The fixed versus random effects debate and how it relates to centering in multilevel modeling

Ellen Hamaker
Utrecht University

Running head: FE VERSUS RE MODELS AND CENTERING IN MULTILEVEL MODELS

The fixed versus random effects debate and how it relates to centering in multilevel modeling

Ellen L. Hamaker[1] and Bengt Muthén[2]

1. Methodology and Statistics, Faculty of Social and Behavioural Sciences, Utrecht University, The Netherlands

2. Muthén and Muthén, Los Angeles, CA, USA

This manuscript was accepted for publication in *Psychological Methods*.

Abstract

In many disciplines researchers use longitudinal panel data to investigate the potentially causal relationship between two variables. However, the conventions and concerns vary widely across disciplines. Here we focus on two concerns, that is: a) the concern about random effects versus fixed effects, which is central in the (micro-)econometrics/sociology literature; and b) the concern about grand mean versus group (or person) mean centering, which is central in the multilevel literature associated with disciplines like psychology and educational sciences. We show that these two concerns are actually addressing the same underlying issue. We discuss diverse modeling methods based on either multilevel regression modeling with the data in long format, or structural equation modeling with the data in wide format, and compare these approaches with simulated data. We extend the multilevel model with random slopes and discuss the consequences of this. Subsequently, we provide guidelines on how to choose between the diverse modeling options. We illustrate the use of these guidelines with an empirical example based on intensive longitudinal data, in which we consider both a time-varying and a time-invariant covariate.

The fixed versus random effects debate and how it relates to centering in multilevel modeling

Some research questions are of such a fundamental nature that they stir up fervent discussions in many fields. A key example of such a question is: How can we establish the effect of one variable on another, when practical and/or ethical limitations make it impossible to perform a randomized experiment?

It has been widely recognized that longitudinal research—in which the same cases (e.g., individuals, families, or companies) are measured repeatedly—forms a useful alternative here. But despite the enormous popularity of longitudinal research in many fields, there has been relatively little exchange of insights on how to best analyze the data that are obtained with it. This lack of communication has led to remarkable differences in what is considered the golden standard for longitudinal data analysis in each discipline: For instance, in psychology such clustered data are oftentimes analyzed using multilevel regression analysis, whereas in other disciplines this approach is considered extremely problematic and researchers are strongly advised to use other techniques instead.

The goal of the current paper is to discuss this seeming schism between disciplines from a psychologist's perspective. In doing so, we want to aid psychological researchers who encounter the stark criticism of multilevel analysis—either in real life or in the literature—and who become alarmed that our field may be relying on inferior statistical methods. We will show that the concern about *fixed effects versus random effects models*, which has led disciplines like sociology, (micro-)economics, and political science to mostly abandon multilevel regression analysis, actually maps perfectly onto the concern of *how to center a level one predictor in multilevel modeling.* The latter dates back to Cronbach (1976), and is often discussed—but perhaps still not known well enough—in the psychological literature. While the connection between these concerns has been pointed out by others before (cf. Allison, 2005; Bell & Jones, 2015; McNeish & Kelley, 2019; Raudenbush, 2009; Snijders & Bosker, 2012; Wooldridge, 2013), many of these treatments are fragmented and superficial, and most of them target researchers from a different disciplinary background. Hence, we aim to complement these accounts with a presentation that should fit more naturally with the knowledge psychological researchers tend to have of multilevel modeling and structural equation modeling.

This paper is organized as follows. In the first two sections we elaborate on each of the two concerns mentioned above separately. In the third section we show analytically that, while these concerns tend to be presented in quite distinct ways, they actually are identical. We also provide a brief summary of other sources that have pointed out this equivalence before. In the fourth section, we perform two simulation studies to show how these models can be estimated based on multilevel regression analysis with the data in long format, and based on structural equation modeling (SEM) with data in wide format. In the fifth section, we discuss what happens when the models are extended with a random slope to accommodate individual differences in the within-person relationships. The sixth section consists of a set of guidelines for choosing between the diverse modeling approaches in practice, and in the seventh section these guidelines are applied to an empirical dataset based on intensive longitudinal measurements. We end with summarizing the main findings, and discussing remaining challenges.

Concern 1: Fixed effects versus random effects models

The longitudinal data we are focusing on in the current paper consist of repeated measures taken from a sample of cases (e.g., individuals, dyads, families, organizations, etc.). Such data are known as *panel data*, but are also sometimes referred to as *longitudinal multilevel data*. A key interest in this kind of research is to establish the effect of a variable x on a variable y. Let y_{it} be the outcome variable for person i at occasion t. A basic model, which is often proposed in this context (e.g., Allison, 2005; Bell & Jones, 2015; Wooldridge, 2002, 2013), can be expressed as

$$y_{it} = c_t + \alpha_i + \beta x_{it} + \gamma z_i + e_{it}, \tag{1}$$

where c_t is the intercept that can vary over time; α_i is a component that captures unobserved differences between individuals; x_{it} represents predictors that vary over time; z_i represents predictors that are invariant over time; and e_{it} is the residual that is assumed to be time and person specific. Typically it is assumed that these residuals are uncorrelated over time, although alternatives are possible.

For ease of presentation and to facilitate later comparisons, we simplify the model to

$$y_{it} = c + \alpha_i + \beta x_{it} + e_{it}, \tag{2}$$

that is, we assume the intercept c does not vary over time, and that there are no time-invariant predictors z_i. We will return to the topic of time-varying intercepts and time-invariant predictors in later sections.

The goal and the challenge

In the (micro-)econometrics, sociological, and political science literature, there has been ample discussion regarding the individual component α_i. The reason for including this component is that it accounts for *unobserved heterogeneity*, also referred to as the *individual effect*, *latent variable*, or *unobserved component* (Allison, 2005; Bollen & Brand, 2010; Bou & Satorra, 2017; Finkel, 1995; Ousey, Wilcox, & Fisher, 2011; Wooldridge, 2002; 2013). It forms a simple and elegant way to control for *time-invariant* omitted variables, and has therefore been considered essential when the goal is to get at the causal connection between x and y.

The discussions have centered around the relationship between α_i and x_{it}, known as the *endogeneity problem*. Using a multilevel regression approach to estimate the model in Equation 2 is based on the underlying assumption that the random intercept α_i is *uncorrelated* with the predictors x_{it}. When this assumption is violated, the predictors x are said to be endogenous, and this leads to bias in the estimation of β, which undermines the purpose of the analysis. Therefore, many researchers have warned against the use of a multilevel regression approach in this context, which they refer to as the *random effects (RE) model*, and the consensus has been that alternative modeling procedures should be preferred, which they refer to as the *fixed effects (FE) model*.[1]

Modeling methods for the RE and FE models

To estimate the RE model, one can simply use a multilevel regression approach for the model in Equation 2, or pooled ordinary least squares with robust standard errors (Wooldridge, 2013). In this model, there is a random intercept for which a mean and variance are estimated, and in addition, a slope estimate is obtained, which describes the relation between the time-varying predictor and the outcome.

To estimate the FE model, researchers can choose from a number of options. First, there is the approach based on *difference scores* (cf. Allison, 2009; Liker, Augustyniak, & Duncan, 1985; Wooldridge, 2002). The rationale behind using difference scores (i.e., $y_{it} - y_{it-1}$), is that the individual's constant α_i drops out. However, this approach requires specialized software for data that consist of more than two time points, in order to accounted for the nested structure of the data then.

Second, one can use *dummy variables* for each individual in what has been referred to as the *least squares dummy variable estimator* (cf. Bollen & Brand, 2010). Because the dummy variables are part of the set of predictors, a correlation between these individual effects and the other predictor(s) does not violate the underlying model assumptions, and it will lead to an unbiased estimate of β. The downside of this approach though is that it requires a separate α_i to be estimated for each individual (without assuming a distribution), such that the number of parameters that needs to be estimated increases with sample size (i.e., the number of cases).

Third, one can center both y and x per person using each individual's sample means, and then run a regression analysis on these within-person centered variables using *pooled ordinary least squares* (e.g., Bou & Satorra, 2017; Wooldridge, 2002). This approach requires an additional correction for the associated degrees of freedom, which is implemented in diverse software packages.

Wooldridge (2013) indicates that in practice, researchers often estimate both the RE model and the FE model, and then use the Hausman test to determine which model is more appropriate. The Hausman test is based on comparing the β estimates from both approaches, where the β from the RE model is obtained under the null hypothesis of no correlation between the intercept α_i and the predictor x_{it}, while the β from the FE model is obtained under the alternative hypothesis that there may be a correlation between the intercept and the predictor (Snijders & Bosker, 2012). When the Hausman test is significant, the FE model should be used; otherwise the RE model should be preferred (Wooldridge, 2013).

While the above methods form the conventional approaches to estimating the FE and RE models, Allison (2005) has shown how these models can also be estimated using structural equation modeling (SEM) with maximum likelihood estimation. His approach is represented in Figure 1, and is based on modeling α_i as a latent variable with the time-varying y_{it}'s as its indicators, and including the time-varying x_{it} as predictors of these indicators. The FE model is specified by allowing the latent variable α_i to be correlated with the x variables (depicted as the grey two-headed arrows in Figure 1); this model leads to an unbiased estimate of β, which is identical to the estimate obtained with traditional FE methods. In contrast, the RE model is obtained by setting the correlations between α_i and the x-variables to zero (i.e., omitting the grey two-headed arrows); this leads to bias in the estimation of β when α_i is actually correlated with x, as shown by Allison (2005).

Insert Figure 1 about here

Clearly, in this SEM approach to estimating the FE model, the α_i's have a distribution, while in the traditional ways to estimate a FE model (i.e., using difference scores, dummy variables, or centered variables), the α_i's are not bounded by a distribution. As a result, the SEM approach allows for the inclusion of time-invariant predictors of α_i in a natural way, whereas this is not possible in the traditional FE model approaches. But despite this advantage, many researchers still rely on the more traditional FE methods.

Concern 2: Separating the within and between slopes in multilevel modeling

Multilevel modeling, which is also referred to as *hierarchical modeling*, *random effects modeling*, *mixed effects modeling*, and *variance components modeling*, is very popular in psychological research. This technique can be used when the data have a multilevel structure with lower units nested in higher units. In panel data, we have repeated measures that are nested within individuals; in this case, the repeated measures are referred to as the within-person level or level 1 units, whereas the persons are referred to as the between-person level or level 2 units.

A fundamental concern that is often discussed in the multilevel literature is that the relationship between two variables may differ across levels (cf. Bolger & Laurenceau, 2013; Enders & Tofighi, 2007; Hamaker & Grasman, 2015; Hoffman, 2014; Kievit, Frankenhuis, Waldorp, & Borsboom, 2013; Kreft, de Leeuw, & Aiken, 1995; Mundlak, 1978; Neuhaus & Kalbfleisch, 1998; Nezlek, 2001; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). For instance, when we consider the relationship between physical exercise and heart rate, we will find a positive within-person relationship in that more strenuous exercise leads to a higher heart rate; but, in contrast, at the between-person level this relationship is likely to be reversed, as people who exercise more on average tend to have a lower heart rate on average due to their overall fitness (Hoffman, 2014).

Although such dramatic differences between the slopes may be exceptional and in most practical scenarios these slopes will have the same sign, the key point here is that there is not *one* relationship between x and y. This was already noted by Cronbach (1976), when he pointed out that asking how x and y are related is not the right question when data are clustered; instead, we need to clarify whether we are interested in the within-cluster or the between-cluster relationship.

Insert Figure 2 about here

In Figure 2 a graphical representation is given in which these two relationships are both positive, but nevertheless different. The ellipses represent data from separate individuals, each of them characterized by the same within-person slope $\beta^{(w)}$. The between-person slope $\beta^{(b)}$ describes the relationship between the person means on the two variables ($\mu_{y,i}$ and $\mu_{x,i}$). Here, these within-person means lie on a straight line; in reality, the person-specific ellipses would typically be scattered around the between-person slope, with the person means forming a between-person ellipse. In the example in Figure 2 we have $\beta^{(b)} > \beta^{(w)} > 0$.

