ is not identified; it is an unobserved latent variable. In BSEM, the models are estimated directly by applying Equation 1.25 to raw data, and $\eta$ is estimated simultaneously using other parameters (Merkle & Wang, 2014; Song & Lee, 2012).

BSEM identifies the model through data augmentation; this way $\eta$ is identify by setting priors which constrain the data space (Gelman et al., 2013; Song & Lee, 2012). A common way to set $\eta$ is to limit the distribution to $N \sim (0, 1)$. Doing this allows us to draw values of $\eta$ from a constrained data space (Song & Lee, 2012; Merkle & Wang, 2014). We can see that this is similar to a method of identification in frequentist SEM, which fixes the mean and variance of $\eta$ (Little, 2013; McArdle & Nesselroade, 2014). The difference is that BSEM actually estimates $\eta$. In the case of an LGC, $\lambda_j$ are fixed values defining the latent intercept and slope. The main difference between frequentist SEM and BSEM is that in BSEM we estimate $\eta_{ik} \sim MVN(\mu, \psi)$, where $\mu$ is a vector of latent means and $\Psi$ is the covariance matrix between the latent intercepts, meaning that in BSEM we estimate the individual values of $\eta_{ik}$ and the characteristic of their distribution ($MVN(\mu, \psi)$). In frequentist SEM, we only estimate the characteristics of the distribution (Song & Lee, 2012).

### 1.4.5 Bayesian estimation

Bayesian methods are estimated through Markov Chain Monte Carlo (MCMC) simulations, which is a general method based on drawing values of all parameters $\theta$ from approximate distributions and adjusting the draws to approximate the target posterior distribution (Gelman et al., 2013). Researchers draw from a Markov chain by sampling sequentially, where the distribution of the

sample draws depends on the last value drawn. MCMC is implemented when it is not possible to sample $\theta$ directly from $p(\theta|y)$. Instead, MCMC samples iteratively in a way that at each step of the process researchers expect to draw from a distribution that becomes closer to $p(\theta|y)$. With MCMC, researchers create a Markov process where the stationary distribution is specified as $p(\theta|y)$. Then the simulation must run long enough so that the distribution of the current draws is close enough to this stationary distribution (Gelman et al., 2013).

There are different MCMC methods used to sample from $p(\theta|y)$ such as the Gibbs sampler, the Metropolis-Hastings algorithm, and the Hamiltonian Monte Carlo (HMC) algorithm (Gelman et al., 2013; Song & Lee, 2012; Merkle & Wang, 2014; Hoffman & Gelman, 2014). Here, I will expound on the Gibbs sampler because it is the one that implemented in the analyses. The Gibbs sampler is know to reach convergence with proper parameter estimates in short time, the models implemented in this research does not fulfill the conditions to benefit from the characteristics of other samplers, these models would achieve the equivalent solutions independently of the sampler (Gelman et al., 2013; Merkle & Wang, 2014; Plummer, 2003).

The Gibbs sampler is defined in terms of subvectors of $\theta$, where the vector of parameters $\theta$ is divided into $d$ subvectors of $\theta$, $\theta = (\theta_1, ..., \theta_d)$. For each iteration the Gibbs sampler cycles through the subvectors of $\theta$, drawing each subset conditional on the value of the other subvectors. At each iteration $t$ the Gibbs sampler has $d$ steps. At iteration $t$, an ordering of the $d$ subvectors is chosen, where each $\theta_j^t$ is sample from the conditional distribution given all the other subvectors of $\theta$:

$$p(\theta_j|\theta_{-j}^{t-1}, y) \tag{1.26}$$

where $\theta_{-j}^{t-1}$ represents all the components of $\theta$ but $\theta_j$ at the current state. Every subvector of $\theta_j$ is updated conditional on the latest values of the other subvectors, which are the iteration $t$ values for the subvectors already updated and the iteration $t - 1$ values of the other subvectors (Gelman et al., 2013).

When estimating a model, the sampler starts a random points of the parameter space, define as

random numbers based on the priors of the model. Since the sampler starts with random number of $\theta$, it needs to sample over several iterations (usually several thousands) until it has reach stationarity of the parameters $\theta$. Once the model has achieve stationarity of $\theta$ we have to exclude the initial set iterations, this way we only make inferences in function of stationary iterations. In order to evaluate convergence the sampler is estimated on independent chains, convergence is evaluated by the variability between the independent chains, a model has found a convergent solution when the independent chains reached the same solution. This is determined using the potential scale reduction factor (PSRF), also know as univariate R-hat (see Gelman & Rubin, 1992; Gelman et al., 2013). It is established that a model converged when R-hat for every parameter $\theta$ is lower than 1.10 (Brooks & Gelman, 1998).

## 1.5   Contribution of this Dissertation

Neither the LDE nor the two-stage methods are perfect. The LDE method allows researchers to separate measurement error from the dynamics of the system, whereas the two-stage method assumes that the measures are without measurement error. This allows for an LDE to have lower bias in the parameters than the two-stage method (Boker, 2007). The two-stage method also has advantages over the LDE because it allows researchers to do mixed-effects models. This is shown by Butner et al. (2005), who easily added other nonlinear parameters. From the SEM framework, there are some restrictions on the relationships that can be implemented between latent constructs. At the same time, the SEM framework allows researchers to estimate some relationships that can not be estimated with the two-stage method. For example; a coupled DLO (Boker, 2007; Steele & Ferrer, 2011; Hu et al., 2014) allows researchers to estimate all the correlations and regression between derivatives of two systems in one model. With the two-stage model, researchers would have to estimate it in several separate models.

Given this comparison between the two-stage method and the LDE, the LDE has more advantages over the two-stage method. More important, the disadvantages of the LDE can be overcome

if researchers develop methods that can account for the mixed-effects for the LDE. There are limitations and needed development of the LDE for it to be applicable in a wider variety of settings. First, the implementation of mixed-effects LDE because individuals may have their own frequency ($\eta$) and damping ($\zeta$), and these may be related to other variables (observed or latent). Other needed development is for LDE to be able to account for nonstationary systems. Nonstationarity may be accounted for in different ways: It could be on the equilibrium, meaning that the equilibrium point may be changing over time. Another form of a nonstationary system would be if the parameters change over time because a process may be adapting itself over time (Boker, 2007).

This research take on the extensions made to the LDE model (Boker et al., 2016), which include: the estimation of subject parameters, and nonstationary equilibrium. This extended LDE model lacks the estimation of the sample parameters and their respective standard error, it requires to use the mean of the individual parameters as the best approximate of the sample parameter (without standard error). Currently implemented with OpenMX it is limited to an indirect inference of the parameters.

From the extended LDE (Boker et al., 2016) this research will develop a Bayesian LDE model that allows to estimate both subject and sample parameters, presenting a direct inference of the parameters, a model that can be also apply to multiple groups as shown in the apply example. The Bayesian implementation of this model presents the advantages of both the two-stage and LDE methods, where we have the benefits of SEM and the flexibility to implement mixed-effects for any subject level parameter, and include non-linear relations between factors or parameters. This research will show this application with an univariate DLO model.

# Chapter 2

# Bayesian Mixed-Effects Nonstationary Latent Differential Equation Model (BMNLDE)

This dissertation intends to develop and test a Bayesian Mixed-Effects Nonstationary Latent Differential Equation Model (BMNLDE). This model is a Bayesian implementation of an extension of the model presented by Boker et al. (2016). This section will explain the BMNLDE model and its Bayesian implementation of the DLO model.

BMNLDE is expressed by the two following equations, where Equation2.1 represents the measurement model from the data model where the embedded data ($y_{id}$) is predicted by the linear relation of the latent derivatives ($g_{ik}$) and latent equilibirum and slope ($b_{jz}$). Where $g_{ik}$ and $b_{jz}$ are approximated through data augmentation and define by the fixed factor loadings matrices $\lambda_{1d}$ and $\lambda_{2d}$.

$$y_{id} = \alpha_d + \lambda_{1d}g_{ik} + \lambda_{2d}b_{jz} + \varepsilon \tag{2.1}$$

$$\ddot{g}_i = \eta_j g_i + \zeta_j \dot{g}_i + \varepsilon_{\ddot{g}} \tag{2.2}$$

35

Equation 2.1 defines the measurement model of the latent derivatives and latent equilibrium intercept as well as the equilibrium slope—where $y_{id}$ is the delay-embedded dimension matrix for dimension $d$, $\alpha_d$ is the intercept for each embedded column $d$ that is fixed to 0, and $\lambda_{1d}$ is the matrix of fixed values to approximate the latent derivatives. For a $d = 5$, $\lambda_{1d}$ would be

$$\lambda_{15} = \begin{bmatrix} 1 & -2\Delta\tau & (-2\Delta\tau)^2/2 \\ 1 & -1\Delta\tau & (-1\Delta\tau)^2/2 \\ 1 & 0 & 0 \\ 1 & 1\Delta\tau & (1\Delta\tau)^2/2 \\ 1 & 2\Delta\tau & (2\Delta\tau)^2/2 \end{bmatrix} \tag{2.3}$$

In Equation 2.1, $g_{ik}$ represent the derivative scores for each row $i$ in the time-delay embedded matrix for derivative $k$. $\lambda_{2d}$ is the factor loading matrix of fixed values used to define the latent equilibrium intercept and slope. For a $d = 5$, $\lambda_{2d}$ would be

$$\lambda_{2d} = JH + C \tag{2.4}$$

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2\Delta\tau & -\Delta\tau & 0 & \Delta\tau & -2\Delta\tau \end{bmatrix} \tag{2.5}$$

$$H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \Delta\tau & \Delta\tau & \Delta\tau & \Delta\tau & \Delta\tau \end{bmatrix} \tag{2.6}$$

For the $j$th occasion of measurement within individual $i$'s data, we must create the matrix $J$ in the following form

$$J = \begin{bmatrix} 0 & 0 \\ 0 & j \end{bmatrix} \tag{2.7}$$

In Equation 2.1, $b_{jz}$ are the scores for the latent factors $z$ equilibrium intercept and slope for

each subject $j$, and $\varepsilon$ is the uniqueness for the columns of the delay embedded matrix, which are fixed to be equal for every column. This is an equality constraint that could be relaxed.

Equation 2.2 expresses the relation between the derivatives that estimate the DLO model. $\ddot{g}_i$ is the second derivative score for each row $i$ in the time-delay embedded matrix, $\eta_j$ are the $\eta$ parameters for each subject $j$, $g_i$ is the zeroth derivative score for each row $i$ in the time-delay embedded matrix, $\zeta_j$ are the $\zeta$ parameters for each subject $j$, $\dot{g}_i$ is the first derivative score for each row $i$ in the time-delay embedded matrix, and $\varepsilon_{\ddot{g}}$ is residual standard deviation from the second derivative in Equation 2.2.

Now that the BMNLDE has been define, the Bayesian specification of priors is the following. The delay-embedded dimension matrix is specified from a normal distribution,

$$y_{id} \sim N(\mu_d, \varepsilon^2) \tag{2.8}$$

where $y_{id}$ is the delay-embedded matrix, with a mean $\mu_d$ for each column $d$ of the matrix. This mean is the predicted value from Equation 2.1, and $\varepsilon^2$ is the residual variance, which will be assumed to be equal for every column of the delay-embedded matrix. The prior for the inverse of $\varepsilon^2$ will be

$$\frac{1}{\varepsilon^2} \sim \gamma(1, 0.5) \tag{2.9}$$

The zeroth and first derivatives will follow a multivariate normal distribution, whereas the second derivative follows a normal distribution, such as

$$(x_i, \dot{x}_i) \sim MVN(\mu_x, \Sigma_x) \tag{2.10}$$

where $\mu_x$ is the vector or derivatives means fixed to 0, and $\Sigma_x$ is the covariance matrix. The covariance matrices will have an Inverse-Wishart prior in the inverse of $\Sigma_x$ (Gelman & Hill, 2007).

$$\Sigma_x \sim Inverse-Wishart(I, df) \tag{2.11}$$

where $I$ is an identity matrix of the same dimensions as $\Sigma_x$, and $df$ is equal to the dimension of $\Sigma_x + 1$.