The goal and the challenge

There are extensive discussions in the multilevel literature about the importance of separating the within-cluster slope from the between-cluster slope (e.g., Enders & Tofighi, 2007; Hoffman, 2014; Kreft et al., 1995; Laurenceau & Bolger, 2005; Nezlek, 2001 Raudenbush & Bryk, 2002). The key concern here is that if the model does not properly account for the fact that there may be different relationships between the predictor and the outcome at different levels, we end up with a slope that is an “uninterpretable blend” of the within-person and between-person slopes (p.139, Raudenbush & Bryk, 2002).

In the multilevel literature it has been well recognized that using the model in Equation 2 as the basis of a multilevel regression analysis will lead to a slope estimate that represents neither the within-cluster slope $\beta^{(w)}$, nor the between-cluster slope $\beta^{(b)}$. If we refer to this estimated slope as $\hat{\beta}$, it has been shown that it is a weighted sum of the estimated slopes at both levels, that is,

$$\begin{aligned}\hat{\beta} &= \psi\hat{\beta}^{(b)} + (1-\psi)\hat{\beta}^{(w)} \\ &= \hat{\beta}^{(w)} + \psi(\hat{\beta}^{(b)} - \hat{\beta}^{(w)})\end{aligned} \tag{3}$$

where ψ can be thought of as a measure indicating the relative amount of variability at the between-cluster level compared to the total variability in the data across all clusters and time points (Mundlak, 1978; Neuhaus & Kalbfleisch, 1998; Raudenbush & Bryk, 2002). Hence, the weight ψ is not the intra-class correlation, which is independent of the number of time points; instead it is a direct function of the number of time points and becomes smaller as the number of time points increases, as we will see later. The weighted sum β is represented as the dashed line in Figure 2, which lies somewhere between the within-person slope $\beta^{(w)}$ and the between-person slope $\beta^{(b)}$.

Multilevel regression methods to obtain $\beta^{(w)}$ and $\beta^{(b)}$

There are several ways to disentangle the within-person and between-person slopes in the context of multilevel modeling. Here we discuss three different approaches: The first two are based on using the within-person sample means on the predictor in a multilevel regression model, while the third approach is based on using latent within-person means in a multilevel structural equation model.

Using the within-person centered predictor $x_{it} - \bar{x}_{i\cdot}$. As explained by Raudenbush and Bryk (2002), and others as well (e.g., Bell & Jones, 2015; Hoffman, 2014; Kreft et al., 1995; Snijders & Bosker, 2012), we can use multilevel modeling to estimate the within-person slope without it being contaminated by the between-person slope. Let $\bar{x}_{i\cdot}$ represent the sample mean of individual i on the variable x, such that $x_{it} - \bar{x}_{i\cdot}$ represents the within-person centered predictor. Using the multilevel model notation from Raudenbush and Bryk (2002), we can specify the following within-person and between-person equations respectively,

$$y_{it} = \beta_{0i} + \beta_1\big(x_{it} - \bar{x}_{i\cdot}\big) + e_{it} \tag{4}$$

$$\beta_{0i} = \gamma_{00} + u_{0i}, \tag{5}$$

where β_{0i} is referred to as a random intercept. By plugging the between-person equation into the within-person equation, we get the combined expression

$$y_{it} = \gamma_{00} + u_{0i} + \beta_1\big(x_{it} - \bar{x}_{i\cdot}\big) + e_{it}. \tag{6}$$

While the latter expression is very similar to the model presented in Equation 2, there is one critical difference: In Equation 6 we use the within-person centered predictor, whereas in Equation 2 we had the uncentered, raw predictor. To understand the consequences of this, we should realize that—by definition—there are *no stable between-person differences* for the within-person centered predictor; consequently, the within-person centered predictor simply *cannot* be correlated with the random intercept (here: β_{0i}), so that the main concern from the FE-RE debate is circumvented here. Put differently, since the within-person centered predictor only contains within-person variance, the regression coefficient β_1 by definition represents the within-person slope between x and y, that is, $\beta_1 = \beta^{(w)}$.

To obtain the between-person slope, the within-person means on the predictor $\bar{x}_{i\cdot}$ can be included as a predictor of β_{0i}, the random intercept at the between-person level. Note that since the within-person centered predictor and the residual term in Equation 4 both have a within-person mean of zero (by definition), the random intercept β_{0i} actually represents the within-person mean on the outcome variable y. We can replace Equation 5 by a between-person equation in which we predict the within-person mean on the outcome from the within-person mean on the predictor, that is,

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\bar{x}_{i\cdot} + u_{0i}, \tag{7}$$

where the slope γ_{01} represents the between-person slope, that is $\gamma_{01} = \beta^{(b)}$. Plugging this between-person expression into the within-person expression given in Equation 4, and replacing γ_{01} with $\beta^{(b)}$, and β_1 with $\beta^{(w)}$, results in the combined expression

$$y_{it} = \gamma_{00} + \beta^{(b)}\bar{x}_{i\cdot} + u_{0i} + \beta^{(w)}\left(x_{it} - \bar{x}_{i\cdot}\right) + e_{it}. \tag{8}$$

This expression will prove useful for comparisons below.

Using the uncentered predictor x_{it} in combination with the sample means $\bar{x}_{i\cdot}$. An alternative approach to obtain the within-person slope without it being biased due to the between-person slope was first proposed by Mundlak (1978) and is known as the *contextual model.* In this approach, the predictor x_{it} is not centered, but the within-person means $\bar{x}_{i\cdot}$ are nevertheless included as a between-person predictor of the random intercept. This can be expressed as

$$y_{it} = \beta^*_{0i} + \beta^*_1 x_{it} + e_{it} \tag{9}$$

$$\beta^*_{0i} = \gamma^*_{00} + \gamma^*_{01}\bar{x}_{i\cdot} + u_{0i}. \tag{10}$$

By including Equation 10 in Equation 9, we get the combined expression

$$y_{it} = \gamma_{00}^* + \gamma_{01}^* \bar{x}_{i\cdot} + u_{0i} + \beta_1^* x_{it} + e_{it}. \tag{11}$$

At first, it may not be obvious why in this model β_1^* represents the within-person slope. However, as explained by Hoffman (2014), the cluster means $\bar{x}_i$ are correlated with the uncentered predictor x_{it}, and therefore they partly account for the same variance in the outcome; as a result, their regression coefficients will reflect only their unique contributions, which in the case of the uncentered predictor is the within-person effect. Hence we have $\beta_1^* = \beta^{(w)}$.

The parameter γ_{01}^* from Mundlak's model represents the difference between the within-person and between-person slopes. To show this, we first replace β_1^* with $\beta^{(w)}$, and rewrite Equation 11 as

$$y_{it} = \gamma_{00}^* + (\gamma_{01}^* + \beta^{(w)})\bar{x}_{i\cdot} + u_{0i} + \beta^{(w)}(x_{it} - \bar{x}_{i\cdot}) + e_{it}. \tag{12}$$

Comparing this expression to the model in Equation 8, it can be seen that these expressions are equivalent with $\gamma_{00}^* = \gamma_{00}$ and $(\gamma_{01}^* + \beta^{(w)}) = \beta^{(b)}$. Hence, $\gamma_{01}^* = \beta^{(b)} - \beta^{(w)}$, that is, Mundlak's γ_{01}^* represents the difference between the slopes.[2] This parameter is referred to as the *compositional* or *contextual effect* (cf. Raudenbush & Bryk, 2002), or as the *incremental between-person effect* (Hoffman, 2014). It can be interpreted as the expected difference on the outcome (e.g., heart rate) when we take two persons who differ one unit in their within-person mean (average level of exercise), and we observe them at an occasion when they have the same score on the predictor (current amount of physical exercise). Put simply, it is the effect of being a different person (see also Raudenbush & Bryk, 2002, p. 141).

Using the latent mean $\mu_{x,i}$ rather than the sample mean $\bar{x}_i$. The multilevel approaches discussed above are based on using the sample mean per person on the predictor, denoted as $\bar{x}_i$. However, Lüdtke et al. (2008) have shown that using the sample means tends to result in bias in the estimation of the between-person slope, especially when the lower level sample size (here: the number of repeated measures) is small. To avoid this bias, we should use latent mean centering, which is based on assuming the individual means on the predictor also come from a distribution, and that the sample mean is not an unbiased estimate of this (Lüdtke et al., 2008). This latent mean centering can be done using software that allows us to combine multilevel modeling with SEM, but—as we will see later—it can also be done in a strictly SEM context.

Let $\mu_{x,i}$ and $\mu_{y,i}$ represent the latent (or true) means of individual i on variables x and y respectively, that is $\mu_{x,i} = E_t[x_{it}]$ and $\mu_{y,i} = E_t[y_{it}]$, where the expectation is taken over time. Then the multilevel model can be expressed as

$$y_{it} = \mu_{y,i} + \beta^{(w)}\left(x_{it} - \mu_{x,i}\right) + e_{it}. \tag{13}$$

$$\mu_{y,i} = \gamma_{00} + \beta^{(b)}\mu_{x,i} + u_{0i}, \tag{14}$$

or in combined form as

$$y_{it} = \gamma_{00} + \beta^{(b)}\mu_{x,i} + u_{0i} + \beta^{(w)}\left(x_{it} - \mu_{x,i}\right) + e_{it}. \tag{15}$$

Similarly, we can also rewrite Mundlak's contextual model using the latent mean instead of the sample mean. Starting with the combined expression given in Equation 11 for this and expressing the contextual effect as $\gamma^*_{01} = \beta^{(b)} - \beta^{(w)}$ and the within-person slope as $\beta^*_1 = \beta^{(w)}$, we get

$$y_{it} = \gamma^*_{00} + (\beta^{(b)} - \beta^{(w)})\mu_{x,i} + u_{0i} + \beta^{(w)}x_{it} + e_{it}, \tag{16}$$

which is identical to Equation 15 with $\gamma_{00} = \gamma^{*}_{00}$. We will use the expression in Equation 16 to show how centering in multilevel modeling is related to the FE versus RE concern.

Two sides of the same coin

While the two concerns presented above—that is, FE versus RE modeling, and how to disentangle the within-person and between-person slopes—may seem unrelated at first, it can be easily shown that they actually deal with the same underlying issue. Here, we first show this analytically, and then point out others who have discussed this connection as well.

Connection

When we compare the key expression from the FE-RE modeling debate (i.e., Equation 2), with the final expression of Mundlak's model (i.e., Equation 16), we see they are identical with

$$c + \alpha_i = \gamma_{00} + \left(\beta^{(b)} - \beta^{(w)}\right)\mu_{x,i} + u_{0i}. \tag{17}$$

This shows that if the model in Equation 16 (which is identical to the model defined in Equations 13 and 14) is the data generating mechanism and the truth is that there are different within- and between-person slopes, then the intercept α_i in Equation 2 will necessarily be correlated with $\mu_{x,i}$. From this it follows that α_i will be correlated with the time-varying predictor x_{it}, because x_{it} consists of the within-person mean $\mu_{x,i} = E_t[x_{it}]$ plus a temporal deviation from this. Hence, when $\beta^{(b)} \neq \beta^{(w)}$, then the concern that is raised by those who advocate a FE approach is certainly legitimate. However, it should also be noted that there are various multilevel modeling approaches that can be taken to successfully tackle this issue, as was elaborated on in the previous section.

Based on Equation 17 we can also easily identify the two specific scenarios under which α_i will *not* be correlated with x_{it}. This is the case when: a) the within-person and between-person slopes are identical (i.e., $\beta^{(b)} = \beta^{(w)}$); or b) there are no stable between-person differences on the predictor (i.e., $\mu_{x,i}$ does not differ across individuals). In those situations, the second term on the right-hand side of Equation 17 drops out, and we are left with $c = \gamma_{00}$ and $\alpha_i = u_{0i}$. This implies that the random intercept α_i is not correlated with x_{it}, and estimating β with a multilevel model based on Equation 2 will result in an unbiased estimate of the within-person slope $\beta^{(w)}$ (cf. Mundlak, 1978).

In all other scenarios, using a multilevel approach based on Equation 2 is problematic: It leads to the infamous "uninterpretable blend" of $\beta^{(w)}$ and $\beta^{(b)}$ discussed in the multilevel literature, and the biased causal effect estimate due to a violation of the assumption of uncorrelated intercept and the predictor as described in the FE-RE modeling literature. Hence, while presented in very different manners, these two concerns are at their core really identical.