The second derivative follows a univariate normal distribution,

$$\ddot{x}_i \sim N(\mu_{\ddot{x}}, \sigma_{\ddot{x}}^2) \tag{2.12}$$

where $\mu_{\ddot{x}}$ is the mean that is define as the predicted value by Equation 2.2, and $\sigma\ddot{x}^2$ is the residual variance of the latent derivative, the prior for the inverse of $\sigma\ddot{x}^2$ is

$$\frac{1}{\sigma_{\ddot{x}}^2} \sim \gamma(1, 0.5) \tag{2.13}$$

The latent individual equilibrium intercept and slope also follow a multivariate normal distribution

$$b_{jz} \sim MVN(\mu_b, \Sigma_b) \tag{2.14}$$

where $\mu_b$ is the vector of means of the equilibrium intercept and slope across $j$ subjects, meaning that it is the sample equilibrium intercept and slope, and $\Sigma_b$ is the covariance matrix between the latent factors. $\Sigma_b$ follows the same type of prior as in Equation 2.11. And the $\mu_b$ have the following prior, for both the equilibrium intercept and slope

$$\mu_b \sim N(0, 100) \tag{2.15}$$

The individual parameters $\eta_j$ and $\zeta_j$ follow a normal distribution,

$$\eta_j \sim N(\mu_\eta, \sigma_\eta^2) \tag{2.16}$$

$$\zeta_j \sim N(\mu_\zeta, \sigma_\zeta^2) \tag{2.17}$$

where $\mu$ is the sample mean of the random parameters, and $\sigma^2$ is their variance. These $\mu$ sample

means had a prior of

$$\mu \sim N(-0.1, 5) \tag{2.18}$$

This is a reasonable weakly informative prior for $\eta$ and $\zeta$ because their value is expected to be closer to zero— this prior sets the 50% highest density of the distribution ranging from $[-1.61, 1.41]$. This prior distribution sets the estimation on the likely data space without leading the estimation. The variances $(\sigma_\eta^2, \sigma_\zeta^2)$ estimated have a prior of

$$\sigma^2 \sim U(0, 1) \tag{2.19}$$

This is a reasonable prior for the simulation, in function of the possible estimates for the latent variances given the data simulation parameters. It is recommended to include proper weakly informative prior distributions (Gelman & others, 2006). This prior distribution should be scaled in function of the data.

All the priors implemented in this model would be consider weakly informative (Gelman et al., 2013; Gelman & others, 2006), meaning that they would not guide the estimation of the parameters, but they represents a reasonable data space for the parameters.

## 2.1   Simulation Study

A simulation study was performed t to evaluate the performance of the proposed BMNLDE model. This simulation tested the BMNLDE model under different levels of the random parameters. The main goal is to identify the robustness of the model to estimate unbiased parameters for different conditions on the random parameters as well as their different combinations.

For this simulation, different values for three of the four random parameters varied, values for $\eta_j$, $\zeta_j$, and $b_s$ (equilibrium slope), from the Equations 2.1 and 2.2. The parameters that varied are the sample means ($\mu_\eta$, $\mu_\zeta$, and $\mu_{bs}$). These parameters are selected because they are the

Table 2.1: Parameters for simulated data that are fixed across conditions

| parameter | value |
|-----------|-------|
| $\mu_{bi}$ | 0 |
| $\sigma_{bi}$ | 2 |
| $\sigma_{bs}$ | 0.3 |
| $\sigma_{\eta}$ | 0.1 |
| $\sigma_{\zeta}$ | 0.1 |
| $\rho_x$ | 0 |
| $\rho_b$ | 0.3 |

main parameters of interest in the DLO model (Boker, 2007), and the variation of $b_s$ could present difficulties in estimating unbiased parameters. It is not expected that the location of $b_i$ (equilibrium intercept) will affect the performance of the BMNLDE model.

The elements of the simulated data that did not vary are shown in Table 2.1. Fixed values for the mean of the equilibrium intercept at 0; as the standard deviation for equilibrium intercept, slope, $\eta$, $\zeta$ as well as the correlations between the latent derivatives (0) and the latent intercept and slope (0.3).

The three parameters that varied across the condition will have four values each (Table 2.3), for $\mu_{\eta}$ ($\eta = -0.4, -0.2, -0.1, -0.05$). These values correspond approximately to 9.9, 14.0, 19.9, and 28.1 observations measured per cycle. The different levels of $\mu_{\zeta}$ show models with no damping and models with high damping ($\zeta = 0, -0.02, -0.05, -0.1$). The different $\mu_{bs}$ generated models with no linear growth as well as models with a clear linear trend ($b_s = 0, 0.05, 0.1, 0.2$). The time series were simulated with random white noise $N \sim (0, 1)$—this is scale, so the signal-to-noise ratio will be 1:1. Sample size and length of the time series was fixed at 50 subjects with 50 time points. The embedding dimension $d$ and $\tau$ was adjusted in function of $\mu_{\eta}$, so that the time-delay embedding corresponds to more than a $1/4$ and less than $1/2$ of the length of the cycle (Boker et al., 2004), the respective $d$ and $\tau$ are presented in Table 2.2. All conditions were crossed (64 cells), and 100 data sets were generated per cell.

For a total of 6400 data sets, each data set was analyzed with the BMNLDE and with the closest model from the frequentist framework. The frequentist model will be estimated with OpenMx

Table 2.2: Embedding dimension and $\tau$ for each value of $\mu_\eta$

| $\mu_\eta$ | $d$ | $\tau$ |
|------|---|---|
| -0.4 | 4 | 1 |
| -0.2 | 6 | 1 |
| -0.1 | 8 | 1 |
| -0.05 | 6 | 2 |

Table 2.3: Parameters for simulated data that vary across conditions

| parameter | values |
|-----------|--------|
| $\mu_\eta$ | -0.4, -0.2, -0.1, -0.05 |
| $\mu_\zeta$ | 0, -0.02, -0.05, -0.1 |
| $\mu_{bs}$ | 0, 0.05, 0.1, 0.2 |

(Neale et al., 2016). This model estimate the individual $\eta$, $\zeta$, and $b_{is}$, the covariance between $x$ and $\dot{x}$ had to be excluded due to convergence problems by including it. The parameters that BMNLDE estimates that OpenMx cannot estimate are the sample mean and standard deviation for the random parameters, the covariances between the derivatives and equilibrium intercept and slope ($\mu_\eta$, $\mu_\zeta$, $\mu_{bi}$, $\mu_{bs}$, $\sigma_\eta$, $\sigma_\zeta$, $\sigma_{bi}$, $\sigma_{bs}$, $\rho_x$, $\rho_b$), the sample mean and standard deviation for the random parameters are approximated by the mean of the subject parameters post estimation, descriptive statistics from the individual estimates (no standard errors) (Boker et al., 2016).

To illustrate the different kinds of oscillating time series that were simulated, the following figures (2.1, 2.2, 2.3) show how the time series changes as a function of one of the parameters ($\eta$, $\zeta$, $b_s$). The extreme state of these simulation conditions tested the robustness of this model in order to estimate proper parameters from processes with no clear cycle behavior. For example, conditions with high slope ($\mu_{bs} = 0.2$), high damping ($\mu_\zeta = -0.1$), and low $\eta$ ($\mu_\eta = -0.05$) present a behavior similar to a linear increase. These conditions will allow us to test the performance of the method under a wide range of processes.

The analysis was be executed using the general Bayesian software JAGS (Plummer, 2003), with its interface R package, R2jags (Su & Yajima, 2015). JAGS implements the Gibbs sampler to approximate the posterior distribution. Convergence of the Markov chains was determined using the potential scale reduction factor (PSRF), also know as univariate R-hat (see Gelman & Rubin,

Figure 2.1: Oscillating time series with varying values of $\eta$
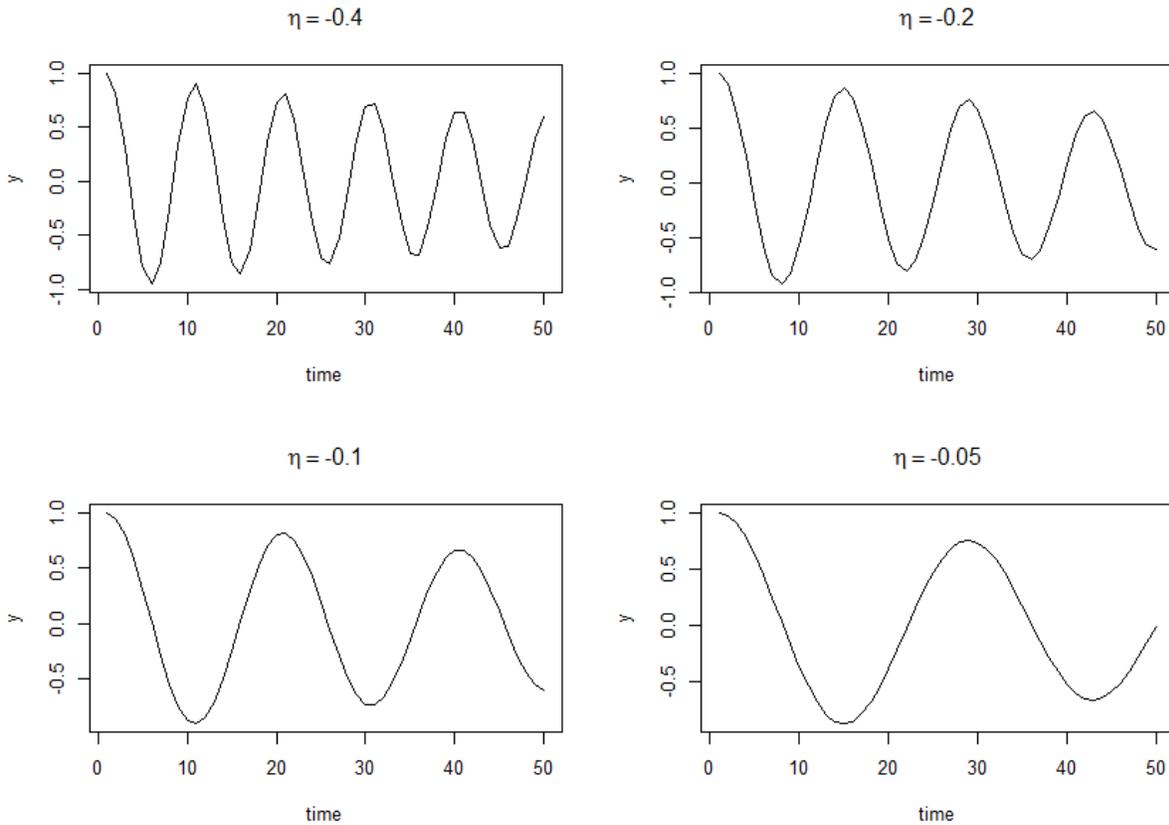


*Note.* All were simulated with $\zeta = -0.02$

1992). It was determined that the model converged when R-hat was lower than 1.10 for every parameter (Brooks & Gelman, 1998). Each data set was run several numbers of iterations, so it could achieve convergence, as in MCMC estimation different data sets may required different number of iterations to converge. Although the number of total iterations may be different, the number of kept iterations for analysis was always the same. The last 5000 iterations were kept from each chain. For every analysis we used 3 chains, meaning that the total number of iterations used for inferences was 15000. The thinning interval for all models was one, which indicated no thinning.