Others about the connection

This connection between the FE-RE models on the one hand, and the different versions of a multilevel model on the other hand, has been noted before. However, most of the sources that deal with this connection are written from a FE-RE rather than a multilevel perspective; this makes these sources less appealing and accessible for an audience of psychologists, who tend to be familiar with multilevel modeling, but not with the FE-RE debate. Furthermore, many of these presentations are rather brief and may be easily overlooked: For instance, Wooldridge spends a little over two pages on it in a book that contains more than 800 pages.

One of the first sources that deals with this connection is the 2005 book by Allison. Allison (2005) proposes a *hybrid model* (which he now refers to as the *within-between model*), in which the predictor is cluster-mean centered. He explicitly mentions the connection with group-mean centering in the multilevel literature, but adds that "this literature has not generally made the connection to fixed effects models nor has it been recognized that group-mean centering controls for all time-invariant covariates." (p.33 Allison, 2005). Raudenbush (2009) also discusses this model from a FE-RE perspective, referring to person mean centering as *adaptive centering*.

A year earlier, Halaby also pointed out that "nothing is gained by distinguishing within- and between-unit variation" (p.520 Halaby, 2004), when the individual component α_i and the time-varying predictors x_{it} are uncorrelated. However, when they are correlated or one is not sure whether they are, his proposal is to include the cluster means on the predictors in the multilevel model (i.e., Mundlak's contextual model), referring to this as "a middle ground" between FE and RE approaches (p.530, Halaby, 2004). Wooldridge (2013) proposes what he refers to as the *correlated random effects (CRE)* approach, which is also identical to Mundlak's contextual multilevel model as presented in Equation 11. Others have argued that the Hausman test, which is widely used to choose between the FE and the RE approach, is actually a test for checking whether the within and between slopes are identical or not (e.g., Snijders & Bosker, 2012, p.57).

The most elaborate and animated treatment of the connection can be found in the recent paper of Bell and Jones (2015). They build a compelling case for multilevel modeling, arguing that, while the problem of endogeneity is very real, the point is that we should simply use the *right* multilevel model to tackle it (i.e., based on person mean centering the time-varying covariate and/or including these means as a predictor at the between-level). They emphasize that presenting this as "a 'hybrid' or 'compromise' approach between FE and RE" (p. 144) is not doing justice to the historical presence of this approach in the multilevel literature. Their paper is not well known among psychologists, but was recently discussed by McNeish and Kelley (2019); however, the latter authors continue to present the multilevel model that disentangles within-person and between-person slopes as a hybrid solution rather than as one of the original versions of multilevel modeling that it truly is.

It is humbling to realize that Mundlak already discussed the issue in great detail in his paper of 40 years ago (cf. Mundlak, 1978). Although multilevel modeling did not exist at the time, researchers were concerned about how to obtain an unbiased estimate of the within-person slope, and they were arguing at the time that a FE approach should be preferred over a RE approach, which was then based on *generalized least squares estimation*. Mundlak (1978) indicated that the whole FE versus RE debate "has been based on an imaginary difference" (p.70), and that the RE model will lead to an estimate of the within-person slope *if the model is specified correctly*. His solution—as we described above—is based on including both the uncentered predictor and the within-person means on the predictor into the model. Unfortunately, this early and highly relevant paper did not have the impact on the FE-RE literature that is should have had.

Simulations

Above, we have seen that there are roughly four ways to approach these data within a multilevel framework. Specifically, we can: 1) use the uncentered predictor to get an estimate of β; 2) center the predictor with the cluster mean to get an estimate of $\beta^{(w)}$; 3) center the predictor with the cluster mean, and use the cluster mean as a predictor for stable differences between the cluster on the outcome variable to get estimates of $\beta^{(w)}$ and $\beta^{(b)}$; and 4) use the uncentered predictor, and include the cluster mean as a predictor for the stable individual specific intercept to get estimates of $\beta^{(w)}$ and the contextual effect $\beta^{(b)} - \beta^{(w)}$. Furthermore, we have seen that Allison also developed SEM approaches that parallel the first two multilevel approaches. In this section we illustrate how each of these four approaches can be implemented in practice using either multilevel regression analysis or structural equation modeling, and compare the results that are obtained with this.

Specifically, we will present two simulation studies. The goal if the first simulation study is to introduce the different ways to model these data, and to see how these are related. The goal of the second simulation is to compare these methods with respect to the inferences based on the slope estimates.

Simulation study 1

The first simulation study is performed to introduce nine different analysis methods: Five of these methods are based on multilevel regression modeling with the data in long format, while four others are based on SEM with the data in wide format. We us a single data set to compare the numerical results obtained with these methods.[3]

Data generation. We simulated the data using R. Specifically, we created a single data set consisting of 4 time points and 5000 cases. We used a large sample size to avoid small sample artifacts (cf. Bou & Satorra, 2017). The within-person slope $\beta^{(w)}$ was set to 1, such that the within-person expression for our model is

$$y_{it} = \mu_{y,i} + 1(x_{it} - \mu_{x,i}) + e_{it}. \tag{18}$$

We set the variance of the within-person centered predictor (i.e., $var(x_{it} - \mu_{x,i})$) to 1, and the variance of the residual (i.e., $var(e_{it})$) also to 1. For the between-person expression we set the grand intercept γ_{00} equal to zero, and the between-person slope $\beta^{(b)}$ equal to 2, so we have

$$\mu_{y,i} = 2\mu_{x,i} + u_{0i}. \tag{19}$$

We set the variance of the person means on the predictor (i.e., $var(\mu_{x,i})$ to 4, and the residual (i.e., $var(u_{0i})$) to 1. Hence, the between-person variability on the predictor is larger than the within-person variability on the predictor (i.e., the ratio of predictor variances is 4:1), and the between-person slope is also larger than the within-person slope (i.e., the ratio of slopes is 2:1).

Nine analysis methods. As discussed by Bou and Satorra (2017), longitudinal data can be organized in two different ways. In the *univariate* approach, which is based on data in *long-format*, the repeated measures of a single case are represented by separate rows that are identified to belong to the same case by a clustering variable. In contrast, in the *multivariate* approach, which is based on the data in *wide-format*, each repeated measure is a separate variable (represented by a different column), and the rows represent independent cases. Bou and Satorra (2017) compared several FE and RE methods from each of these two categories. The methods we use here partly overlap with the methods they used. In total, we consider nine different methods, which are summarized in the left part of Table 1.

In the category of univariate/long-format approaches, we make use of five distinct multilevel models. Method L1 is based on the raw score x_{it} (i.e., no centering), which is identical to the univariate method for the RE model used by Bou and Satorra (2017), and which is known to result in a slope β that forms a mix of the within-person and between-person slopes (Bell & Jones, 2015; Enders & Tofighi, 2007; Kreft et al., 1995; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Method L2 uses cluster-mean centering of the predictor, and is arguably the gold standard in the multilevel literature at this point. It is known to result in an estimate of the within-person slope (Bell & Jones, 2015; Enders & Tofighi, 2007; Kreft et al., 1995; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Methods L3a and L3b are both based on cluster-mean centering of the predictor and adding the cluster mean at the between-person level as a predictor of the random intercept. This corresponds to the hybrid model proposed by Allison (2005). It results in the estimation of both the within-person slope and the between-person slope (Bell & Jones, 2015; Kreft et al., 1995; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012), and is an approach that is also often used in multilevel modeling. The difference between the two versions considered here is that L3a is based on using the sample mean, while L3b is based on latent centering (Lüdtke et al., 2008). Finally, method L4 is Mundlak's contextual model, which is what Wooldridge referred to as the CRE model (Wooldridge, 2013). It is based on the raw score predictor x_{it}, with the cluster means included as a predictor of the random intercept. This model is known to lead to estimates of the within-person slope $\beta^{(w)}$ and the contextual effect

$\gamma^*_{01} = \beta^{(b)} - \beta^{(w)}$ (Bell & Jones, 2015; Kreft et al., 1995; Mundlak, 1978; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012).

For the multivariate/wide-format methods, we make use of these same two SEM models used by Bou and Satorra (2017), which were proposed by Allison (2005; see also Bollen & Brand, 2010), and which are represented in Figure 1. In both models, the random intercept α_i is modeled as a latent variable with the y's as its indicators. Method W1 is a model in which this random intercept is not allowed to correlate with the x-variables; this is thus referred to as a RE model, and is likely to result in a biased estimate of the effect of x on y. Method W2 is a model that allows for a correlation between the random intercept α_i and the x-variables; hence, it is referred to as a FE approach, and it is known to lead to an unbiased estimate of the within-person slope $\beta^{(w)}$ (Allison, 2005; Bollen & Brand, 2010; Bou & Satorra, 2017).

We consider two alternative multivariate/wide-format SEM methods, which are represented in Figure 3. Method W3 is closely related to the SEM model proposed by Curran, Lee, Howard, Lane, and MacCallum (2012).[4] It is based on modeling the time-invariant parts of x_{it} and y_{it} as latent variables, $\mu_{x,i}$ and $\mu_{y,i}$ respectively, where the latter is regressed on the former, allowing us to estimate both the within-person slope $\beta^{(w)}$ and between-person slope $\beta^{(b)}$. Hence, this method is akin to method L3b. Finally, method W4 can be considered the SEM version of Mundlak's contextual model, in which y_{it} is regressed on x_{it}, while the time-invariant part of the predictor (i.e., $\mu_{x,i}$) is used as a predictor of the intercept α_i. This model thus provides estimates of the within-person slope $\beta^{(w)}$ and the contextual effect $\gamma^*_{01} = \beta^{(b)} - \beta^{(w)}$.

Insert Figure 3 about here

Results. For methods L1, L2, L3a and L4, we compared the results obtained with Mplus (Muthén & Muthén, 1998-2017), to those obtained with lmer in R (using full maximum likelihood estimation, rather than the default restricted maximum likelihood estimation; Bates, Mächler, Bolker, & Walker, 2015). The differences were ignorable (i.e., only a few differences on the third decimal of standard errors); note however that these differences can become larger when the number of cases is smaller. For method L3b we used Mplus, as it requires latent mean centering, which is not possible with lmer. For the methods in wide-format we also used Mplus, and we imposed constraints over time (as the data were also generated using these constraints); this resulted in the exact same model fit for methods W2-W4. The parameter estimates for all nine methods are presented in the right part of Table 1. The following comments can be made about this.

First, the two methods that have been identified as versions of the RE model (i.e., L1 and W1), both result in a slope estimate of 1.320; this can be thought of as a blend of the within-person slope of 1.00 and the between-person slope of 2.00. All the other methods result in a within-person slope estimate of 1.012, which will be identical to the within-person slope estimate obtained with FE methods.

Second, methods L3a, L3b, and W3 result in an estimate of the between-person slope. Method L3a is based on using the sample mean to center the predictor, and this results in a between-person slope estimate of 1.938. Methods L3b, and W3 are both based on latent centering—either in univariate/long-format or in multivariate/wide-format—and they both result in the same between-person slope estimate of 1.996. This is in line with results reported in Lüdtke et al. (2008), who showed that latent mean centering solves the bias associated with sample mean centering in multilevel models.

Third, methods L4 and W4 result in the estimation of the contextual effect, which is the difference between the between-person slope and the within-person slope. By considering the significance of this parameter, one can test whether there is a difference in the within- and between-person slope, and thus whether a FE model is necessary, or that a RE model suffices (Bell & Jones, 2015; Wooldridge, 2013). From the contextual effect and the within-person slope, we can compute the between-person slope; for method L4 this is 1.012 + 0.926 = 1.938, which is identical to the between-person slope estimate from method L3a, which is also based on using the sample mean; for method W4 it is 1.012 + 0.984 = 1.996, which is identical to the between-person slope obtained with methods based on latent mean centering (i.e., methods L3b, and W3).

In sum, the simulation presented here shows that there are four different slopes that may be estimated, that is: a) the within-person slope $\beta^{(w)}$, which can be obtained with all methods except for L1 and W1; b) the between-person slope $\beta^{(b)}$, whose point estimate depends on whether or not latent mean centering was used; c) the weighted average between these two β, which is only obtained with L1 and W1; and d) the contextual effect $\gamma^*_{01} = \beta^{(b)} - \beta^{(w)}$, which also depends on whether or not latent mean centering was used. In general, these four slopes will not be the same. It is therefore crucial researchers understand which slope they are estimating, and whether this is in fact the slope one is interested in.