The performance of the BMNLDE model was evaluated at the subject and sample-parameter level. The models were evaluated by their capability to estimate unbiased parameters and by

Figure 2.2: Oscillating time series with varying values of $\zeta$



$\zeta = 0$

$\zeta = -0.02$

$\zeta = -0.05$

$\zeta = -0.1$
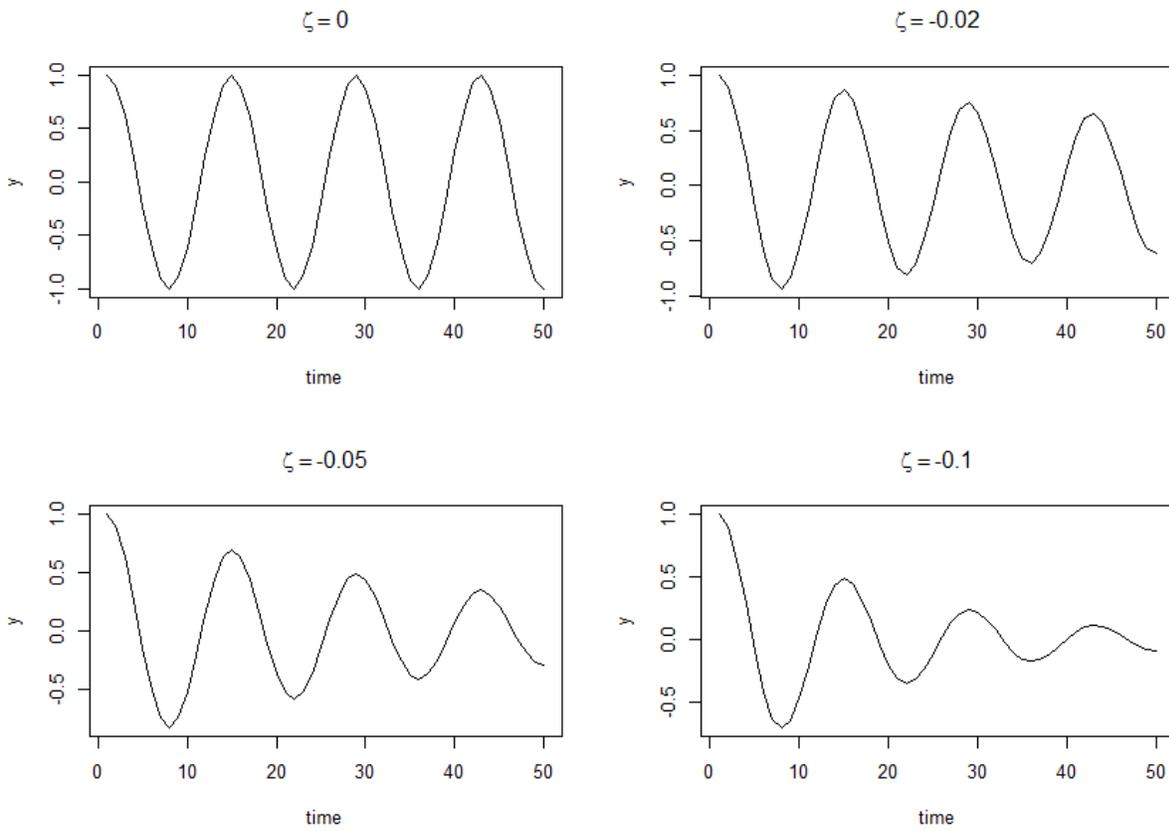
*Note.* All were simulated with $\eta = -0.2$

Figure 2.3: Oscillating time series with varying values of $b_s$

**Equilibrium slopes**
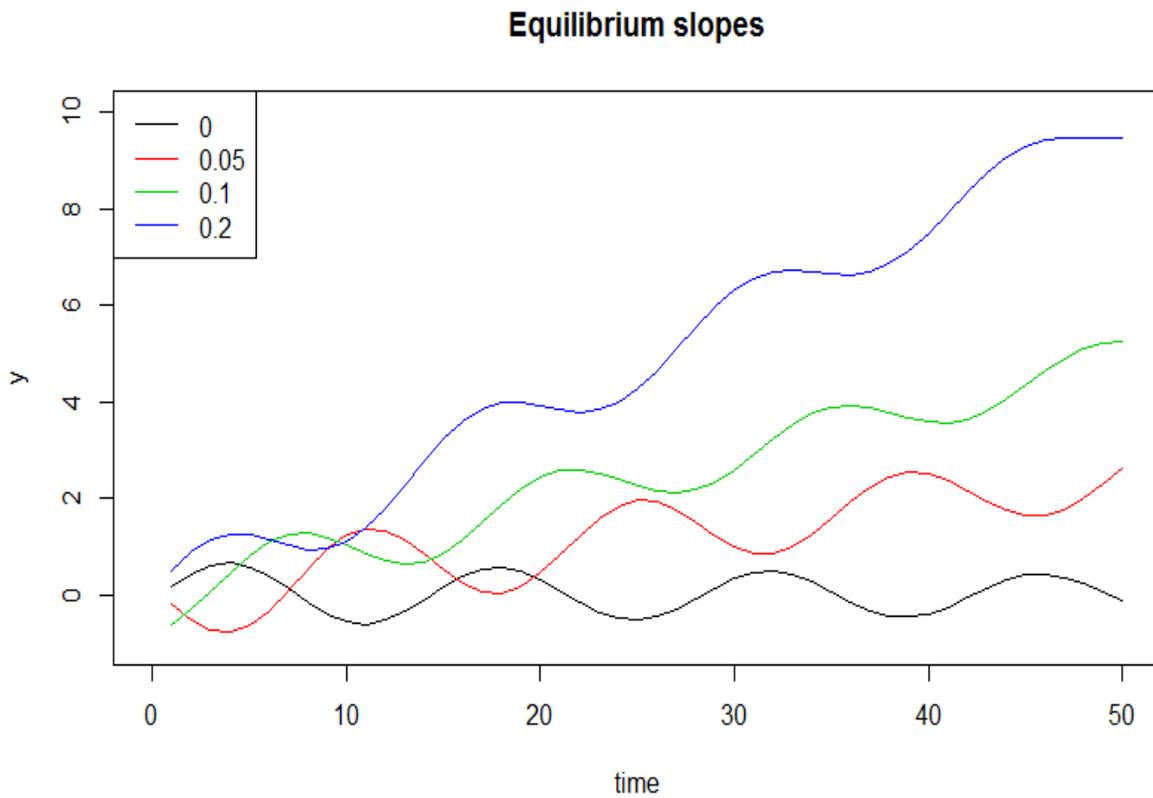


*Note.* All were simulated with $\eta = -0.2$ and $\zeta = -0.02$

the inclusion of the population value within the 95% Credible Interval (95% CI). The bias was measure by the distance of each estimated parameter from the population parameter $Bias = \theta - \hat{\theta}$, evaluate the quality of the estimator through the Mean Square Error (MSE) which is defined as $MSE = Bias^2 + Var(\hat{\theta})$, and the 95% CI was implemented to find if the population value is within the 95% CI and calculate the proportion of cases that include it from the different conditions. Because this model estimates subject and sample level parameters, the model will be evaluated at both levels. For the subject level we will use the average bias across all subjects for the four random parameters ($\eta$, $\zeta$, $b_i$, $b_s$), and for the sample we will evaluate the bias and 95% CI for $\mu_\eta$, $\mu_\zeta$, $\mu_{bi}$, $\mu_{bs}$, $\rho_x$, $\rho_b$, $\sigma_{bi}$, $\sigma_{bs}$, $\sigma_\eta$, $\sigma_\zeta$.

## 2.2 Results

### 2.2.1 Convergence

For the BMNLDE model, it was established that the model converged when all the parameters had a R-hat lower than 1.10. For the full set of conditions, the BMNLDE model converged 100% out of the 6400 data sets. The model did not estimate outliers parameters, outliers were defined as parameters 10 standard deviation away from the population mean. The closest frequentist model to BMNLDE had to be estimated without the covariance between derivatives, since introducing it generated convergence problems. The model presented 93.3% of proper convergence, 92.8% of the data sets converge without outliers. OpenMx model was sensitive to starting values, to avoid non-convergence due to this sensitivity the models were allowed to run up to seven times with different starting values, the model was stopped when it converged and not forced to run the seven times. The starting values were random numbers from normal distributions centered at the true population values.

Table 2.4: Average Bias between the two estimation methods

| parameter | BMNLDE | OpenMx | cohen $d$ | $\Delta$ |
|---|---|---|---|---|
| $b_{ji}$ | 0.053 | 0.065 | 0.070 | -0.012 |
| $b_{js}$ | 0.004 | 0.004 | 0.093 | 0.000 |
| $\eta_j$ | -0.024 | -0.027 | 0.163 | 0.003 |
| $\zeta_j$ | 0.024 | 0.026 | 0.085 | -0.003 |
| $\mu_{bi}$ | 0.049 | 0.063 | 0.040 | -0.014 |
| $\mu_{bs}$ | 0.005 | 0.004 | 0.013 | 0.001 |
| $\mu_\eta$ | -0.024 | -0.027 | 0.185 | 0.003 |
| $\mu_\zeta$ | 0.024 | 0.027 | 0.085 | -0.003 |
| $\rho_b$ | 0.057 | NA | NA | NA |
| $\rho_x$ | -0.007 | NA | NA | NA |
| $\sigma_{bi}$ | 0.161 | 0.171 | 0.037 | -0.010 |
| $\sigma_{bs}$ | 0.030 | -0.003 | 1.139 | 0.033 |
| $\sigma_\eta$ | -0.042 | -0.009 | 1.002 | -0.033 |
| $\sigma_\zeta$ | -0.038 | 0.049 | 3.982 | -0.087 |

## 2.2.2 Bias

The average bias was compare between the two estimation methods. Table 2.4 shows the average bias for the common parameters between models (excluding $\rho_b$ and $\rho_x$). The difference between the two methods is small. BMNLDE has lower bias in the parameters of the DLO, while OpenMx has lower bias in the standard deviations of the random effect $\sigma_{bs}$ and $\sigma_\eta$. OpenMx does not estimate the sample mean and standard deviations of the random effects in the model, they are a post calculation from the individual parameters.

An Analysis of Variance (ANOVA) was implemented to determine which varying conditions of the simulation have an effect on the bias. Given the large sample size, the $p - value$ was not used to decide which factors had a meaningful effect. The effect size $\eta_p^2$ was use to make a decision about which conditions had a meaningful effect on bias (Maxwell & Delaney, 2003). Table 2.5 shows the $\eta_p^2$ effect size for the bias of the parameters of interest. An effect was considered meaningful with an $\eta_p^2 \geq 0.06$, considered a medium effect size. ANOVA included all conditions and all interactions (two and three way), Table 2.5 only includes $\eta_p^2$ for the effect of the varying conditions for the parameters $\mu_\eta$, $\mu_\zeta$, and their interaction; the $\eta_p^2$ related to the varying of $\mu_{bs}$ are not included since it showed no effect ($\eta_p^2 < 0.06$) for any parameter.

Table 2.5: $\eta_p^2$ effect of varying conditions on bias

| parameter | BMNLDE | | | OpenMx | | |
|---|---|---|---|---|---|---|
| | $\mu_\eta$ | $\mu_\zeta$ | $\mu_\eta * \mu_\zeta$ | $\mu_\eta$ | $\mu_\zeta$ | $\mu_\eta * \mu_\zeta$ |
| $b_{ji}$ | 0.334 | 0.001 | 0.003 | 0.364 | 0.002 | 0.003 |
| $b_{js}$ | 0.647 | 0.005 | 0.005 | 0.635 | 0.004 | 0.005 |
| $\eta_j$ | 0.527 | 0.017 | 0.006 | 0.460 | 0.015 | 0.005 |
| $\zeta_j$ | 0.679 | 0.357 | 0.166 | 0.788 | 0.007 | 0.226 |
| $\mu_{bi}$ | 0.070 | 0.001 | 0.007 | 0.076 | 0.001 | 0.007 |
| $\mu_{bs}$ | 0.022 | 0.000 | 0.008 | 0.020 | 0.000 | 0.009 |
| $\mu_\eta$ | 0.526 | 0.016 | 0.006 | 0.467 | 0.012 | 0.004 |
| $\mu_\zeta$ | 0.563 | 0.258 | 0.104 | 0.698 | 0.004 | 0.147 |
| $\rho_b$ | 0.400 | 0.002 | 0.003 | NA | NA | NA |
| $\rho_x$ | 0.297 | 0.408 | 0.088 | NA | NA | NA |
| $\sigma_{bi}$ | 0.397 | 0.002 | 0.007 | 0.402 | 0.001 | 0.007 |
| $\sigma_{bs}$ | 0.000 | 0.002 | 0.004 | 0.001 | 0.003 | 0.003 |
| $\sigma_\eta$ | 0.704 | 0.002 | 0.003 | 0.919 | 0.010 | 0.005 |
| $\sigma_\zeta$ | 0.314 | 0.025 | 0.032 | 0.233 | 0.048 | 0.021 |

$\mu_\eta$ is the varying condition that has the greater effect on the bias estimates, while two parameters are not affected by it or any other condition ($\mu_{bs}$ and $\sigma_{bs}$). In the case of $\mu_\zeta$ eight parameters are not affected by it, and five are affected by it. For their interaction ($\mu_\eta * \mu_\zeta$), three parameters are affected by it ($\zeta_j$, $\mu_\zeta$, and $\rho_x$).

For the parameters that are affected by both $\mu_\eta$ and $\mu_\zeta$ we can see (Table 2.6) that the trend is that as $\mu_\eta$ gets closer to zero and $\mu_\zeta$ gets further from zero the bias increases. Meaning that the models with high damping and low frequency are less accurate. In the case of the parameters that are only affected by $\mu_\eta$ (Table 2.7), the results follow a similar trend where the bias increase as $\mu_\eta$ approaches 0. Comparing the two estimation methods (Table 2.6 and Table 2.7) we can see little difference between them, the bias estimates in general are smaller for the BMNLDE, but not substantially.

For the parameters that are not affected by any condition, their respective average bias was $\mu_{bs} = 0.004$ and $\sigma_{bs} = 0.029$ for BMNLDE, and $\mu_{bs} = 0.004$ and $\sigma_{bs} = -0.003$ for OpenMx, the population values for these parameters were $\mu_{bs} = 0, 0.05, 0.1, 0.2$ and $\sigma_{bs} = 0.3$.

The bias of the random effects means can be calculated as $z - scores$, looking at the distance

Table 2.6: Average Bias for BMNLDE and OpenMx across $\mu_\eta$ and $\mu_\zeta$

| $\mu_\eta$ | population | $-0.40$ | $-0.20$ | $-0.10$ | $-0.05$ |
|---|---|---|---|---|---|
| **BMNLDE** | | | | | |
| $\mu_\zeta = 0$ | | | | | |
| $\zeta_j$ | $\approx 0$ | 0.000 | 0.002 | 0.020 | 0.031 |
| $\mu_\zeta$ | 0 | -0.001 | -0.001 | 0.023 | 0.033 |
| $\rho_x$ | 0 | 0.000 | 0.001 | 0.020 | 0.022 |
| $\mu_\zeta = -0.02$ | | | | | |
| $\zeta_j$ | $\approx -0.02$ | 0.003 | 0.006 | 0.023 | 0.038 |
| $\mu_\zeta$ | $-0.02$ | 0.003 | 0.009 | 0.023 | 0.038 |
| $\rho_x$ | 0 | -0.009 | -0.005 | 0.005 | 0.015 |
| $\mu_\zeta = -0.05$ | | | | | |
| $\zeta_j$ | $\approx -0.05$ | 0.006 | 0.011 | 0.031 | 0.055 |
| $\mu_\zeta$ | $-0.05$ | 0.006 | 0.013 | 0.030 | 0.051 |
| $\rho_x$ | 0 | -0.021 | -0.020 | -0.014 | 0.011 |
| $\mu_\zeta = -0.10$ | | | | | |
| $\zeta_j$ | $\approx -0.10$ | 0.009 | 0.020 | 0.044 | 0.079 |
| $\mu_\zeta$ | $-0.10$ | 0.010 | 0.020 | 0.044 | 0.078 |
| $\rho_x$ | 0 | -0.044 | -0.050 | -0.041 | 0.003 |
| **OpenMx** | | | | | |
| $\mu_\zeta = 0$ | | | | | |
| $\zeta_j$ | $\approx 0$ | 0.000 | 0.006 | 0.043 | 0.055 |
| $\mu_\zeta$ | 0 | 0.000 | 0.002 | 0.046 | 0.056 |
| $\mu_\zeta = -0.02$ | | | | | |
| $\zeta_j$ | $\approx -0.02$ | -0.005 | 0.008 | 0.042 | 0.059 |
| $\mu_\zeta$ | $-0.02$ | -0.005 | 0.010 | 0.043 | 0.060 |
| $\mu_\zeta = -0.05$ | | | | | |
| $\zeta_j$ | $\approx -0.05$ | -0.011 | 0.005 | 0.046 | 0.073 |
| $\mu_\zeta$ | $-0.05$ | -0.010 | 0.008 | 0.045 | 0.069 |
| $\mu_\zeta = -0.10$ | | | | | |
| $\zeta_j$ | $\approx -0.10$ | -0.027 | 0.002 | 0.051 | 0.092 |
| $\mu_\zeta$ | $-0.10$ | -0.025 | 0.002 | 0.052 | 0.092 |

Table 2.7: Average Bias for BMNLDE and OpenMx across $\mu_\eta$ values

| $\mu_\eta$ | population | $-0.40$ | $-0.20$ | $-0.10$ | $-0.05$ |
|---|---|---|---|---|---|
| **BMNLDE** | | | | | |
| $b_{ji}$ | $\approx 0$ | 0.037 | 0.019 | -0.020 | 0.176 |
| $b_{js}$ | $\approx 0, 0.05, 0.1, 0.2$ | 0.000 | 0.001 | 0.006 | 0.010 |
| $\eta_j$ | $\approx -0.40, -0.20, -0.10, -0.05$ | -0.014 | -0.010 | -0.033 | -0.038 |
| $\mu_{bi}$ | 0 | 0.043 | 0.017 | -0.050 | 0.188 |
| $\mu_\eta$ | $-0.40, -0.20, -0.10, -0.05$ | -0.015 | -0.009 | -0.032 | -0.038 |
| $\rho_b$ | 0.3 | 0.021 | 0.002 | -0.007 | 0.212 |
| $\sigma_{bi}$ | 2 | 0.038 | 0.039 | 0.099 | 0.467 |
| $\sigma_\eta$ | 0.1 | -0.031 | -0.027 | -0.041 | -0.068 |
| $\sigma_\zeta$ | 0.1 | -0.035 | -0.055 | -0.028 | -0.035 |
| **OpenMx** | | | | | |
| $b_{ji}$ | $\approx 0$ | 0.037 | 0.020 | -0.008 | 0.193 |
| $b_{js}$ | $\approx 0, 0.05, 0.1, 0.2$ | 0.000 | 0.001 | 0.005 | 0.009 |
| $\eta_j$ | $\approx -0.40, -0.20, -0.10, -0.05$ | -0.023 | -0.010 | -0.035 | -0.040 |
| $\mu_{bi}$ | 0 | 0.043 | 0.019 | -0.041 | 0.205 |
| $\mu_\eta$ | $-0.40, -0.20, -0.10, -0.05$ | -0.024 | -0.009 | -0.036 | -0.040 |
| $\sigma_{bi}$ | 2 | 0.050 | 0.047 | 0.100 | 0.468 |
| $\sigma_\eta$ | 0.1 | 0.052 | -0.007 | -0.029 | -0.056 |
| $\sigma_\zeta$ | 0.1 | 0.061 | 0.035 | 0.064 | 0.041 |

between the estimated parameter and the population means in function of the standard deviations of the random effects. Table 2.8 presents the average $z-scores$ of the estimated parameters in function of the population means. We can see that when the estimates are relative to the population distribution, the bias is low. For the equilibrium intercept and slope the estimates are lower that 0.04 standard deviations away from the population mean. For $\mu_\eta$ and $\mu_\zeta$ the estimated parameters are at 0.26 standard deviations away from the mean, these are in function of a small standard deviation ($\sigma = 0.1$), meaning that in average the estimates are 26% of a 0.1 standard deviations away from the population mean. For the mean of the random effects, the estimated parameters are close to the population mean, showing good parameter recovery.

## 2.2.3   95% Credible Interval coverage

The 95% Credible Interval coverage indicates how precise is the model to lead us to the right inference about the population values. The 95% CI coverage was not calculated for $b_{ji}$, $b_{js}$, $\eta_j$,

Table 2.8: Average $z-score$ of the estimated parameters from the population means

| parameter | $\sigma$ | BMNLDE | OpenMx |
|---|---|---|---|
| $\mu_{bi}$ | 2 | 0.025 | 0.032 |
| $\mu_{bs}$ | 0.3 | 0.015 | 0.013 |
| $\mu_\eta$ | 0.1 | -0.236 | -0.268 |
| $\mu_\zeta$ | 0.1 | 0.237 | 0.265 |

Table 2.9: 95% Credible Interval coverage for BMNLDE across $\mu_\eta$ and $\mu_\zeta$

| $\mu_\eta$ | $-0.40$ | $-0.20$ | $-0.10$ | $-0.05$ |
|---|---|---|---|---|
| $\mu_\zeta = 0$ | | | | |
| $\mu_\zeta$ | 100 | 99.25 | 79.00 | 43.00 |
| $\rho_x$ | 100 | 100 | 98.00 | 98.00 |
| $\mu_\zeta = -0.02$ | | | | |
| $\mu_\zeta$ | 100 | 95.75 | 78.75 | 24.50 |
| $\rho_x$ | 100 | 100 | 100 | 100 |
| $\mu_\zeta = -0.05$ | | | | |
| $\mu_\zeta$ | 100 | 96.75 | 62.50 | 5.50 |
| $\rho_x$ | 100 | 98.00 | 99.00 | 100 |
| $\mu_\zeta = -0.10$ | | | | |
| $\mu_\zeta$ | 100 | 94.25 | 30.25 | 0 |
| $\rho_x$ | 99.00 | 76.75 | 85.50 | 99.00 |

$\zeta_j$; because these are the average bias from subjects $j$ parameters. The 95% CI coverage will be calculated for the sample parameters, these results will be presented as the previous section, where the parameters will be divided by which conditions of the simulation affect the bias. The 95% CI coverage can not be calculated for the OpenMx results since the sample statistics are descriptive post calculations from the subjects it does not provide standard errors or Confidence Intervals.

For the parameters that are not affected by the conditions the 95% CI coverage was $\mu_{bs} = 96.41$ and $\sigma_{bs} = 90.22$. For the parameters that are affected by the conditions in the simulation, the 95% CI coverage decreases as $\mu_\eta$ decreases and $\mu_\zeta$ increases (Table 2.9 and Table 2.10). Parameters that are particularly problematic are $\mu_\zeta$, $\mu_\eta$, $\sigma_\eta$, and $\sigma_\zeta$. The condition where $\sigma_\eta$ approaches 0 have unreliably low 95% CI coverage for these parameters.

Table 2.10: 95% Credible Interval coverage for BMNLDE across $\mu_\eta$ values

| $\mu_\eta$ | $-0.40$ | $-0.20$ | $-0.10$ | $-0.05$ |
|---|---|---|---|---|
| $\mu_{bi}$ | 92.12 | 93.69 | 92.94 | 85.25 |
| $\mu_\eta$ | 97.75 | 90.06 | 6.94 | 0 |
| $\rho_b$ | 86.56 | 85.00 | 84.56 | 29.12 |
| $\sigma_{bi}$ | 93.88 | 93.81 | 93.94 | 46.56 |
| $\sigma_\eta$ | 81.94 | 40.00 | 0.06 | 0 |
| $\sigma_\zeta$ | 99.81 | 13.88 | 69.25 | 52.00 |

Table 2.11: Mean Square Error for the BMNLDE and OpenMx models

| parameter | BMNLDE | OpenMx |
|---|---|---|
| $\mu_{bi}$ | 0.125 | 0.126 |
| $\mu_{bs}$ | 0.002 | 0.002 |
| $\mu_\eta$ | 0.001 | 0.001 |
| $\mu_\zeta$ | 0.001 | 0.002 |
| $\rho_b$ | 0.024 | NA |
| $\rho_x$ | 0.001 | NA |
| $\sigma_{bi}$ | 0.107 | 0.108 |
| $\sigma_{bs}$ | 0.002 | 0.001 |
| $\sigma_\eta$ | 0.002 | 0.002 |
| $\sigma_\zeta$ | 0.002 | 0.003 |

### 2.2.4 Mean Square Error

Mean Square Error (MSE) allow us the evaluate the quality of the estimator, and is a good metric to compare estimation methods. Table 2.11 shows the MSE for the parameters for both BMNLDE and OpenMx models. We can see that these parameters have a low MSE for every parameter, and that the MSE does not show differences between the two estimation methods. Given the calculation of the MSE, the results indicate that bias is the main contributor to it; and the two estimation methods are equally stable.

## 2.3 Discussion

The simulation was design to test the performance of the BMNLDE model to estimate proper parameters under different characteristics of an oscillation process. The simulation tested both the performance to estimate subject and sample parameters. Varying conditions were in function of

$\mu_\eta$, $\mu_\zeta$, and $\mu_{bs}$. Testing the model under different conditions of damping, slope, and frequency. The BMNLDE model was compare to the closest frequentist model estimated with OpenMx.

The simulation tested the BMNLDE estimating a single group DLO model. The BMNLDE can be implemented into a wide variety of relations between derivatives and equilibrium intercept and slope. It can be apply to estimate couple models, where the derivatives are estimated for several constructs of interested and related between each other. The coupled effects between constructs can be estimated as mixed-effects. As will be shown in the next section, it can also be apply to multiple group models and compare parameters of interest between groups. The BMNLDE is a general method to estimate mixed-effects differential equations.

In general the model presents appropriately low bias. Making the estimated parameters relative to the population distribution identify that the means of the random effects are 0.25 from the population means in cases where the standard deviation is 0.1 ($\mu_\eta$, and $\mu_\zeta$), where in cases where the standard deviation is higher ($\mu_{bi} = 2$, and $\mu_{bs} = 0.3$) the estimates are under 0.03 standard deviations way from the population means. The condition that has more effect on bias is the different values of $\mu_\eta$. As $\mu_\eta$ approaches 0 bias tends to increase in several parameters, where data sets with $\mu_\eta = -0.05$ shows problems to estimate accurate parameters. This indicates that the model is less precise for oscillations with low frequency. For the parameters that are affected by the different values of $\mu_\zeta$, we find that bias increase as the damping of the model increases. Leading that data sets with low frequency and high damping are difficult to estimate properly. The equilibrium slope does not affect the proper recover of parameters.