Simulation study 2

The second simulation study is performed to see whether the equivalences between the diverse analyses methods also hold in more realistic sample sizes, and whether the statistical inferences are similar across the different methods. Hence, instead of using a single replication based on a large number of cases, we generated 1,000 replication based on 100 cases. We did this for both a small number of time points, $T = 4$, and with a large number of time points, $T = 40$; these two forms of data are referred to as panel data and intensive longitudinal data respectively (cf. Collins, 2006).

Data generation and analysis. In the second study, we again simulated data with R. The within-person equation was characterized by a within-person slope of 1, that is,

$$y_{it} = \mu_{y,i} + 1(x_{it} - \mu_{x,i}) + e_{it}. \tag{20}$$

We set the variance of the within-person centered predictor (i.e., $var(x_{it} - \mu_{x,i})$) to 1, and the variance of the residual (i.e., $var(e_{it})$) to 4. For the between-person expression we set the grand intercept γ_{00} equal to zero, and the between-person slope $\beta^{(b)}$ equal to 0.5, so we have

$$\mu_{y,i} = 0.5\mu_{x,i} + u_{0i}. \tag{21}$$

We set the variance of the person means on the predictor (i.e., $var(\mu_{x,i})$ to 4, and the residual (i.e., $var(u_{0i})$) to 4. As a result, the intra-class correlation is 0.8 for x and 0.5 for y. Moreover, the correlation between x and y is 0.45 at both levels.

Results. For intensive longitudinal data, it is impractical (or simply impossible) to use the wide-format approaches; therefore, we only performed the long-format approaches when $T = 40$. For $T = 4$, we applied all the possible modeling methods, but based on the results from the previous simulation, we focus primarily on the results for methods L3a, L3b and W3 here. In Table 2 the results based on 4 time points (left), and on 40 time points (right) are presented. It contains: a) the average slope estimate across 1000 replications (i.e., $\bar{\hat{\theta}}$); b) the standard deviation across 1000 replications of the point estimates (i.e., $SD(\hat{\theta})$); c) the average standard error (i.e., $\bar{SE}$), which should be close to the standard deviation of the point estimates; and d) the coverage rate of the 95% confidence intervals of the point estimates, which should be close to 0.95.

When considering the results for $T = 4$, we see that for the within-person slope, the standard error tends to be overestimated (i.e., mean standard error is larger than standard deviation of the point estimates), which results also in inflated coverage rates: These are 0.958 and 0.960 for the two long-format approaches L3a and L3b respectively, and even 0.965 for the wide-format approach W3. However, this is accompanied by superior results obtained by W3 for the estimation of the between-person slope: In comparison to method L3a, method W4 leads to an average estimate for $\beta^{(b)}$ that is closer to the true value, and in comparison to both long-format approaches L3a and L3b, method W4 leads to a coverage rate for this parameter that is close to the desired 0.95 (i.e., it is 0.947), while the long-format approaches result in lower coverage rates (i.e., 0.932 when using L3a based on the sample mean, and 0.935 when using L3b based on latent mean).

When comparing the results obtained for $T = 4$ and $T = 40$, it is noteworthy that for the within-person slope $\beta^{(w)}$, when $T = 4$ the average standard error was larger than the standard deviation of the sampling distribution, but that when $T = 40$ this has actually reversed; this explains why when $T = 4$ the coverage rate was larger than 0.95, while for $T = 40$ is has become smaller than 0.95. Moreover, while latent mean centering (L3b) is certainly preferred over sample mean centering (L3a) when $T = 4$, this is no longer the case when $T = 40$. Finally, we see that for the between-person slope $\beta^{(b)}$ both long-format approaches L3a and L3b result in an average standard error that underestimates the sampling standard deviation, and that this discrepancy increases as T increases; as a result, the coverage rates for this parameter are lower than the preferred 0.95, and are lower for $T = 40$ than for $T = 4$.

Possibilities and consequences of random slopes

So far we have assumed that there was one within-person slope $\beta^{(w)}$, which applies to all individuals. However, it is also possible that each person is characterized by their own slope $\beta_i^{(w)}$, which comes from a normal distribution with a mean (i.e., the fixed or average effect) and a variance (making it a random effect). Here we briefly present a key model that includes a random slope, and then discuss how this affects the model equivalences we presented above for models with a common slope for all cases.

Model with a random slope

The possibility to incorporate *random slopes* is identified as one of the key advantages of the multilevel approach to longitudinal data (Bell & Jones, 2015). Such models are considered especially useful in the analyses of intensive longitudinal data, which consist of many repeated measurements per person and where the focus is primarily on individual differences in the effect of a time-varying covariate on an outcome (e.g., Bolger & Laurenceau, 2013; Geschwind, Peeters, Drukker, van Os, & Wichers, 2011; Wichers et al., 2009). Intensive longitudinal data are becoming increasingly more popular in the social and behavioral sciences, as well as other fields like health research and epidemiology, where it is obtained with techniques known as *ambulatory assessments*, *ecological momentary assessment*, and *experience sampling* (Trull & Ebner-Priemer, 2013; Hamaker & Wichers, 2017). However, random slopes may also be of interest in more traditional panel research, consisting of a relatively small number of repeated measures.

The multilevel model with a random slope can be expressed as

$$y_{it} = \mu_{y,i} + \beta_i^{(w)}\left(x_{it} - \mu_{x,i}\right) + e_{it} \tag{22}$$

$$\mu_{y,i} = \gamma_{00} + \beta^{(b)}\mu_{x,i} + u_{0i} \tag{23}$$

$$\beta_i^{(w)} = \gamma_{10} + \gamma_{11}\mu_{x,i} + u_{1i} \tag{24}$$

where the third equation shows that the random slope can be a function of the mean in x. The model as expressed here can be estimated using any multilevel approach that allows for latent mean centering. Note that when x is grand mean centered, this implies that the mean of $\mu_{x,i}$ is zero, so that γ_{10} in Equation 24 represents the average within-person slope.

Suppose that x_{it} is a measure of stress (which is grand mean centered), and y_{it} is a measure of anxiety in individual i at occasion t. Then the model specified in Equations 22-24 can inform us on three critical relationships, captured by $\beta^{(b)}$, γ_{10}, and γ_{11}. First, when $\beta^{(b)} > 0$, this implies that individuals that on average report more stress also tend to experience more anxiety than individuals lower on average stress. Second, when $\gamma_{10} > 0$, this implies that for the average person a temporary elevated level of stress (i.e., $x_{it} - \mu_{x,i}$) predicts a temporary increase in anxiety. Third, when $\gamma_{11} > 0$ this implies that individuals with a higher average stress level also tend to respond more strongly to temporary increases in stress than individuals with a lower average stress level.

Consequences of a random slope for model equivalences

To extend Mundlak's model with a random slope, we can write

$$y_{it} = \mu_{y,i} + \beta_i^* x_{it} + e_{it} \tag{25}$$

$$\mu_{y,i} = \gamma_{00}^* + \gamma_{01}^* \mu_{x,i} + u_{0i} \tag{26}$$

$$\beta_i^* = \gamma_{10}^* + \gamma_{11}^* \mu_{x,i} + u_{1i}. \tag{27}$$

The question now is whether this model is still identical to the model in Equations 22-24. To this end, we need to get the combined expressions for both models, which means we plug the expressions for $\mu_{y,i}$ and $\beta_i^{(w)}$ into the level 1 expression. For simplicity, we assume that $\gamma_{11} = 0$ and $\gamma_{11}^* = 0$ here, so that the expression for the random slope simplifies to $\beta_i^{(w)} = \gamma_{10} + u_{1i}$ and $\beta_i^* = \gamma_{10}^* + u_{1i}$.

Then we can write Equation 22 as

$$y_{it} = \gamma_{00} + \left(\beta^{(b)} - \gamma_{10}\right)\mu_{x,i} + \gamma_{10} x_{it} + u_{0i} + u_{1i} x_{it} - u_{1i}\mu_{x,i} + e_{it} \tag{28}$$

and Equation 25 as

$$y_{it} = \gamma_{00}^* + \gamma_{01}^* \mu_{x,i} + \gamma_{10}^* x_{it} + u_{0i} + u_{1i}^* x_{it} + e_{it}. \tag{29}$$

In comparing Equations 28 and 29 we see that, even when $\gamma_{01}^* = \beta^{(b)} - \gamma_{10}$ and $u_{1i} = u_{1i}^*$, these two expression will not be the same: The first expression includes the term $u_{1i}\mu_{x,i}$, which is not included in the latter. This shows that, when we extend the models with a random slope, they are no longer reparametrizations of each other; instead, these models will have a different fit, and the parameters of the one cannot be expressed as functions of the parameters of the other (see also Snijders & Bosker, 2012).

Guidelines for model selection

To support the decision making process for what modeling method to use, Table 3 provides an overview of features that are present or absent in each approach. This overview is not exhaustive, but we consider these the most important ones to consider.

First of all, it is indicated in the first three columns whether the modeling method results in an estimate of the within-person slope $\beta^{(w)}$, the between-person slope $\beta^{(b)}$, and/or the contextual effect γ_{01}. Note that if we get two of these three, the third can be computed from the other two, which we have then indicated in parentheses. In general, we believe the within-person and between-person slopes are most interesting, although there may research areas where the contextual effect is the main feature of interest.

Second, we indicated whether standardized within-person slope and/or the standardized between-person slope can be obtained, as these may be of interest as measures of the effect sizes at each level. Standardized results are not commonly included in multilevel modeling, because there are diverse ways to standardize slopes. Schuurman, Ferrer, de Boer-Sonnenschein, and Hamaker (2016) argued that standardizing a within-person slope should be done with the within-person standard deviations, while between-person slopes should be standardized with between-person standard deviations. This has now been implemented in Mplus, and can be obtained for the model with a common slope using maximum likelihood estimation, and for models with a random slope using the Bayesian multilevel module (cf. Asparouhov, Hamaker, & Muthén, 2018).

In the SEM approaches it is fairly easy to obtain standardized results, but one should be careful with these for the same reason as discussed above: Depending on how the model is specified, standardization will occur using the total, within-person or between-person standard deviation. Of the methods presented, only W3 will lead to a between-person slope that is standardized using the between-person standard deviations for both x and y. However, the within-person slope in this approach is standardized with the within-person standard deviation on x and the total standard deviation on y (see also Figure 3), which is thus not a true within-person standardization. Instead, to obtain the within-person standardized slope, the y variable needs to be decomposed in the same manner that the x variable was decomposed, with a separate within-person component.

Third, we distinguish between panel data, which consist of a relatively small number of repeated measures (say $T < 10$), and intensive longitudinal data, which consist of a relatively large number of repeated measures. Both forms of data can be handled using a long-format, multilevel approach, but the wide-format SEM approach becomes impractical when the number of repeated measures increases: It is inconvenient to specify the model then, and the inversion of the associated covariance matrix may become computationally impossible.

Fourth, we may be interested in extending these models with a random slope that varies across individuals (as discussed in the previous section), or with slopes and variances that vary over time. Individual differences in the slope in general require a long-format approach. Time-varying parameters are easily realized in wide-format, simply by relaxing the constraints over time. In long-format, time-varying parameters can be realized through the inclusion of dummy variables that represent different waves, and allowing for interactions with the predictor so the within-person slope can vary over time.

Fifth, in all models it is easy to include time-invariant covariates as predictors of the random intercept of y. However, the substantive interpretation of the random intercept—and thus the prediction of it—depends on the centering of the time-varying covariate x. If x was centered per person, then the intercept at the within-person level represents the individual mean, which can be interpreted as the person's trait or equilibrium. However, when x is not centered per person, the intercept at the within-person level represents the expected score on y for person i at occasion t, when their predictor x_{it} is zero. This may be rather uninformative, especially if a zero score is not even possible on x. The mean on y for person i can be expressed as a function of the intercept, the (random) slope, and the mean on x for person i. Hence, the predictive relationship between a time-invariant covariate and the mean will in general be different than that between the time-invariant covariate and the intercept. Hence, when the interest is in the effect of a time-invariant covariate, we advice person mean centering of the time-varying covariate in the within-person equation.