$\mu_\eta$ is related to bias, since with 50 time points the condition $\mu_\eta = -0.05$ would complete less than two full cycles, while the condition $\mu_\eta = -0.4$ has completed five full cycles. Having more complete cycles allows the model to identify the oscillating process with more precision. The results indicate that five complete cycles are needed to estimate stable results, although more research is needed to determine clearly how much a cycle has to be sample.

The inferences made from the model, are also affected by the different population values of $\mu_\eta$, $\mu_\zeta$, following the same trend, data sets with low frequency and high damping are less likely to

52

give a correct inference about the population values. The parameters with serious problems in this area are $\mu_\zeta$, $\mu_\eta$, $\sigma_\eta$, and $\sigma_\zeta$, where the 95% CI coverage is unreliably low as $\mu_\eta$ approaches 0 and $\mu_\zeta$ increases. More research is needed to identify this problem and a parametrization that provide proper inferences.

Comparing the BMNLDE and OpenMx results related to bias they are equivalent in the recovery of the parameters. The MSE does not identify one estimation method as superior over the other, as both present equivalent values. The differences between the estimation methods are (a) BMNLDE model is capable to estimate all the parameters at the subject and sample levels, while OpenMx can not estimate the sample level parameters, which have to be calculated from a post processing of the subject level data. The main disadvantage of this is that OpenMx can not provide standard errors or Confidence intervals to make inferences about these parameters, while BMNLDE is capable to estimate each parameter and include a Credible Interval to inferences about it. (b) OpenMx is not capable to estimate the correlation between equilibrium intercept and slope, and the inclusion of the correlation between the zeroth and first derivatives lead to convergence problems. While the BMNLDE is capable to estimate both of these correlations without convergence problems. (c) OpenMx is sensitive to the provided starting values, it was necessary to provide starting values close to the true population values for the models to converge. (d) Computing time and resources are commonly refer to as a drawback of bayesian estimation, in this case the BMNLDE was faster to converged to a final solution than OpenMx. BMNLDE took from 15 minutes to 2 hours, while OpenMx took from 1 hour to 10 hours. BMNLDE was allowed to run as many iterations needed to converged, while OpenMx was allow to run up to seven times (with different starting values) to converge. (d) Finally, the BMNLDE model allow us to make direct inferences about the parameters based in Bayesian estimation (more research is needed to identify under which conditions these inferences are reliable), while the OpenMx model is limited to indirect inferences in function of the null hypothesis.

The simulation shows the flexibility of the BMNLDE model to estimate mixed-effects differential equations from the SEM framework, the BMNLDE model is proper model to incorporate

intra and inter individual variation. Extending the previous model that allowed the inclusion of subject parameters and equilibrium intercept and slope. The present model is capable to estimate subject and sample level parameters. The bayesian framework showed to be flexible to estimate a model that can not be estimated from the frequentist framework, and identify the main differences between it and the closest frequentist model. It was able to demonstrate circumstances in which the model estimate proper parameters, also show the limitation of the model, in which conditions the model tends to increase bias of the estimated parameters and provide low 95% CI coverage. It is useful for researchers to understand the advantages and limitations of models.

## 2.4   Limitations and future directions

A limitation of this simulation is the conditions that are the parameters varied. To have a more comprehensive understanding of the BMNLDE model it is needed to tested on more conditions. Future research has to extend the varying parameters to continue testing the BMNLDE model. Characteristics that should be tested to extend the present study are: sample size, length of time series, $\mu_\eta$, $\rho_x$, $\sigma_\eta$, $\sigma_\zeta$, priors, complete cycles.

Sample size and length of the time series should be tested to find how low the sample size can be and still estimate proper parameters. It would be expected for these two characteristics to interact, meaning that lower samples would benefit on longer time series, looking to identify an interaction in estimation precision between sample size and length of the time series. The Bayesian framework is known to have the possibility to estimate proper models for low samples, which would lead to a higher effect of the model priors into the model. In the case of intensive longitudinal data, is generally complicated to have a large sample size and still have a long time series. A typical example would be like the one presented by Watts et al. (2016) and the applied example in the next chapter, where the groups are small with long time series.

Since the population value that has the greater effect on bias was $\mu_\eta$, it would be of interest to extend the understanding of its effect on the quality of the model estimates. Given the difference

in precision at low values $\mu_\eta = -0.05$ it is of interest to understand the behavior of the model with more detail to identify at which point of $\mu_\eta$ the model lose accuracy, at which level of $\mu_\eta$ bias increases.

Test the model with different number of complete cycles is needed to identify the minimum number of cycles needed to estimate accurate estimates. This would lead to an important recommendation for apply research. When collecting data this will give direct recommendation on how long the time series need to be collected.

Since $\rho_x$ is a parameter that may interact to affect the oscillation process, it would of interest to estimate the BMNLDE model when the oscillation is also affected by $\rho_x$. Given that the present study did not vary population values of $\sigma_\eta$, $\sigma_\zeta$; it is of interest to identify under which values these $\sigma$ the model improves their estimation, or if a different parametrization of these standard deviations may improve their proper recovery.

The BMNLDE model was implemented for every condition with the same priors, weakly informative priors. Bayesian models can be affected by the selected priors. It would be of interest of know how sensitive are the parameters to different priors; for example, a set of priors could focus on low frequency oscillations. There are several characteristic of the priors that could be modified, such as functional form, mean or variability of the distributions. This is drawback of the flexibility provided by the bayesian framework, since the model can be parametrized in different ways and the decisions made can have impact on the quality of the estimation. This require more research into the different model characteristics and how they may interact with other factors such as sample size and length of the time series.

# Chapter 3

# Sedentary Behavior in Older Adults

Physical activity (PA) decreases and sedentary time increases with age (DiPietro, 2001). Several factors influence physical activity in older adults (e.g., physiological, psychological, knowledge, social support, and environmental factors). Characteristics including self-motivation, confidence, and peer reinforcement are related to adoption and adherence to physical activity. Age-related physical changes in weight-bearing, strength, and flexibility are influential in the preservation of bone and lean mass, are among the highest priority targets for intervention with regard to maintaining functional ability and independence (DiPietro, 2001).

Physical activity is positively related to increased cognitive function and overall health outcomes for older adults and mitigate chronic diseases including obesity, high blood glucose levels, type 2 diabetes, and cardiovascular problems (Chastin & Granat, 2010; DiPietro, 2001). Modern lifestyles are conducive to sedentary behavior (SB), which also reduces the likelihood of health behavior engagement (Chastin & Granat, 2010). SB refers to activities that do not increase energy expenditure above the resting level and includes sleeping, siting, lying down, watching TV, and other forms of screen-based entertainment (Proper et al., 2011). Physical activity is a determinant of successful aging, which is defined as the absence of disease and disease related disability, high functional capacity, and active engagement in life (Baker et al., 2009). Overall, physical activity is related to fewer symptoms of age-related diseases, contributing to successful aging and is a critical

component of lifestyle interventions for older adults (Wang et al., 2015). Healthy lifestyle is recommended for controlling risk factor for cardiovascular disease (obesity, hypertension, diabetes, metabolic syndrome, and hypercholesterolemia) and reduces risk for stroke (Gandy & DeKosky, 2013). Improved health behaviors will likely lead to decreased rates of vascular dementia and possibly AD, through the mitigation of AD related vasculopathy (Gandy & DeKosky, 2013).

Increased physical activity limits age-related decline in cardiovascular function and thus lessens brain hypoxia and consequent cognitive decline (DiPietro, 2001). Age-related cognitive decline is related to progressive brain tissue loss, while greater physical activity is related to greater tissue volume in the white matter of the corona radiata extending into the parietal-occipital junction (Ho et al., 2011). Older adults who are physically active tend to have better cognitive performance and reduced risk of cognitive impairment (Colcombe & Kramer, 2003). Physical activity and Lifelong physical activity is related to better performance on measures of executive function, but not consistently related to measures of episodic memory (Eggermont et al., 2009; Gajewski & Falkenstein, 2015a; Gajewski & Falkenstein, 2015b).

The measurement of sedentary behavior and physical activity has improved due to the use of objective measures including the activPAL and the ActiGraph (Kang & Rowe, 2015; Kim et al., 2015). Accelerometers are generally classified as energy-expenditure or postural devices. Energy-expenditure devices (i.e., ActiGraph) measure the frequency and amplitude of accelerations generated from ambulatory movements. The data generated by energy-expenditure devices is usually presented as activity counts for a defined time interval (i.e., epochs). Several activity thresholds have been proposed to establish different levels of PA (e.g., light, moderate, vigorous). Using the ActiGraph, sedentary behavior (SB) is classified when the activity equals $< 100$ counts per minute. Postural devices (i.e., activPAL) use an inclinometer to detect postural information (e.g., sitting, or standing). The monitors are placed on different parts of the body depending on the objective. The activPAL distinguishes postural changes more accurately since it is placed on the thigh or lower back, which are more sensitive to changes in sitting and standing compared to monitors placed on th wrist or hip. Sitting and lying during waking hours is operationalized as SB (Kim et al., 2015).

Little methodological work exists to inform how researchers should use accelometry data. The first step in the use of accelometry data is to divide the raw activity data into epochs and to summarize the data into segments of predetermined length (e.g., 1 hour). Even when the monitors collect physical activity data over the course of several days, the data is commonly used as cross-sectional data because the repeated measures are summarized in total numbers, such as the percentage of time spent sedentary (Chastin & Granat, 2010; Chastin et al., 2010; van der Niet et al., 2015; Godfrey et al., 2014; Janssen & Cliff, 2015). Epoch-based data is limited in that epochs reflect the summation of observations across an arbitrary window of continuous time (Janssen & Cliff, 2015).

Existing methods for the analysis of longitudinal behavioral data include the use of the coefficient of variation (Cavanaugh et al., 2010), Fourier transformations to identify periodicity of physical activity (Rowlands et al., 2015), or longitudinal analyses using the generalized linear models (Ridgers et al., 2015). These longitudinal analyses are limited in that they summarize the data in epochs. For example, Ridgers et al. (2015) and Rowlands et al. (2015) summarized the data for each day, looking only at variations between days. The selection of epoch length is related to the kind of question that can be asked. These methods lack the use of continuous-time models, meaning that the results and inferences are limited to the chosen epoch length.

The BMNLDE model will be applied to a real data set. Consisting of time series from the physical activity of older adults who wore an activPAL monitor, which has been shown to produce a valid measure of physical activity and sedentary behavior (Grant et al., 2008; Reid et al., 2013; Aguilar-Farias et al., 2015; Kim et al., 2015). The goal of the present analysis is to describe the SB of healthy older adults and older adults with mild Alzheimer Dementia (AD), and to identify differences in patterns of sedentary behavior between these groups. SB is expected to follow an oscillating form over time (Rowlands et al., 2015), and SB is expected to vary around an average (equilibrium) SB. It is also expected for the SB to oscillate within days, and in the case of this sample, it is not expected to have an oscillation between days (i.e., we expect variation across hours of the day but not across days of the week). It is of interest to implement a model that identifies the SB oscillating process within continuous time models so that the results will not be

dependent on the chosen epoch. It is also of interest to find out how much subjects vary from their group means, accounting for inter-individual variability. At the same time we need to be able to compare the healthy and demented (AD) older adults.

The Damped Linear Oscillator (DLO) model is a model of interest to describe the SB in this sample because the DLO describes the oscillating behavior of SB while identifying a damping effect, if the oscillation tends to stop oscillating at a constant equilibrium, meaning that would identify if subjects tend to stop oscillating at a constant time spent standing. However, a damping effect is not expected in the present sample as it is expected for the cyclic behavior to be constant over time. The goal of the present study was to find the oscillating process of SB within days and to identify a cycle of SB on the order of hours. The estimation of the equilibrium allows for the comparison of the average SB times for each subject and group. The estimation of the equilibrium slope will determine whether the average SB time changes across time. In this sample, we expected to identify equilibrium differences between healthy and AD groups. The implementation of the BMNLDE allows for the analysis of SB time series data with the objective of incorporating intra and interindividual variability. Within the framework of a Differential Equation (DE), the cyclical dynamics of SB can be modeled while taking into account intraindividual variation from a continuous-time perspective. By implementing the mixed-effect approach, we can allow each participant to follow their own cycle and to estimate the mean level of the overall sample. Given that the BMNLDE model estimates sample level parameters, it can also be used to compare group means for the random parameters—making it particularly suited to compare the healthy and AD groups.

## 3.1   Data Description

SB is operationalized based on body position, where standing or stepping positions will be accounted as non-sedentary, and sitting and lying down will be counted as SB. To identify the body position we used the activPAL monitor worn on the dominant thigh (Kim et al., 2015). The mon-

Table 3.1: Description of the sample

|  | Healthy | AD | Total |
|---|---|---|---|
| Sample size | 48 | 37 | 85 |
| % of males | 33.3 | 64.9 | 47.1* |
| % of white | 95.8 | 86.5 | 91.8 |
| Age (SD) | 73.3 (6.8) | 73.3 (7.7) | 73.3 (7.2) |
| Education (SD) | 17.1 (3.4) | 15.7 (2.8) | 16.5 (3.2) |
| BMI (SD) | 26.6 (4.2) | 26.5 (4.6) | 26.5 (4.4) |
| VO2 max (SD) | 1.6 (0.4) | 1.6 (0.5) | 1.6 (0.5) |
| Number of valid 10 minute epochs (range) | 455.1 (394 - 503) | 440.0 (325 - 505) | 448.6 (325 - 505) |

* indicates a group mean difference at a $p < .01$

itor was worn by 91 participants for a minimum of 5 days. At the same time, the participants maintained a sleep diary to validate the activPAL data for waking hours. Of the 91, six subjects experienced monitor malfunctions, so their data were not included. The final sample size was 85.

The activPAL data was wear-time validated. Wear-time validation is a preprocess to verify that the monitor was worn and to define sleep time. The wear-time validation was completed using the sleep diaries; we excluded any time that the subjects were sleeping during the night. Sleep time during the day (naps) were not excluded, but counted as SB. Thus, data represent activity during daytime waking hours. A day was considered a valid day when it included at least 10 hours of valid wear-time per day. Subjects' number of valid days varied from 5 to 12. To keep consistent the number of valid days, we selected the first 5 valid days for each subject. This follows the best practices to monitor and evaluate habitual physical activity, to include days with at least 10 hours of valid time, and at least four days to reliably estimate the habitual physical activity (Gretebeck & Montoye, 1992; Trost et al., 2005).

The activPAL monitors describe the physical activity (e.g., sitting, standing, walking) in 15 second epochs. For the analysis we used 10 minute epochs, summing up the seconds of standing time for each 10 minute epoch. By using 10 minute epochs, every subject has at least 300 measures. The 10 minute epochs allowed the data to follow a variety of cycles that could be from 1 to 7 hour cycles.

The sample descriptive statistics are shown in Table 3.1. This indicates that the two groups are

similar in characteristics like age, years of education, body max index (BMI), and VO2max (maximum volume of oxygen) during a treadmill test. Where the two groups differ is the percentage of males, the healthy group has lower percentage of males. This shows that the groups are similar in characteristics that are related to their physical performance, and there are no sample differences that would impair ability to be physically active in either group. The number of 10 minute epochs shows that there is variation on how much time the subjects spent awake for the 5 valid days, the minimum time-series is of 325 epochs (54.2 hours), and a maximum length of 505 epochs (84.2 hours).

## 3.2   Analysis

The first step was defining the right embedding dimension $d$ for this data. We followed the method proposed by Hu et al. (2014) to estimate the model with different $d$, selecting the $d$ once $\eta$ has stabilized, looking for $d$ with respect to the Nyquist limit, which states that $d$ must be less than one half the period of the oscillation one wishes to estimate (Boker et al., 2004). For this, we detrended and mean centered each time series to estimate the series of LDE in their simple form without mixed-effects or equilibrium changes. These analyses were performed with the R package OpenMx (Neale et al., 2016). This is in function of time because the BMNLDE takes longer to estimate than the simple form of DLO, but this model characteristic is not be affected by the other characteristics of the model.

Once we defined the $d$, we used the original time series to estimate the BMNLDE model. All the analyses were run using JAGS (Plummer, 2003) with its R interface package R2jags (Su & Yajima, 2015), and the models were allowed to run until they reached convergence. We determined that the model converges when R-hat is lower than 1.10 (Brooks & Gelman, 1998). The models were allowed to run for as many iterations as needed to converge, were run for 3 chains, and saved the last 2000 iterations of each chain with no thinning; which gave us 6000 iterations to make inferences about the posterior distributions of the parameters.

In the first model, we estimated the BMNLDE for the whole sample, and in a second model, we estimated the BMNLDE with group (i.e., healthy and AD) mean differences for the four random parameters ($\mu_\eta$, $\mu_\zeta$, $\mu_{bi}$, $\mu_{bs}$), and standard deviation differences for these random parameters, this will indicate if the subjects from one group vary more than other from their respective group means. By comparing the posterior draws, we calculated posterior distribution of the mean group differences. We compared these two models with the Deviance Information Criteria (DIC; Gelman et al., 2013) which was calculated by R2jags. These information criteria helped us to identify the best fitting model based on the conditional log-likelihood.

### 3.2.1 Model priors

The embedded data follows a normal distribution

$$y_{id} \sim N(\mu_d, \varepsilon^2) \tag{3.1}$$

where $y_{id}$ is the delay-embedded matrix, with a mean $\mu_d$ for each column $d$ of the matrix. This mean is the predicted value from Equation 2.1, and $\varepsilon^2$ is the residual variance. The prior for the inverse of $\varepsilon^2$ was

$$\frac{1}{\varepsilon^2} \sim \gamma(1, 0.5) \tag{3.2}$$

The zeroth and first derivatives will follow a multivariate normal distribution, whereas the second derivative follows a normal distribution, such as

$$(x_i, \dot{x}) \sim MVN(\mu_x, \Sigma_x) \tag{3.3}$$

where $\mu_x$ is the vector or derivatives means fixed to 0, and $\Sigma_x$ is the covariance matrix. The prior for $\Sigma_x$ will be

$$\Sigma_x \sim Inverse - Wishart(I, 3) \tag{3.4}$$

where $I$ is an identity matrix of the same dimensions as $\Sigma_x$.

The subject $\eta$ and $\zeta$ follow the following structure

$$\eta_j \sim N(\mu_\eta, \sigma_\eta^2) \tag{3.5}$$

$$\zeta_j \sim N(\mu_\zeta, \sigma_\zeta^2) \tag{3.6}$$

where *mu* represents the group mean, and $\sigma^2$ represents the variance of the subjects from the group mean. $\mu_\eta$ and $\mu_\zeta$ will have the following priors

$$\mu_\eta \sim N(-0.1, 5) \tag{3.7}$$

$$\mu_\zeta \sim N(-0.1, 5) \tag{3.8}$$

While the variances $\sigma_\eta^2$ and $\sigma_\zeta^2$ have the following priors

$$\sigma_\eta^2 \sim U(0, 1) \tag{3.9}$$

$$\sigma_\zeta^2 \sim U(0, 1) \tag{3.10}$$

The subject equilibrium and equilibrium slope have the following structure

$$b_{is} \sim MVN(\mu_{bibs}, \Sigma_{bibs}) \tag{3.11}$$

where $\mu_{bibs}$ is the vector of group means, and $\Sigma_{bibs}$ is there covariance matrix. $\mu_{bibs}$ priors are

$$\mu_{bi} \sim N(217, 100000) \tag{3.12}$$

$$\mu_{bs} \sim N(0, 1000) \tag{3.13}$$

for the prior of the equilibrium we use the average number of seconds for every 10 minute epochs spent standing aver the whole sample. While the prior for $\Sigma_{bibs}$ is

$$\Sigma_{bibs} \sim Inverse-Wishart(I, 3) \tag{3.14}$$

where $I$ is an identity matrix of the same dimensions as $\Sigma_{bibs}$.

## 3.3 Results

To identify the proper embedding dimension ($d$), the DLO model was estimated for the sample with different $d$. Modeling it with mean center and detrended time-series, given the starting value sensitivity of OpenMx the models were allowed to run up to 100 times with different starting values until they reach convergence. The starting values were drawn from random distributions. $\eta$ stabilized at $\eta = -0.12$ on $d = 7$ were increasing $d$ meant a change in $\eta$ lower than 0.01. For the implementation of BMNLDE, $d = 7$ was selected. Following the Nyquist limit for $d = 7$ and 10 minute epochs, means that we expect to find oscillations of at least 2.3 hours.

### 3.3.1 Overall sample model

The overall sample model converged properly, this model was estimated twice, the first one was a complete model as describe in the previous sections. The second model estimated the zeroth ($x$) and first ($\dot{x}$) derivatives as univariate normal instead of multivariate normal. The second model was estimated because the multiple group model could not be estimate with this multivariate distribution. It was found that the problem was specifically in the AD group. Comparing the DIC for these two models, the first model shows better fit ($DIC_{dif} = 1398.8$). When comparing the parameter estimates between the two models, the parameters are equivalent. Since the parameters are equivalent, the correlation between the derivatives was 0 ($\rho x = 0.000002, 95\% CI = -0.014, 0.013$), and

Table 3.2: Overall sample BMNLDE results

| parameter | mean | 2.5% | 97.5% |
|-----------|------|------|-------|
| $\mu_\eta$ | -0.137 | -0.163 | -0.111 |
| $\mu_\zeta$ | -0.001 | -0.028 | 0.026 |
| $\sigma_\eta$ | 0.121 | 0.103 | 0.141 |
| $\sigma_\zeta$ | 0.118 | 0.101 | 0.137 |
| $\mu_{bi}$ | 220.154 | 200.791 | 240.326 |
| $\mu_{bs}$ | -0.002 | -0.007 | 0.004 |
| $\sigma_{bi}$ | 93.546 | 79.947 | 109.069 |
| $\sigma_{bs}$ | 0.025 | 0.021 | 0.029 |
| $\rho_b$ | -0.508 | -0.655 | -0.334 |
| $\sigma_x$ | 164.363 | 162.841 | 165.840 |
| $\sigma_{\dot{x}}$ | 41.840 | 41.406 | 42.271 |
| $\sigma_{\ddot{x}}$ | 25.553 | 24.864 | 25.569 |
| $\sigma_\varepsilon$ | 145.532 | 144.992 | 146.064 |
| $\lambda$ | 2.828 | 2.594 | 3.144 |

in order to have comparable results between the overall sample and multiple group the second model is preferred. The results for the overall sample model are shown in Table 3.2.

With these results we can see that the model has no damping ($\mu_\zeta = -0.001$), and has basically an equilibrium slope of 0 ($\mu_{bs} = -0.002$). The equilibrium for their average physical activity is below the 50% of the possible time ($\mu_{bi} = 220.154$), with the chosen 10 minute epochs the maximum seconds of standing for each epoch is 600. $\lambda$ indicates the periodicity of the oscillation, indicating that on average the period of the oscillation is 2.828 hours. The equilibrium intercept and slope have a negative correlation, indicating that subjects with a higher equilibrium have a more negative equilibrium slope. This means that the subjects that have a higher mean level of physical activity tend to decrease their average level with time. The $\sigma$ of the random parameters ($\sigma_\eta$, $\sigma_\zeta$, $\sigma_{bi}$, $\sigma_{bs}$) represent the inter-individual variability around the group mean. The derivatives are represented by their variability ($\sigma_x$, $\sigma_{\dot{x}}$, $\sigma_{\ddot{x}}$), for the second derivative ($\ddot{x}$) it represents the residual variability. Finally, $\sigma_\varepsilon$ represents the residual $\sigma$ for the embedded data.

### 3.3.2 Multiple group model

Once the overall sample model was established, it represents our reference model to compare the multiple group model. Now we are interested in estimating the model for both groups to compare meaningful parameters. Comparing the fit of the models with DIC, the multiple group model fit the data better ($DIC_{dif} = 5733.03$). This indicates that accounting for the different oscillating characteristics for each group is a better way to represent the data. The results for the multiple group model are presented in Table 3.3.

Results show that frequency of the periodicity of both groups is similar, $AD_\lambda = 2.78$ and $HOA_\lambda = 2.88$ hours. Neither group presents damping or equilibrium slopes. The AD groups presents higher $\eta$ which lead to the slightly faster frequency of the cycle, while the Healthy group presents higher equilibrium. AD subjects have higher variability for their $\eta$ and $\zeta$ which means that for these parameters the AD subjects have wider variation between each other than the Healthy group, and the Healthy subjects present higher variability for their equilibrium. These are the trends between the groups, but we cannot say that these differences are different from 0.