To conclude, from Table 3 it can be seen that methods L3a and L3b include most of the features: These methods can accommodate traditional panel data as well as intensive longitudinal data, include random slopes, separate within from between slopes, and provide standardized slopes per level. When choosing between these two options, we prefer L3b, based on latent within-person centering, as it overcomes the bias associated with sample mean centering, which is especially relevant for panel data.

Empirical application

We make use of a dataset that was made available by Bringmann et al. (2013), and that originates from Geschwind et al. (2011).[5] These data come from 128 individuals who reported on their momentary somberness at random occasions during the day, ten times a day for 6 days, resulting in a maximum of 60 measurements per person. Participants were also asked to report on the most important event that took place since the previous beep, and indicate the degree to which this was pleasant or unpleasant.

We will consider the pleasantness of events (EV) as a predictor for somberness (SO). It is important to note that due to the instructions that were used (i.e., "at this moment", versus "since the previous beep", there is an implicit lagged relationship between these variables, even though they were measured at the same time. Furthermore, we will also include a measure of the personality trait neuroticism (N), a time-invariant covariate that was measured prior to the intensive longitudinal data collection period, and investigate whether this trait predicts individual differences in means on both the event variable and the somberness measurements.

Descriptives

Three individuals had no variation on the outcome variable; they were by default excluded from the analysis by Mplus. The intraclass correlation for SO was 0.33, indicating that 67% of the observed variance in somberness is within-person variance, while 33% can be considered as trait-like, between-person variance. In contrast, the intraclass correlation of pleasantness of events was 0.09, implying that only 9% of the variance in reported pleasantness of events can be considered as stemming from stable, between-person differences, while 91% of the variability is within-person.

As indicated, the maximum number of time points was 60; however, most participants had missing data. The smallest number of observations per person was 20, and the average number of observations per person was 47.9.

Analysis

Based on Table 3, we see that since we are dealing with intensive longitudinal data, the long-format approaches are suitable. Moreover, we want to obtain both within-person and between-person slopes, which narrows our options down to L3a and L3b. We prefer latent mean centering (L3b) over observed mean centering (L3a), as earlier work has shown latent mean centering to be superior (Asparouhov & Muthén, 2018); note though that—in line with Figure 3 of Asparouhov and Muthén (2018)—our simulations have shown that this advantage disappears with increasing number of time points.

The within-person model we use can be expressed as

$$EV_{it} = \mu_{EV,i} + EV_{it}^{(w)}$$
$$SO_{it} = \mu_{SO,i} + \beta^{(w)} EV_{it}^{(w)} + e_{it}$$

where the first equation decomposes the predictor into a within-person part $EV_{it}^{(w)}$ and a between-person part $\mu_{EV,i}$, while the second equation decomposes the outcome variable in a between-person part $\mu_{SO,i}$ and a within-person part, and the latter part is to some extent predicted by $EV_{it}^{(w)}$, and partly unpredicted (i.e., e_{it}.

At the between-person level we specify the model as

$$\mu_{EV,i} = \gamma_{00} + \gamma_{01} N_i + u_{0i},$$
$$\mu_{SO,i} = \gamma_{10} + \gamma_{11} N_i + \beta^{(b)} \mu_{EV,i} + u_{1i}$$

which implies that the individual mean on the outcome ($\mu_{SO,i}$) is regressed on the individual mean on the predictor ($\mu_{EV,i}$) and the time-invariant covariate neuroticism. The mean on the predictor ($\mu_{EV,i}$) is also regressed on the time-invariant covariate neuroticism, which implies neuroticism has a direct effect (i.e., γ_{11}) and an indirect effect (i.e., $\gamma_{01}\beta^{(b)}$) on somberness.

Results

The within-person slope $\beta^{(w)}$ is estimated to be -0.204 ($SE = 0.014$, $p < 0.001$), while the between-person slope $\beta^{(b)}$ is estimated to be -0.796 ($SE = 0.139$, $p < 0.001$). The first slope describes the comparison on somberness at two occasions *within the same person*, where these occasions differ one unit in momentary pleasantness of events. The second slope describes the comparison on average level of somberness *between two different individuals*, where these individuals differ one unit in their average pleasantness of events (while also controlling for neuroticism). It thus shows that comparing two individuals that differ one unit on pleasantness of events leads to a larger expected difference in somberness than comparing two occasions within the same individual that differ one unit on pleasantness of events.

To determine the relative importance of pleasantness of events at each level however, we should consider the standardized results. The standardized slope at the within-person level is -0.260 ($SE = 0.016$, $p < 0.001$), and at the between-person level it is -0.441 ($SE = 0.074$, $p < 0.001$). Hence, while the unstandardized between-person slope was about four times larger than the unstandardized within-person slope, the standardized between-person slope is less than twice as large as the standardized within-person slope. This is related to the fact that the two variables have such different intraclass correlations.

Finally, neuroticism is a predictor of both the persons' average levels of pleasantness of events and somberness. The (between-level) standardized direct effect of neuroticism on somberness is 0.329 ($SE = 0.073$, $p < 0.001$), and the (between-level) standardized indirect effect through pleasantness of events is 0.114 ($SE = 0.052$, $p = 0.028$), such that the total (between-level) standardized effect is 0.443 ($SE = 0.070$, $p < 0.001$). This implies that individuals who score higher on neuroticism tend to experience higher levels of somberness, which is only partly mediated through the experience of (un)pleasantness of events.

Discussion

In the current paper we have pulled together several strands of literature on how to model clustered longitudinal data. First, we have discussed the fundamental connection between the FE-RE debate and centering of a time-varying covariate in longitudinal research. Although this connection has been known for over 40 years, it has gone largely unnoticed and is rarely covered in longitudinal modeling teachings. Moreover, most of the existing publications on this topic target researchers from disciplines like sociology, (micro-)econometrics, and political sciences, as reflected by the assumed background knowledge and the jargon that is used. The current paper complements these treatments by using the modeling methods and terminology that are common in psychology instead. Second, we have elaborated on the connection between the long-format multilevel modeling approach and the wide-format SEM approach, and showed how these approaches are related. In doing so, we have once again underscored the importance of separating within-person fluctuations from the between-person differences in both approaches, through centering the time-varying predictor and/or including the cluster means as a predictor at the between-person level.

With this paper we aimed to target multilevel users who have become alarmed after being confronted with the rather negative portrayal of random effects models in other disciplines, which seem to condemn multilevel modeling altogether. We show that this criticism is only relevant to *one particular form of multilevel modeling*: Multilevel modeling based on the raw level 1 predictors or on grand mean centering the level 1 predictors. Indeed, this approach fails to correctly separate the within-person level from the between-person level, which has been widely acknowledged in the multilevel literature (e.g., Bolger & Laurenceau, 2013; Enders & Tofighi, 2007; Hoffman, 2014; Kreft et al., 1995; Nezlek, 2001; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012; Wang & Maxwell, 2015). Yet, this problem is easily overcome by using within-person centering of the time-varying predictor and/or including the cluster mean as a predictor at the between level. We have shown multiple multilevel and SEM methods that can be used to reach this goal.

There are several model extensions the reader is likely to be interested in, most notably with *trends*, *lagged reciprocal relationships*, and the combination of these two. When both x and y are characterized by trends (e.g., increasing or decreasing trajectories, or repetitive cycles) that are left unaccounted for in the model, this will lead to bias in the estimation of $\beta^{(w)}$ (Liu & West, 2016; Wang & Maxwell, 2015). Wang and Maxwell (2015) showed that to account for a trend, one can either: a) detrend the predictor; b) detrend the predictor and the outcome; or c) include the trend as a covariate. These three approaches lead to very similar results. However, Wang and Maxwell (2015) also indicate that if the trend in y is actually the result of a trend in x, then including time as a predictor may obscure this effect. Hence, whether or not to explicitly account for a trend is an important question, which should be considered of a theoretical rather than a statistical nature (Wang & Maxwell, 2015).

Another way in which we may want to extend the model is by allowing for reciprocal lagged effects between x and y, in combination with autoregressvive effects for both variables. Such models are referred to as *dynamic panel models* in the (micro-)econometrics literature. They are closely related to the cross-lagged panel model (Rogosa, 1980), which is a rather popular approach in disciplines like psychology. However, as pointed out repeatedly, the latter model does not account for unobserved heterogeneity (Berry & Willoughby, 2016; Finkel, 1995; Hamaker, Kuiper, & Grasman, 2015; Liker et al., 1985). There have been several approaches that combine cross-lagged and autoregressive relationships with components that account for unobserved heterogeneity in the form of a random intercept, either presented as FE versus RE modeling (e.g., Allison, 2005; Allison, Williams, & Moral-Benito, 2017; Bianconcini & Bollen, 2018; Bollen & Brand, 2010), or focusing explicitly on how to separate stable, between-person differences from temporal, within-person fluctuations (cf. Curran, Howard, Bainter, Lane, & McGinley, 2013; Curran et al., 2012; Hamaker et al., 2015; McArdle & Grimm, 2010; Ormel & Schaufeli, 1991).

A thorough discussion of models that combine both trends and cross-lagged relationships in the context of SEM is provided by Usami, Murayama, and Hamaker (2019). Because SEM is based on data in wide-format, these modeling methods are necessarily limited to relatively small numbers of waves. Recent software innovations for intensive longitudinal data include dynamic structural equation modeling (DSEM; Asparouhov et al., 2018) and residual dynamic structural equation modeling (RDSEM; Asparouhov & Muthén, in press) in Mplus, and continuous time structural equation modeling (`ctsem`; Driver, Oud, & Voelkle, 2017; Driver & Voelkle, in press) in R. These approaches separate stable between-person differences and possible trends from the dynamics of within-person (residual) fluctuations, while also allowing for random slopes.

In closing, we hope this paper has equipped the reader with the necessary knowledge to navigate and appreciate different strands of literature on longitudinal data analysis, and to make informed decisions about which modeling approach to use in practice. While there may be salient differences in our backgrounds, how we approach our data, and the way we express our concerns in different disciplines, we can trust that—in general—we are trying to solve the same fundamental problems; moreover, our solutions may in fact be more similar than apparent at first sight.

Table 1: Modeling methods used and the slope estimates for a simulated data set consisting of 4 waves and 5000 cases.

Method	$\beta^{(w)} = 1$	$\beta^{(b)} = 2$	β	$\gamma_{01} = \beta^{(b)} - \beta^{(w)}$
L1: $y_{it} = \gamma_{00} + u_{0i} + \beta x_{it} + e_{it}$			1.320 (0.011)	
L2: $y_{it} = \gamma_{00} + u_{0i} + \beta^{(w)}(x_{it} - \bar{x}_{i\cdot}) + e_{it}$	1.012 (0.008)			
L3a: $y_{it} = \gamma_{00} + \beta^{(b)}\bar{x}_{i\cdot} + u_{0i} + \beta^{(w)}(x_{it} - \bar{x}_{i\cdot}) + e_{it}$	1.012 (0.008)	1.938 (0.008)		
L3b: $y_{it} = \gamma_{00} + \beta^{(b)}\mu_{x,i} + u_{0i} + \beta^{(w)}(x_{it} - \mu_{x,i}) + e_{it}$	1.012 (0.008)	1.996 (0.009)		
L4: $y_{it} = \gamma^{*}_{00} + \gamma^{*}_{01}\bar{x}_{i\cdot} + u_{0i} + \beta^{(w)}x_{it} + e_{it}$	1.012 (0.008)			0.926 (0.012)
W1: $y_{it} = c + \alpha_i + \beta x_{it} + e_{it}$ $\quad \mathrm{cor}(\alpha_i, x_{it}) = 0$			1.320 (0.010)	
W2: $y_{it} = c + \alpha_i + \beta^{(w)}x_{it} + e_{it}$ $\quad \mathrm{cor}(\alpha_i, x_{it}) \neq 0$	1.012 (0.008)			
W3: $y_{it} = \gamma_{00} + \beta^{(b)}\mu_{x,i} + u_{0i} + \beta^{(w)}(x_{it} - \mu_{x,i}) + e_{it}$	1.012 (0.008)	1.996 (0.009)		
W4: $y_{it} = \gamma^{*}_{00} + \gamma^{*}_{01}\mu_{x,i} + u_{0i} + \beta^{(w)}x_{it} + e_{it}$	1.012 (0.008)			0.984 (0.013)

Slope estimates with standard errors in parentheses as obtained with nine modeling methods for a single dataset generated with the model in Equations 13 and 14. Methods L1-L4 are based on multilevel regression analysis with the data in univariate/long-format, while methods W1-W5 are based on SEM with the data in multivariate/wide-format. Methods L1 and W1 are RE models that do not allow for a correlation between the predictor and the intercept; all other models account for a correlation between the random intercept α_i and the predictor x_{it}, either through: a) centering the predictor (methods L2, L3a, L3b, and W3); b) explicitly including the correlation (method W2); or c) using a contextual modeling approach (methods L4 and W4). Resulting slopes are: the within-person slope $\beta^{(w)}$, the between-person slope $\beta^{(b)}$, the weighted average β, and the contextual effect γ_{01}.