To characterize the group differences, we built the posterior distribution of the differences by subtracting the groups values for each of the saved iterations (6000). This allowed us to make direct inference about the differences of interest. The group differences are presented in Table 3.4. Table 3.4 shows both mean and variance differences for the random parameters. Using the 95% Credible Intervals we can see that every parameter difference includes 0, which indicates that the groups differences are not distinguishable from 0. The model did not find meaningful differences between groups. To see in greater detail, we calculated the cohen's $d$ as an effect size of mean differences for the random parameters. Cohen's $d$ was again calculated for each saved iteration of the posterior distributions, allowing us to estimate Credible Intervals for it. In Table 3.5 we can see that the biggest cohen' $d$ is for the equilibrium slope difference, this is not a meaningful result since it is an artifact of the low variability ($\sigma_{bs}$) of the slopes (Table 3.3), since the cohen' $d$ is a standardized mean difference in function of the standard deviation low variability lead to a cohen' $d$ that looks large in function of the low $\sigma_{bs}$. For $\eta$ and $\zeta$ the effect size is small, basically 0.

Table 3.3: Multiple-group BMNLDE results

| parameter | mean | 2.5% | 97.5% |
|---|---|---|---|
| AD group | | | |
| $\mu_\eta$ | -0.144 | -0.201 | -0.087 |
| $\mu_\zeta$ | -0.001 | -0.058 | 0.058 |
| $\sigma_\eta$ | 0.177 | 0.141 | 0.221 |
| $\sigma_\zeta$ | 0.173 | 0.138 | 0.218 |
| $\mu_{bi}$ | 211.301 | 184.029 | 238.987 |
| $\mu_{bs}$ | -0.006 | -0.013 | 0.002 |
| $\sigma_{bi}$ | 84.676 | 66.991 | 106.859 |
| $\sigma_{bs}$ | 0.023 | 0.018 | 0.029 |
| $\rho_b$ | -0.546 | -0.743 | -0.288 |
| $\sigma_x$ | 162.085 | 159.815 | 164.351 |
| $\sigma_{\dot{x}}$ | 41.593 | 40.939 | 442.246 |
| $\sigma_{\ddot{x}}$ | 24.640 | 24.061 | 25.211 |
| $\sigma_\varepsilon$ | 146.229 | 145.411 | 147.038 |
| $\lambda$ | 2.758 | 2.339 | 3.571 |
| Healthy group | | | |
| $\mu_\eta$ | -0.132 | -0.176 | -0.089 |
| $\mu_\zeta$ | 0.000 | -0.046 | 0.045 |
| $\sigma_\eta$ | 0.153 | 0.125 | 0.188 |
| $\sigma_\zeta$ | 0.153 | 0.125 | 0.188 |
| $\mu_{bi}$ | 226.614 | 198.228 | 255.478 |
| $\mu_{bs}$ | 0.001 | -0.006 | 0.009 |
| $\sigma_{bi}$ | 100.222 | 81.936 | 122.802 |
| $\sigma_{bs}$ | 0.027 | 0.022 | 0.033 |
| $\rho_b$ | -0.504 | -0.688 | -0.273 |
| $\sigma_x$ | 166.075 | 164.079 | 168.038 |
| $\sigma_{\dot{x}}$ | 42.008 | 41.430 | 42.575 |
| $\sigma_{\ddot{x}}$ | 25.660 | 25.194 | 26.140 |
| $\sigma_\varepsilon$ | 145.022 | 144.356 | 145.714 |
| $\lambda$ | 2.880 | 2.498 | 3.518 |

Table 3.4: Group differences for the random parameters

| parameter | mean | 2.5% | 97.5% |
|---|---|---|---|
| $\mu_\eta$ | -0.012 | -0.084 | 0.060 |
| $\mu_\zeta$ | -0.001 | -0.074 | 0.075 |
| $\sigma_\eta$ | 0.024 | -0.026 | 0.078 |
| $\sigma_\zeta$ | 0.020 | -0.029 | 0.074 |
| $\mu_{bi}$ | -15.313 | -54.314 | 23.106 |
| $\mu_{bs}$ | -0.007 | -0.018 | 0.004 |
| $\sigma_{bi}$ | -15.625 | -43.966 | 12.527 |
| $\sigma_{bs}$ | -0.004 | -0.012 | 0.004 |

Table 3.5: Cohen's $d$ for the mean group differences of the random parameters

| parameter | mean | 2.5% | 97.5% |
|---|---|---|---|
| $\mu_\eta$ | -0.074 | -0.505 | 0.363 |
| $\mu_\zeta$ | -0.007 | -0.452 | 0.461 |
| $\mu_{bi}$ | -0.165 | -0.584 | 0.243 |
| $\mu_{bs}$ | -0.273 | -0.699 | 0.150 |

And for the equilibrium, the effect size is considered small, the equilibrium means are in average $-0.165$ standard deviations apart.

## 3.4 Discussion

The BMNLDE was applied to sedentary data from older adults to describe their oscillating behavior and compare groups of healthy and demented (Alzheimer's) older adults. Sedentary data was collected from the activPAL monitor, which provides information every 15 seconds. The use of data from this kind of monitor creates new challenges to analyze it. The BMNLDE is presented as a method to analyze this detailed data from a continuous-time perspective with mixed-effects. This provides a model that can answer questions about the intra and interindividual variability. The intraindividual variability is usually not accounted for in the analysis of this data, were the focus has been on summary statistics instead of looking at the behaviors across time. The interindividual variability allows to estimate the the deviations of the subjects from the group mean. The BMNLDE model is flexible and powerful to account for intra and interindividual variability, and

group comparisons (mean, standard deviations, and correlations).

Due to differences between older adults with and without AD, a multiple group model was a better fit to the data, where we can identify the oscillating characteristics for each group. The BMNLDE estimates an oscillation on the sedentary behavior with a period of 2.76 and 2.89 hours for the AD and Healthy groups respectively, meaning that every 2.8 hours the cycle of physical activity starts repeating, the oscillation period is not meaningfully different between groups. The process has no damping for either group, meaning that the oscillation does not decrease amplitude to approach the equilibrium. The equilibrium for both groups is under the median of possible standing time for each 10 minute epoch, meaning that their average level of activity is closer to being sedentary rather than longer periods of time standing. The model allows us to estimate the inter-individual variability, thus we can look at how much subjects vary from their group mean. When we compared the two groups, there are no group differences different from 0, either in group means or standard deviations. Indicating that the random parameters that describe the oscillating process do not differ between the Healthy and AD subjects. The average oscillation can be equated, also the variation from the respective group means did not differ between groups. Thus group mean parameters and inter-individual variation within the groups can be assumed to be equal.

The lack of differences in physical activity between healthy and AD subjects is not inconsistent with the definition of cognitive impairment, given that in early stages of Alzheimer's disease physical function typically remains intact despite changes in cognitive function. These subjects would be classified with mild or early stage Alzheimer's (Petersen, 2003; Selkoe & Schenk, 2003; Jicha & Carr, 2010). AD should be seen as an illness that progresses in a continuum between *normal* and *demented* (Selkoe & Schenk, 2003). The first diagnosis is guided by cognitive symptoms which in its early stages it is characterized by memory complaints, objective memory impairment, normal general cognition, and preserved activities of daily living (Petersen, 2003; Petersen et al., 2009). This final characteristic, that the subjects are still able have functional and physical independence means that the brain damage hasn't affected the motor skills. The motor skills of subjects who suffer Alzheimer's are likely to remain unaffected in this early stage, thus there may be no observ-

able differences in their physical activity. It is possible that we would see greater differences if our participants were in a more advanced stage of the disease where the motor skills have been affected by it's progression. These results corroborate the definition of early stages of AD, showing how subjects with memory problems and cognitive decline do not show change in physical activity in comparison with healthy older adults (Petersen, 2003).

Future research could apply the BMNLDE to compare different groups, such as a comparison between men and women, as Watts et al. (2016) have found differences between these groups. Also, the different random parameters could be treated as second level outcomes and be predicted by different variables. The random parameters could be related to cognitive performance, social characteristics, or emotional variables.

# References

Aguilar-Farias, N., Brown, W. J., Olds, T. S., & Peeters, G. G. (2015). Validity of self-report methods for measuring sedentary behaviour in older adults. *Journal of Science and Medicine in Sport*, 18(6), 662–666.

Ait-Sahalia, Y. (2009). Estimating and Testing Continuous-Time Models in Finance: The Role of Transition Densities. *Annual Review of Financial Economics*, 1(1), 341–359.

Andersen, K. E. & Hojbjerre, M. (2005). A population-based Bayesian approach to the minimal model of glucose and insulin homeostasis. *Statistics in Medicine*, 24(15), 2381–2400.

Baker, J., Meisner, B. A., Logan, A. J., Kungl, A.-M., & Weir, P. (2009). Physical activity and successful aging in Canadian older adults. *Journal of Aging and Physical Activity*, 17(2), 223–235.

Bertenthal, B. I. (2007). Dynamical Systems: It is about time! In S. M. Boker & M. J. Wenger (Eds.), *Data Analytic Techniques for Dynamical Systems*, Notre Dame Series on Quantitative Methodology. Lawrence Erlbaum Associates, Inc.

Boker, S., Neale, M., & Rausch, J. (2004). Latent differential equation modeling with multivariate multi-occasion indicators. In *Recent developments on structural equation models* (pp. 151–174). Springer.

Boker, S. M. (2007). Specifying Latent Differential Equations Models. In S. M. Boker & M. J.

Wenger (Eds.), *Data Analytic Techniques for Dynamical Systems*, Notre Dame Series on Quantitative Methodology. Lawrence Erlbaum Associates, Inc.

Boker, S. M., Molenaar, P. C. M., & Nesselroade, J. R. (2009). Issues in intraindividual variability: Individual differences in equilibria and dynamics over multiple time scales. *Psychology and Aging*, 24(4), 858–862.

Boker, S. M. & Nesselroade, J. R. (2002). A Method for Modeling the Intrinsic Dynamics of Intraindividual Variability: Recovering the Parameters of Simulated Oscillators in Multi-Wave Panel Data. *Multivariate Behavioral Research*, 37(1), 127–160.

Boker, S. M., Staples, A. D., & Hu, Y. (2016). Dynamics of Change and Change in Dynamics. *Journal for Person-Oriented Research*, 2(1-2), 34–55.

Brooks, S. & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434–455.

Brown, C. (2007). *Differential Equations: A modeling Approach*. Number 150 in Quantitative Applications in the Social Sciences. Sage Publications.

Butner, J., Amazeen, P. G., & Mulvey, G. M. (2005). Multilevel Modeling of Two Cyclical Processes: Extending Differential Structural Equation Modeling to Nonlinear Coupled Systems. *Psychological Methods*, 10(2), 159–177.

Cavanaugh, J. T., Kochi, N., & Stergiou, N. (2010). Nonlinear Analysis of Ambulatory Activity Patterns in Community-Dwelling Older Adults. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 65A(2), 197–203.

Chastin, S. & Granat, M. (2010). Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. *Gait & Posture*, 31(1), 82–86.

Chastin, S. F., Baker, K., Jones, D., Burn, D., Granat, M. H., & Rochester, L. (2010). The pattern of

habitual sedentary behavior is different in advanced Parkinson's disease. *Movement Disorders*, 25(13), 2114–2120.

Colcombe, S. & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults a meta-analytic study. *Psychological science*, 14(2), 125–130.

Deboeck, P. R., Boker, S. M., & Bergeman, C. S. (2008). Modeling Individual Damped Linear Oscillator Processes with Differential Equations: Using Surrogate Data Analysis to Estimate the Smoothing Parameter. *Multivariate Behavioral Research*, 43(4), 497–523.

Deboeck, P. R., Montpetit, M. A., Bergeman, C. S., & Boker, S. M. (2009). Using derivative estimates to describe intraindividual variability at multiple time scales. *Psychological Methods*, 14(4), 367–386.

DiPietro, L. (2001). Physical activity in aging changes in patterns and their relationship to health and function. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(suppl 2), 13–22.

Donnet, S., Foulley, J.-L., & Samson, A. (2010). Bayesian Analysis of Growth Curves Using Mixed Models Defined by Stochastic Differential Equations. *Biometrics*, 66(3), 733–741.

Duan, J. A., Gelfand, A. E., & Sirmans, C. (2009). Modeling space-time data using stochastic differential equations. *Bayesian Analysis*, 4, 733–758.

Efron, B. (1986). Why Isn't Everyone a Bayesian? *The American Statistician*, 40(1), 1.

Eggermont, L. H. P., Milberg, W. P., Lipsitz, L. A., Scherder, E. J. A., & Leveille, S. G. (2009). Physical Activity and Executive Function in Aging: The MOBILIZE Boston Study: PHYSICAL ACTIVITY AND EXECUTIVE FUNCTION IN AGING. *Journal of the American Geriatrics Society*, 57(10), 1750–1756.

Gajewski, P. D. & Falkenstein, M. (2015a). Lifelong physical activity and executive functions in older age assessed by memory based task switching. *Neuropsychologia*, 73, 195–207.