Table 2: Slope estimates for simulated data with 1000 replications, each consisting of 100 cases at 4 waves and 40 waves.

		T=4				T=40			
Parameter	Method	$\bar{\hat{\theta}}$	$SD(\hat{\theta})$	$\bar{SE}$	cov. rate	$\bar{\hat{\theta}}$	$SD(\hat{\theta})$	$\bar{SE}$	cov. rate
β	L1	0.7568	0.0800	0.0832	0.176	0.9617	0.0317	0.0312	0.746
	W1	0.7571	0.0801	0.0827	0.173				
$\beta^{(w)} = 1$	L2	1.0036	0.1099	0.1139	0.958	0.9998	0.0322	0.0318	0.944
	W2	1.0036	0.1099	0.1154	0.965				
	L3a	1.0036	0.1099	0.1139	0.958	0.9998	0.0322	0.0318	0.944
	L3b	1.0036	0.1098	0.1139	0.960	0.9998	0.0322	0.0318	0.944
	W3	1.0036	0.1099	0.1154	0.965				
	L4	1.0036	0.1099	0.1139	0.958	0.9998	0.0319	0.0318	0.945
	W4	1.0036	0.1099	0.1154	0.965				
$\beta^{(b)} = 0.5$	L3a	0.5318	0.1091	0.1076	0.932	0.4994	0.1076	0.0999	0.927
	L3b	0.5015	0.1169	0.1150	0.935	0.4962	0.1083	0.1005	0.926
	W3	0.5014	0.1171	0.1163	0.947				
$\beta^{(b)} - \beta^{(w)} = -0.5$	L4	-0.4717	0.1596	0.1568	0.939	-0.5004	0.1101	0.1048	0.938
	W4	-0.5022	0.1700	0.1693	0.944				

Results obtained for slope estimates from 1000 replications generated with the model in Equations 13 and 14 and analyzed using nine different modeling methods (see Table 1 for the methods). The slopes are: the weighted average β, the within-person slope $\beta^{(w)}$, the between-person slope $\beta^{(b)}$, and the contextual effect $\beta^{(b)} - \beta^{(w)}$. Columns contain the the average point estimate $\bar{\hat{\theta}}$, the standard deviation of the point estimates $SD(\hat{\theta})$, the average standard error of the point estimates $\hat{SE}$, and the coverage rate of the 95% confidence interval of the slope.

Table 3: Guidelines for choosing a modeling method

	Slopes			**Standardized slopes**		**Data suitability**		**Extensions**	
	Within slope $\beta^{(w)}$	Between slope $\beta^{(b)}$	Contextual effect $\beta^{(b)} - \beta^{(w)}$	Standardized within slope	Standardized between slope	Panel data	Intensive longitudinal data	Person-specific slopes	Time-varying slopes
L1	-	-	-	-	-	+	+	+	$(+)^e$
L2	+	-	-	$(+)^c$	-	+	+	+	$(+)^e$
L3a	+	+	$(+)^b$	$(+)^c$	$(+)^c$	+	+	+	$(+)^e$
L3b	+	+	$(+)^b$	$(+)^c$	$(+)^c$	+	+	+	$(+)^e$
L4	+	$(+)^a$	+	-	-	+	+	+	$(+)^e$
W1	-	-	-	-	-	+	-	-	+
W2	+	-	-	-	-	+	-	-	+
W3	+	+	$(+)^b$	$(+)^d$	+	+	-	-	+
W4	+	$(+)^a$	+	-	-	+	-	-	+

Features of the diverse long-format, multilevel modeling methods (L1-L4) and wide-format, structural equation modeling methods (W1-W4), where + implies present and - implies absent. Some features are denoted as partly present (+), meaning: ^{a} between-person slope can be computed using the contextual effect and the within-person slope; ^{b} contextual effect can be computed using the within-person slope and the between-person slope; ^{c} standardized parameters are typically not included in multilevel software, but can be obtained in Mplus with maximum likelihood estimation for models without a random slope, and with Bayesian estimation for models with random slopes; [d] standardized within-person slope can be obtained using a further decomposition of the outcome variable into a within-person and a between-person part, akin to what is done with the predictor; and [e] time-varying slopes can be accomplished by adding dummy variables for occasions, and including interactions between these dummy variables and the within-person predictor.

References

Allison, P. D. (2005). *Fixed effects regression methods for longitudinal data using SAS*. Cary, NC: SAS Instiute Inc.

Allison, P. D. (2009). *Fixed effects regression models*. Thousand Oaks, CA: Sage. doi: 10.4135/9781412993869

Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum likelihood for cross-lagged panel models with fixed effects. *Sociological Research for a Dynamic World*, *3*, 1-17.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation modeling. *Structural Equation Modeling*, *25*, 359-388. doi: 10.1080/10705511.2017.1406803

Asparouhov, T., & Muthén, B. (2018). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling: A Multidisciplinary Journal*. doi: 10.1080/10705511.2018.1511375

Asparouhov, T., & Muthén, B. (in press). Comparison of models for the analysis of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effect modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, *3*, 133-153. doi: 10.1017/psrm.2014.7

Berry, D., & Willoughby, M. T. (2016). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development*, 1-24. doi: 10.1111/cdev.12660

Bianconcini, S., & Bollen, K. A. (2018). The latent variable-autoregressive latent trajectory model: A general framework for longitudinal data analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*, 791-808. doi: 10.1080/10705511.2018.1426467

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research.* New York, NY: The Guilford Press.

Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A structural equations approach. *Sociol Forces*, *89*, 1-34.

Bou, J. C., & Satorra, A. (2017). Univariate versus multivariate modeling of panel data: Model specification and goodness-of-fit testing. *Organizational Research Methods*, 1-47. doi: 10.1177/1094428117715509

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, . . . Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE*, *8*, e60188, 1-13. doi: 10.1371/journal.pone.0060188

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

Cronbach, L. J. (1976). *Research on classrooms and schools: Formulations of questions, design and analysis.* Stanford, CA: Stanford Evaluation Consortium. Retrieved from `https://eric.ed.gov/?id=ED135801`

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, S. (2013). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, *82*(5), 879-894. doi: 10.1037/a0035297

Curran, P. J., Lee, T., Howard, A. L., Lane, S., & MacCallum, R. (2012). Disaggregating witin-person and between-person effects in mutlilevel and structural equation growth models. In J. R. Harring & G. R. Hancock (Eds.), *Advances in longitudinal methods in the social and behavioral sciences* (p. 217-253). Charlotte, NC: Information Age Publishing Inc.

Driver, C. C., Oud, J. H. L., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*, *77*, 1-35. doi: 10.18637/jss.v077.i05

Driver, C. C., & Voelkle, M. C. (in press). Hierarchical Bayesian continuous time dynamic modeling. *Psychological Methods*. doi: 10.1037/met0000168

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

Finkel, S. E. (1995). *Causal analysis with panel data*. Thousand Oaks, CA: Sage.

Geschwind, N., Peeters, F., Drukker, M., van Os, J., & Wichers, M. (2011). Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: A randomized controlled trial. *Journal of Consulting and Clinincal Psychology*, *79*, 618-628. doi: 10.1037/a0024595

Halaby, C. N. (2004). Panel models in sociological research: Theory intro practice.

Annual Review of Socioloy, *30*, 507-544. doi: 10.1146/annurev.soc.30.012703.110629

Hamaker, E. L., & Grasman, R. P. P. P. (2015). To center or not to center? investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology*, *5*, 1492. doi: 10.3389/fpsyg.2014.01492

Hamaker, E. L., Kuiper, R., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102-116. doi: 10.1037/a0038889

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science*, *26*, 10-15. doi: 10.1177/0963721416666518

Hoffman, L. (2014). *Longitudinal analysis: Modeling within-person fluctuations and change*. New York, NY: Taylor & Francis.

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpsons paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*, 513. doi: 10.3389/fpsyg.2013.00513

Kreft, I. G. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, *30*, 1-21.

Laurenceau, J.-P., & Bolger, N. (2005). Using diary methods to study marital and family processes. *Journal of Family Psychology*, *19*, 86-97.

Liker, L. K., Augustyniak, S., & Duncan, G. J. (1985). Panel data and models of change: A comparison of first difference and conventional two-wave models. *Social Science Research*, *14*, 80–101.

Liu, Y., & West, S. G. (2016). Weekly cycles in daily report data: An overlooked issue. *Journal of Personality*, *84*, 560-579. doi: 10.1111/jopy.12182

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*, 203-229. doi: 10.1037/a0012869

McArdle, J. J., & Grimm, K. J. (2010). Five steps in latent curve and latent change score modeling with longitudinal data. In K. van Montfort, J. Oud, & A. Satorra (Eds.), *Longitudinal research with latent variables* (p. 245-274). Heidelberg, Germany: Springer-Verlag.

McNeish, D., & Kelley, K. (2019). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, *24*, 20-35. doi: 10.1037/met0000182

Mundlak, Y. (1978). Pooling of time-series and cross-sectional data. *Econometrica*, *46*, 69-85.

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide. Eighth edition.* Los Angeles, CA: Muthén & Muthén.

Neuhaus, J. M., & Kalbfleisch, J. D. (1998). Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics*, *54*, 638-645. doi: 10.2307/3109770

Nezlek, J. B. (2001). Multilevel random coefficient analyses of event- and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin*, *27*, 771-785.

Ormel, J., & Schaufeli, W. B. (1991). Stability and change in psychological distress and their relationship with self-esteem and locus of control: A dynamic equilibrium model. *Journal of Personality and Social Psychology*, *60*, 288-299. doi: 10.1037/0022-3514.60.2.288

Ousey, G. C., Wilcox, P., & Fisher, B. S. (2011). Something old, something new: Revisiting competing hypothesis of the victimization-offending relationship among adolescents. *Journal of Quantitative Criminology*, *27*, 53-84. doi: 10.1007/s10940-010-9099-1

Raudenbush, S. W. (2009). Adaptive centering with random effects: An alternative to the fixed effects model for studying time-varying treatments in school settings. *Education Finance and Policy*, *4*, 468-491. doi: 10.1162/edfp.2009.4.4.468

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Rogosa, D. R. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245–258.

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*, 206-221. doi: 10.1037/met0000062

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage Publishers.

Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, *9*, 151-176. doi: 10.1146/annurev-clinpsy-050212-185510

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of cross-lagged models. *Psychological Methods*. doi: 10.1037/met0000210

Wang, L., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Multivariate Behavioral Research*, *20*, 63-83. doi: 10.1037/met0000030

Wichers, M. C., Barge-Schaapveld, D. G. C. M., Nicolson, N. A., Peeters, F., de Vries, M., Mengelers, R., & van Os, J. (2009). Reduced stress-sensitivity or increased reward experience: The psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology*, *34*, 923-931. doi: 10.1038/npp.2008.66

Wooldridge, J. M. (2002). *Econometrics analysis of cross sectional and panel data*. Cambridge, MA: The MIT Press.

Wooldridge, J. M. (2013). *Introductory econometrics: A modern apporach* (5th ed.). Mason, OH: South-Western, Cengage Learning.

Footnotes

[1]It is important to note that the terms *fixed effect* and *random effect*, tend to be used differently in these disciplines than how they are typically used in the psychological literature. For more details on this, see https://statmodeling.stat.columbia.edu/2005/01/25/why_i_dont_use/, and Bou and Satorra (2017).