Gajewski, P. D. & Falkenstein, M. (2015b). Long-term habitual physical activity is associated with lower distractibility in a Stroop interference task in aging: Behavioral and ERP evidence. *Brain and Cognition*, 98, 87–101.

Gander, M. P. & Stephens, D. A. (2007). Stochastic volatility modelling in continuous time with general marginal distributions: Inference, prediction and model selection. *Journal of Statistical Planning and Inference*, 137(10), 3068–3081.

Gandy, S. & DeKosky, S. T. (2013). Toward the Treatment and Prevention of Alzheimer's Disease: Rational Strategies and Recent Progress. *Annual Review of Medicine*, 64(1), 367–383.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis*. Texts in Statistical Science. Boca Raton, FL: Chapman & Hall/CRC, third edition.

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*, volume Analytical methods for social research. New York: Cambridge University Press.

Gelman, A. & others (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, 1(3), 515–534.

Gelman, A. & Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.

Godfrey, A., Lord, S., Galna, B., Mathers, J. C., Burn, D. J., & Rochester, L. (2014). The association between retirement and age on physical activity in older adults. *Age and Ageing*, 43(3), 386–393.

Grant, P. M., Dall, P. M., Mitchell, S. L., & Granat, M. H. (2008). Activity monitor accuracy in measuring step number and cadence in community-dwelling older adults. *Journal of aging and physical activity*, 16(2), 201.

Gretebeck, R. J. & Montoye, H. J. (1992). Variability of some objective measures of physical activity. *Medicine and Science in Sports and Exercise*, 24, 1167–1172.

Guo, G. & Zhao, H. (2000). Multilevel modeling for binary data. *Annual review of sociology*, (pp. 441–462).

Higham, D. J. (2011). Stochastic ordinary differential equations in applied and computational mathematics. *IMA Journal of Applied Mathematics*, 76(3), 449–474.

Ho, A. J., Raji, C. A., Becker, J. T., Lopez, O. L., Kuller, L. H., Hua, X., Dinov, I. D., Stein, J. L., Rosano, C., Toga, A. W., & Thompson, P. M. (2011). The effects of physical activity, education, and body mass index on the aging brain. *Human Brain Mapping*, 32(9), 1371–1382.

Hoffman, M. D. & Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(Apr), 1593–1623.

Hong, H., Carlin, B. P., Shamliyan, T. A., Wyman, J. F., Ramakrishnan, R., Sainfort, F., & Kane, R. L. (2013). Comparing Bayesian and Frequentist Approaches for Multiple Outcome Mixed Treatment Comparisons. *Medical Decision Making*, 33(5), 702–714.

Hox, J. J. (2010). *Multilevel Analysis: Techniques and Applications*. Quantitative Methodology Seriers. Routledge, second edition.

Hu, Y. (2013). *Dynamic Process for Ordinal Data Simulation and Application*. University of Virginia.

Hu, Y., Boker, S., Neale, M., & Klump, K. L. (2014). Coupled latent differential equation with moderators: Simulation and application. *Psychological Methods*, 19(1), 56–71.

Huang, Y., Wu, H., & Acosta, E. P. (2010). Hierarchical Bayesian inference for HIV dynamic differential equation models incorporating multiple treatment factors. *Biometrical Journal*, 52(4), 470–486.

Israel, V. P. & Migon, H. S. (2012). Stochastic models for greenhouse gas emission rate estimation from hydroelectric reservoirs: a Bayesian hierarchical approach. *Journal of Applied Statistics*, 39(5), 1069–1086.

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. West Sussex, PO19 8SQ, United Kingdom: John Wiley and Sons.

Jaeger, J. & Lambert, P. (2013). Bayesian P-spline estimation in hierarchical models specified by systems of affine differential equations. *Statistical Modelling*, 13(1), 3–40.

Janssen, X. & Cliff, D. P. (2015). Issues Related to Measuring and Interpreting Objectively Measured Sedentary Behavior Data. *Measurement in Physical Education and Exercise Science*, 19(3), 116–124.

Jicha, G. A. & Carr, S. A. (2010). Conceptual evolution in Alzheimer's disease: implications for understanding the clinical phenotype of progressive neurodegenerative disease. *Journal of Alzheimer's Disease*, 19(1), 253–272.

Kang, M. & Rowe, D. A. (2015). Issues and Challenges in Sedentary Behavior Measurement. *Measurement in Physical Education and Exercise Science*, 19(3), 105–115.

Kaplan, D. (2014). *Bayesian Statistics for the Social Sciences*. Methodology in the Social Sciences. New York, NY: The Guilfors Press.

Kim, Y., Barry, V. W., & Kang, M. (2015). Validation of the ActiGraph GT3x and activPAL Accelerometers for the Assessment of Sedentary Behavior. *Measurement in Physical Education and Exercise Science*, 19(3), 125–137.

Klim, S., Mortensen, S. B., Kristensen, N. R., Overgaard, R. V., & Madsen, H. (2009). Population stochastic modelling (PSM)-An R package for mixed-effects models based on stochastic differential equations. *Computer Methods and Programs in Biomedicine*, 94(3), 279–289.

Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, 14(7), 293–300.

Little, T. D. (2013). *Longitudinal Structural Equation Modeling*. Methodology in the Social Sciences. New York, NY: The Guilfors Press.

Matthews, R. A. (2001). Why should clinicians care about Bayesian methods? *Journal of statistical planning and inference*, 94(1), 43–58.

Maxwell, S. E. & Boker, S. M. (2007). Multilevel Models of Dynamical Systems. In S. M. Boker & M. J. Wenger (Eds.), *Data Analytic Techniques for Dynamical Systems*, Notre Dame Series on Quantitative Methodology. Lawrence Erlbaum Associates, Inc.

Maxwell, S. E. & Delaney, H. D. (2003). *Designing experiments and analyzing data: A model comparison perspective*. Mahwah, N.J: Lawrence Erlbaum.

McArdle, J. J. & Nesselroade, J. R. (2014). *Longitudinal Data Analysis Using Structural Equation Models*. Washington, DC: American Psychological Association.

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press.

Merkle, E. & Wang, T. (2014). Bayesian Factor Models for the Analysis of Experimental Data. *Available at SSRN 2536989*.

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, 2(4), 201–218.

Murray, L. (2009). Bayesian learning of continuous time dynamical systems with applications in functional magnetic resonance imaging.

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kickpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, 81(2), 535–549.

Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods*, 16(4), 468–490.

Oud, J. H. L. (2007). Continuous Time Modeling of Reciprocal Relationships in the Cross-Lagged Panel Design. In S. M. Boker & M. J. Wenger (Eds.), *Data Analytic Techniques for Dynamical Systems*, Notre Dame Series on Quantitative Methodology. Lawrence Erlbaum Associates, Inc.

Parslow, J., Cressie, N., Campbell, E. P., Jones, E., & Murray, L. (2013). Bayesian learning and predictability in a stochastic nonlinear dynamical model. *Ecological applications*, 23(4), 679–698.

Petersen, R. C. (2003). Mild cognitive impairment clinical trials. *Nature Reviews Drug Discovery*, 2(8), 646–653.

Petersen, R. C., Roberts, R. O., Knopman, D. S., Boeve, B. F., Geda, Y. E., Ivnik, R. J., Smith, G. E., & Jack, C. R. (2009). Mild Cognitive Impairment: Ten Years Later. *Archives of Neurology*, 66(12).

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2015). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-120.

Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-Effects Models in S an S-Plus*. Statistics and Computing. Springer New York.

Plummer, M. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling.

Prague, M., Commenges, D., Guedj, J., Drylewicz, J., & Thiebaut, R. (2013). NIMROD: A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. *Computer Methods and Programs in Biomedicine*, 111(2), 447–458.

Proper, K. I., Singh, A. S., van Mechelen, W., & Chinapaw, M. J. (2011). Sedentary Behaviors and Health Outcomes Among Adults. *American Journal of Preventive Medicine*, 40(2), 174–182.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reid, N., Eakin, E., Henwood, T., Keogh, J., Senior, H., Gardiner, P., Winkler, E., & Healy, G. (2013). Objectively Measured Activity Patterns among Adults in Residential Aged Care. *International Journal of Environmental Research and Public Health*, 10(12), 6783–6798.

Ridgers, N. D., Timperio, A., Cerin, E., & Salmon, J. (2015). Within- and between-day associations between children's sitting and physical activity time. *BMC Public Health*, 15(1).

Rowlands, A. V., Gomersall, S. R., Tudor-Locke, C., Bassett, D. R., Kang, M., Fraysse, F., Ainsworth, B., & Olds, T. S. (2015). Introducing novel approaches for examining the variability of individuals' physical activity. *Journal of Sports Sciences*, 33(5), 457–466.

Selkoe, D. J. & Schenk, D. (2003). A <span style="font-variant:small-caps;">LZHEIMER'S</span> D <span style="font-variant:small-caps;">ISEASE</span>: Molecular Understanding Predicts Amyloid-Based Therapeutics. *Annual Review of Pharmacology and Toxicology*, 43(1), 545–584.

Soetaert, K., Petzoldt, T., & Setzer, R. W. (2010). Solving differential equations in r: Package desolve. *Journal of Statistical Software*, 33(9), 1–25.

Song, X.-Y. & Lee, S.-Y. (2012). *Bayesian and Advanced Bayesian Structural Equation Modeling: With Applications in the Medical and Behavioral Sciences*. Probability and Statistics. West Sussex, United Kindom: JohnWiley & Son, Ltd.

Steele, J. S. & Ferrer, E. (2011). Latent Differential Equation Modeling of Self-Regulatory and Coregulatory Affective Processes. *Multivariate Behavioral Research*, 46(6), 956–984.

Stramer, O. & Roberts, G. O. (2007). On Bayesian analysis of nonlinear continuous-time autoregression models. *Journal of Time Series Analysis*, 28(5), 744–762.

Su, Y.-S. & Yajima, M. (2015). *R2jags: Using R to Run 'JAGS'*. R package version 0.5-7.

Tiberio, S. S. (2008). *The effects of misspecified measurement intervals in Multivariate Latent Differential Equation models*. PhD thesis, University of Notre Dame.

Tomassini, L., Reichert, P., Künsch, H. R., Buser, C., Knutti, R., & Borsuk, M. E. (2009). A smoothing algorithm for estimating stochastic, continuous time model parameters and its application to a simple climate model. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 58(5), 679–704.

Tornoe, C. W., Agerso, H., Jonsson, E., Madsen, H., & Nielsen, H. A. (2004). Non-linear mixed-effects pharmacokinetic/pharmacodynamic modelling in NLME using differential equations. *Computer Methods and Programs in Biomedicine*, 76(1), 31–40.

Trost, S. G., Mciver, K. L., & Pate, R. R. (2005). Conducting Accelerometer-Based Activity Assessments in Field-Based Research:. *Medicine & Science in Sports & Exercise*, 37(Supplement), S531–S543.

van der Niet, A. G., Smith, J., Scherder, E. J., Oosterlaan, J., Hartman, E., & Visscher, C. (2015). Associations between daily physical activity and executive functioning in primary school-aged children. *Journal of Science and Medicine in Sport*, 18(6), 673–677.

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, 17(2), 176–192.

von Oertzen, T. & Boker, S. M. (2010). Time delay embedding increases estimation precision of models of intraindividual variability. *Psychometrika*, 75(1), 158–175.

Wagenmakers, E.-J., Lee, M., Lodewyckx, T., & Iverson, G. J. (2008). Bayesian versus frequentist inference. In *Bayesian evaluation of informative hypotheses* (pp. 181–207). Springer.

Wang, B., Li, Z., Miao, M., Liu, X., Wang, H., Liang, G., Liu, Z., Sun, Y., & Yuan, W. (2015). The Relationship Between Physical Activity and Aging Symptoms Among Community-Dwelling Men Aged 40–70 Years in Shanghai, China. *Journal of Physical Activity and Health*, 12(1), 87–92.

Wang, L., Cao, J., Ramsay, J. O., Burger, D. M., Laporte, C. J. L., & Rockstroh, J. K. (2014). Estimating mixed-effects differential equation models. *Statistics and Computing*, 24(1), 111–121.

Wang, Y., Eskridge, K. M., & Zhang, S. (2008). Semiparametric mixed-effects analysis of PK/PD models using differential equations. *Journal of Pharmacokinetics and Pharmacodynamics*, 35(4), 443–463.

Watts, A., Walters, R. W., Hoffman, L., & Templin, J. (2016). Intra-Individual Variability of Physical Activity in Older Adults With and Without Mild Alzheimer's Disease. *PLOS ONE*, 11(4), e0153898.