[2]The equivalence between these two expressions only holds when there is a fixed slope; when the slope is random, it no longer holds (cf. Snijders & Bosker, 2012).

[3]The codes for simulating and analyzing these data are provided as Online Supporting Material.

[4]This SEM approach can also be thought of as the latent growth curve model with a time-varying covariate proposed by Curran et al. (2012), but without the random slope for time, and without the constraints on the grand means/intercepts over time.

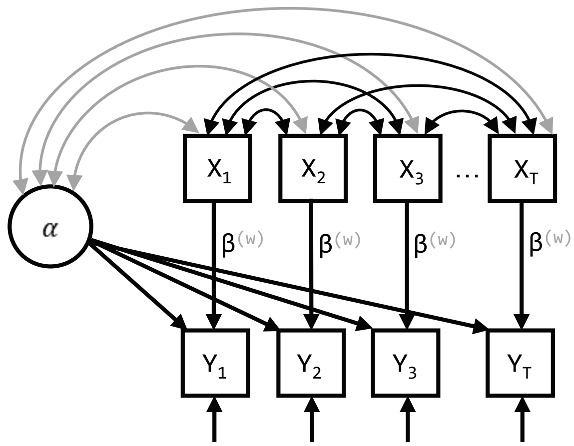
[5]The data were downloaded from this website: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0060188. We considered only the pre-treatment phase in our analyses. We obtained permission from N. Geschwind and M. Wichers to use these data.

Figure Captions

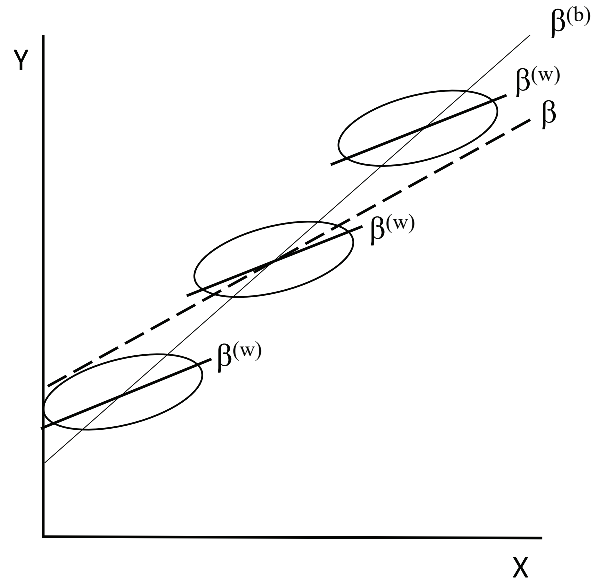
Figure 1. Allison's SEM approaches to Equation 2, in which the correlations between the intercept α_i and the predictors x_{it} are either freely estimated (as in the FE approach) or fixed to zero (as in the RE approach). This changes the slope estimate when regressing y_{it} on x_{it}.

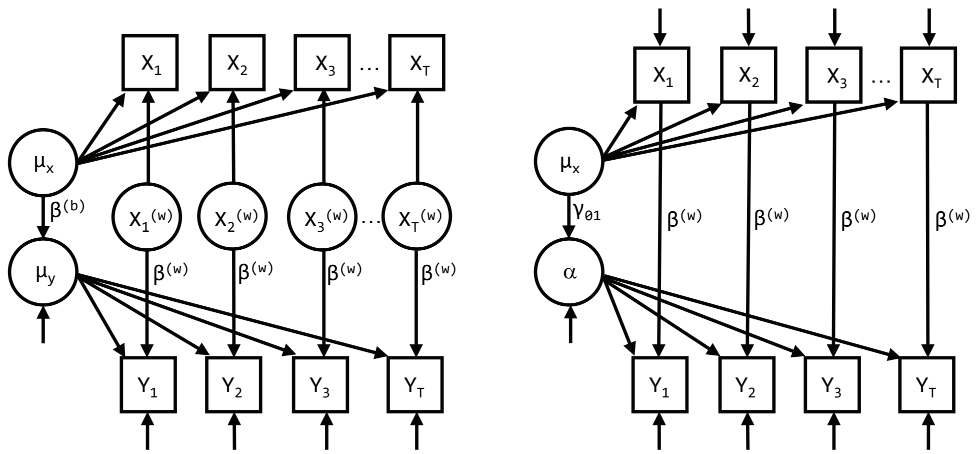
Figure 2. Example of clustered data with different within-person and between-person slopes. Ellipses represent clusters, with $\beta^{(w)}$ as the within-person slope; the cluster means (central points within the ellipses) form the between-person level, with $\beta^{(b)}$ as the between-person slope. The dashed slope γ_{01} is a weighted average of the within-person and between-person slopes.

Figure 3. Two structural equation modeling approaches to longitudinal data in wide-format, which allow for a relationship between the random intercept and the predictor. Model on the left represents method W3 in Table 1, and is based on centering the predictor per person using the latent mean μ_{xi}, and regressing the outcome variable on the within-person centered predictor (denoted as $x_{it}^{(w)}$). This results in the within-person slope $\beta^{(w)}$. In addition, the person mean on the outcome is regressed on the person mean on the predictor, resulting in the between-person slope $\beta^{(b)}$. Model on the right represents method W4 in Table 1, and is based on regressing the outcome variable on the predictor, while using the latent mean per person on the predictor to predict the intercept of the outcome. This results in the within-person slope $\beta^{(w)}$, and the contextual effect $\gamma_{01} = \beta^{(b)} - \beta^{(w)}$.



α
X_1
X_2
X_3
…
X_T
$\beta^{(w)}$
$\beta^{(w)}$
$\beta^{(w)}$
$\beta^{(w)}$
Y_1
Y_2
Y_3
Y_T


Y
X
β(b)
β(w)
β
β(w)
β(w)



X_1 X_2 X_3 … X_T
μ_x
$\beta^{(b)}$
μ_y
$X_1^{(w)}$ $X_2^{(w)}$ $X_3^{(w)}$ … $X_T^{(w)}$
$\beta^{(w)}$ $\beta^{(w)}$ $\beta^{(w)}$ $\beta^{(w)}$
Y_1 Y_2 Y_3 Y_T
X_1 X_2 X_3 … X_T
μ_x
γ_{01}
α
$\beta^{(w)}$ $\beta^{(w)}$ $\beta^{(w)}$ $\beta^{(w)}$
Y_1 Y_2 Y_3 Y_T

Supplementary Materials

for

The fixed versus random effects debate and how it relates to centering in multilevel modeling

E. L. Hamaker & B. Muthén

R-code for simulating (and analyzing) a single data set

and Mplus input for analyzing a single data set

R-code

```
#######################################
#### SIMULATE DATA WITH DIFFERENT ####
###### WITHIN AND BETWEEN SLOPES #####
#######################################

rm(list=ls())
library(lme4)
set.seed(3859)

N<-5000     # number of clusters
n<-4        # number of observations within a cluster

Y <- matrix(,N,n) # to store the outcome
X <- matrix(,N,n) # to store the predictor
X.TrueCent <- matrix(,N,n) # to store the true centered predictor


########################
#### SET PARAMETER VALUES ####
########################

# PREDICTOR
sdX.within <- sqrt(1)        # within-person variance
sdX.between <- sqrt(4)  # between-person variance

# INTERCEPT LEVEL 2
g.00 <- 0               # Grand intercept
g.01 <- 2               # between-cluster slope
sd.u0 <- 1              # SD of residuals intercept at level 2

# SLOPE LEVEL 2
g.10 <- 1               # fixed within-cluster slope
sd.u1 <- 0              # SD of within-cluster slope at level 2

# RESIDUALS AT LEVEL 1
sd.e <- 1               # residual SD at level 1

#######################
#### SIMULATE DATA ####
#######################

for (j in 1:N)
{     # sample mean on x in cluster j
      X.mean.j <- rnorm(1,mean=0,sd=sdX.between)
      # sample x in cluster j
      X.j <- rnorm(n,mean=X.mean.j,sd=sdX.within)
      X[j,] <- X.j
      # sample slope in cluster j (here identical across individuals)
      b1.j <- rnorm(1,g.10,sd.u1)
      # sample intercept in cluster j (level 2 expression)
      b0.j <- rnorm(1,g.00,sd.u0) + g.01*X.mean.j
      # sample y (level 1 expression)
      Y[j,] <- b0.j + b1.j*(X.j-X.mean.j) + rnorm(n,0,sd.e)
}
```

```
########################
### DATA PREPARATION ###
########################

Y <- c(t(Y))
X <- c(t(X))
Cluster <- rep(1:N, each=n)
Time <- rep(1:n,N)

#####################################
### ESTIMATE SAMPLE CLUSTER MEANS ###
###### AND CENTER X WITH THIS #######
#####################################

cluster.means <- NA
cluster.j <- 0
for (j in 1:N){
      cluster.mean <- rep(mean(X[((cluster.j*n)+1):((cluster.j*n)+n)]),n)
      cluster.means <- c(cluster.means,cluster.mean)
      cluster.j <- cluster.j + 1
}
cluster.means <- as.vector(cluster.means)[-1]

X.cent <- X - cluster.means
data.cent <- as.data.frame(cbind(Y, X.cent,X, Cluster, cluster.means))

########################
### ESTIMATE MODELS ####
########################

# L1: RAW X (RE MODEL)
summary(lmer(Y ~ X + (1 |Cluster), data = data.cent, REML = F))

# L2: CENTERED X (FE MODEL)
summary(lmer(Y ~ X.cent + (1 |Cluster), data = data.cent, REML = F))

# L3a: CENTERED X PLUS SAMPLE MEANS ON X (WITHIN-BETWEEN MODEL)
summary(lmer(Y ~ X.cent + cluster.means + (1 |Cluster), data = data.cent,
REML = F))

# L4: RAW X PLUS SAMPLE MEANS ON X (MUNDLAK'S CONTEXTUAL MODEL)
summary(lmer(Y ~ X + cluster.means + (1 |Cluster), data = data.cent, REML =
F))

##################
### WRITE DATA ###
##################

write.table(cbind(round(data.cent,5),Time),
file="BWdataT4.dat",col.names=FALSE,row.names=FALSE)
```

Method L1

```
TITLE:  Multilevel model

DATA:       FILE = BWdataT4.dat;

VARIABLE:       NAMES = y xcent x clus clusmean;
                USEVARIABLE = y x;
                WITHIN = x;
                BETWEEN = ;
                CLUSTER = clus;

ANALYSIS:   TYPE = TWOLEVEL;

MODEL:
        %WITHIN%
        y ON x;
        %BETWEEN%
        y;
```

Mplus code method L2: Multilevel model with within-person centered predictor

```
DATA:         FILE = BWdataT4.dat;

VARIABLE:     NAMES = y xcent x clus clusmean;
              USEVARIABLE = y xcent;
              WITHIN = xcent;
              BETWEEN =;
              CLUSTER = clus;

ANALYSIS:     TYPE = TWOLEVEL;

MODEL:        %WITHIN%
              y ON xcent;
              %BETWEEN%
              y;
```

Or center the predictor within Mplus with the sample mean per cluster (using the DEFINE statement):

```
DATA:         FILE = BWdataT4.dat;

VARIABLE:     NAMES = y xcent x clus clusmean;
              USEVARIABLE = y x;
              WITHIN = x;
              BETWEEN = ;
              CLUSTER = clus;

DEFINE:       CENTER x (GROUPMEAN);

ANALYSIS:     TYPE = TWOLEVEL;

MODEL:        %WITHIN%
              y ON x;
              %BETWEEN%
              y;
```

Mplus code for method L3a: Multilevel model with centered X variable and sample mean as a between level predictor:

```
DATA:          FILE = BWdataT4.dat;

VARIABLE:      NAMES = y xcent x clus clusmean;
               USEVARIABLE = y xcent clusmean;
               WITHIN = xcent;
               BETWEEN = clusmean;
               CLUSTER = clus;

ANALYSIS:      TYPE = TWOLEVEL;

MODEL:         %WITHIN%
               y ON xcent;
               %BETWEEN%
               y ON clusmean;
```

Or center the predictor within Mplus with the sample mean per cluster, and save these cluster means (meanx) to use as a between level predictor (using the DEFINE statement):

```
DATA:          FILE = BWdataT4.dat;

VARIABLE:      NAMES = y xcent x clus clusmean;
               USEVARIABLE = y x meanx;
               WITHIN = x;
               BETWEEN = meanx;
               CLUSTER = clus;

DEFINE:        CENTER x (GROUPMEAN);
               meanx = CLUSTER_MEAN (x);

ANALYSIS:      TYPE = TWOLEVEL;

MODEL:         %WITHIN%
               y ON x;
               %BETWEEN%
               y ON meanx;
```

Mplus code for method L3b: Multilevel model with centered X variable using latent mean centering

```
DATA:          FILE = BWdataT4.dat;

VARIABLE:      NAMES = y xcent x clus clusmean;
               USEVARIABLE = y x;
               WITHIN = ;
               BETWEEN = ;
               CLUSTER = clus;

ANALYSIS:      TYPE = TWOLEVEL;

MODEL:         %WITHIN%
               y ON x;
               %BETWEEN%
               y ON x;
```

Mplus code for method L4: Multilevel model with raw X variable and cluster means (clusmean) as between level predictor:

```
DATA:       FILE = BWdataT4.dat;

VARIABLE:       NAMES = y xcent x clus clusmean;
                USEVARIABLE = y x clusmean;
                WITHIN = x;
                BETWEEN = clusmean;
                CLUSTER = clus;

ANALYSIS:   TYPE = TWOLEVEL;

MODEL:
        %WITHIN%
        y ON x;
        %BETWEEN%
        y ON clusmean;
```

Or computing the cluster mean (meanx) inside Mplus using the DEFINE statement:

```
DATA:           FILE = BWdataT4.dat;

VARIABLE:       NAMES = y xcent x clus clusmean;
                USEVARIABLE = y x meanx;
                WITHIN = x;
                BETWEEN = meanx;
                CLUSTER = clus;

DEFINE:         meanx = CLUSTER_MEAN (x);

ANALYSIS:       TYPE = TWOLEVEL;

MODEL:          %WITHIN%
                y ON x;
                %BETWEEN%
                y ON meanx;
```

Mplus code for method W1: Model with a random intercept that is uncorrelated with X (the RE model)

```
DATA:           FILE = BWdataT4.dat;

DATA LONGTOWIDE: ! Convert the long-format data file to wide-format
                LONG = y | x;
                WIDE = y1-y4 | x1-x4;
                IDVARIABLE = clus;
                REPETITION = Time (1-4);

VARIABLE:       NAMES = y xcent x clus clusmean time;
                USEVARIABLE = y1-y4 x1-x4;

MODEL:          alpha by y1-y4@1;      ! random effect
                y1-y4 PON x1-x4 (b1);
                alpha WITH x1-x4@0;  ! fixing the correlations to zero

                ! Stationarity constraints (optional):
                ! this allows for comparing it to the long format analyses
                x1-x4 WITH x1-x4 (r);
                x1-x4 (vx);
                [y1-y4] (b0);
                [x1-x4] (mu);
                y1-y4 (v);
```

Mplus code for method W2: Model with a random intercept that is correlated with X (the FE model)

```
DATA:           FILE = BWdataT4.dat;

DATA LONGTOWIDE: ! Convert the long-format data file to wide-format
                LONG = y | x;
                WIDE = y1-y4 | x1-x4;
                IDVARIABLE = clus;
                REPETITION = Time (1-4);

VARIABLE:       NAMES = y xcent x clus clusmean time;
                USEVARIABLE = y1-y4 x1-x4;

MODEL:          alpha by y1-y4@1;     ! random effect
                y1-y4 PON x1-x4 (b1);

                ! Stationarity constraints (optional):
                ! this allows for comparing it to the long format analyses
                x1-x4 WITH x1-x4 (r);
                x1-x4 (vx);
                [y1-y4] (b0);
                [x1-x4] (mu);
                y1-y4 (v);
```

Mplus code for method W3: Model with a random intercept that is correlated with X (the FE model)

```
DATA:           FILE = BWdataT4.dat;

DATA LONGTOWIDE: ! Convert the long-format data file to wide-format
                LONG = y | x;
                WIDE = y1-y4 | x1-x4;
                IDVARIABLE = clus;
                REPETITION = Time (1-4);

VARIABLE:       NAMES = y xcent x clus clusmean time;
                USEVARIABLE = y1-y4 x1-x4;

MODEL:          alpha by y1-y4@1;      ! person-specific factor for y
                eta BY x1-x4@1;        ! person-specific factor for x
                alpha ON eta;          ! between-person regression

                ! Create within-person centered X variables
                cx1 BY x1; cx2 BY x2; cx3 BY x3; cx4 BY x4;

                ! Overrule Mplus defaults
                x1-x4@0;                         ! measurement error variances fix to zero
                eta cx1-cx4 WITH cx1-cx4@0; ! default covariances fixed to zero

                y1-y4 PON cx1-cx4 (b1); ! within-person regression

                ! Stationarity constraints (optional)
                ! this allows for comparing it to the long format analyses
                [y1-y4] (b0);   ! stationarity constraints on intercepts of y
                [x1-x4] (mu);   ! stationarity constraints on means of x
                y1-y4 (v);      ! stationarity constraints on residual variances of y
                cx1-cx4 (cvx);  ! stationarity constraints on variances of within-person part of x
```

Mplus code for method W4: Regressing the outcome on the raw X, while including the within-person mean on X as a predictor of the random intercept

```
DATA:           FILE = BWdataT4.dat;

DATA LONGTOWIDE:
                LONG = y | x;
                WIDE = y1-y4 | x1-x4;
                IDVARIABLE = clus;
                REPETITION = Time (1-4);

VARIABLE:       NAMES = y xcent x clus clusmean time;
                USEVARIABLE = y1-y4 x1-x4;

MODEL:          alpha by y1-y4@1;     ! person-specific factor for y
                eta BY x1-x4@1;       ! person-specific factor for x
                alpha ON eta;         ! between-person regression

                y1-y4 PON x1-x4 (b1); ! within-person regression

                ! Stationarity constraints (optional)
                ! this allows for comparing it to the long format analyses
                [y1-y4] (b0);   ! stationarity constraints on intercepts of y
                [x1-x4] (mu);   ! stationarity constraints on means of x
                y1-y4 (v);      ! stationarity constraints on residual variances of y
                x1-x4 (cvx);    ! stationarity constraints on residual variances of x
```

R-code for simulating data for 1000 replications and Mplus input to analyze these data

```
######################################
#### SIMULATE DATA WITH DIFFERENT ####
###### WITHIN AND BETWEEN SLOPES #####
######################################

rm(list=ls())
set.seed(2328)

R <- 1000   # numbser of replications
N <- 100    # number of clusters
N <- 4      # number of observations within a cluster

Lnames <- NA
Wnames <- NA

##############################
#### SET PARAMETER VALUES ####
##############################

# PREDICTOR
sdX.within <- sqrt(1)   # within-person variance
sdX.between <- sqrt(4)  # between-person variance

# INTERCEPT LEVEL 2
g.00 <- 0               # Grand intercept
g.01 <- .5              # between-cluster slope
sd.u0 <- 2              # SD of residuals intercept at level 2

# SLOPE LEVEL 2
g.10 <- 1               # fixed within-cluster slope
sd.u1 <- 0              # SD of within-cluster slope at level 2

# RESIDUALS AT LEVEL 1
sd.e <- 2               # residual SD at level 1

v.y.within <- g.10^2*sdX.within^2 + sd.e^2
v.y.between <- g.01^2*sdX.between^2 + sd.u0^2

v.y.between/(v.y.between+v.y.within)

#######################
#### SIMULATE DATA ####
#######################


for (r in 1:R){

Y <- matrix(,N,n) # to store the outcome
X <- matrix(,N,n) # to store the predictor

      for (j in 1:N)
      {     # sample mean on x in cluster j
      X.mean.j <- rnorm(1,mean=0,sd=sdX.between)
      # sample x in cluster j
            X.j <- rnorm(n,mean=X.mean.j,sd=sdX.within)
            # sample slope in cluster j (here identical across individuals)
```

```
            b1.j <- rnorm(1,g.10,sd.u1)
            # sample intercept in cluster j (level 2 expression)
            b0.j <- rnorm(1,g.00,sd.u0) + g.01*X.mean.j
            # sample y (level 1 expression)
            Y[j,] <- b0.j + b1.j*(X.j-X.mean.j) + rnorm(n,0,sd.e)
            # save x and the within-person centered
            # x using the true mean on x
      X[j,] <- X.j
      }

      dataW <- cbind(Y,X)

      ###################################
      ### CONVERT DATA TO LONG FORMAT ###
      ###################################

      YL <- c(t(Y))
      XL <- c(t(X))
      Cluster <- rep(1:N, each=n)
      Occasion <- rep(1:n,N)
      dataL <- as.data.frame(cbind(YL, XL, Cluster, Occasion))

      ##################
      ### WRITE DATA ###
      ##################

      write.table(round(dataW,5),file=paste0("N100T4Wrep",
                  r,".dat"),col.names=FALSE,row.names=FALSE)

      write.table(round(dataL,5),file=paste0("N100T4Lrep",
                  r,".dat"),col.names=FALSE,row.names=FALSE)

      Lnames <- c(Lnames,paste0("N100T4Lrep",r,".dat"))
      Wnames <- c(Wnames,paste0("N100T4Wrep",r,".dat"))

}


##############################################
### SAVE DATA IN LONG AND WIDE FORMAT ########
##############################################


write.table(Lnames[-1],quote=FALSE,file="N100T4Lreplist.dat",
            col.names=FALSE,row.names=FALSE)

write.table(Wnames[-1],quote=FALSE,file="N100T4Wreplist.dat",
            col.names=FALSE,row.names=FALSE)
```

Mplus code to compute the intraclass correlation

```
TITLE:          Simulation study based on 1000 replications
                Data in long format

DATA:           FILE = N100T4Lreplist.dat;      ! Panel data with T=4
                ! FILE = N100T40Lreplist.dat;  ! Intensive longitudinal data with T=40
                TYPE = MONTECARLO;

VARIABLE:       NAMES = y x clus;
                USEVARIABLE = y x;
                WITHIN = ;
                BETWEEN = ;
                CLUSTER = clus;

ANALYSIS:       TYPE = TWOLEVEL;

MODEL:          %WITHIN%
                y (vyw);
                x (vxw);
                %BETWEEN%
                y (vyb);
                x (vxb);

MODEL CONSTRAINT:
                new (iccY); iccY=vyb/(vyb+vyw);
                new (iccX); iccX=vxb/(vxb+vxw);

OUTPUT:         TECH1;
```

Mplus code method L1

```
TITLE:          Simulation study based on 1000 replications
                Data in long format

DATA:           FILE = N100T4Lreplist.dat;      ! Panel data with T=4
                ! FILE = N100T40Lreplist.dat;  ! Intensive longitudinal data with T=40
                TYPE = MONTECARLO;

VARIABLE:       NAMES = y x clus;
                USEVARIABLE = y x;
                WITHIN = x;
                BETWEEN = ;
                CLUSTER = clus;

ANALYSIS:       TYPE = TWOLEVEL;

MODEL:
                %WITHIN%
                y ON x*1;
                %BETWEEN%
                y*1;
                [y*0];

OUTPUT:         TECH1;
```

Mplus code method L2

```
TITLE:        Simulation study based on 1000 replications
              Data in long format

DATA:         FILE = N100T4Lreplist.dat;      ! Panel data with T=4
              ! FILE = N100T40Lreplist.dat;  ! Intensive longitudinal data with T=40
              TYPE = MONTECARLO;

VARIABLE:     NAMES = y x clus;
              USEVARIABLE = y x;
              WITHIN = x;
              BETWEEN = ;
              CLUSTER = clus;

DEFINE:       CENTER x (GROUPMEAN);

ANALYSIS:     TYPE = TWOLEVEL;

MODEL:        %WITHIN%
              y ON x*1;
              x*1;
              y*1;
              %BETWEEN%
              y*1;
              [y*0];

OUTPUT:       TECH1;
```

Mplus code method L3a

```
TITLE:        Simulation study based on 1000 replications
              Data in long format

DATA:         FILE = N100T4Lreplist.dat;     ! Panel data with T=4
              ! FILE = N100T40Lreplist.dat;  ! Intensive longitudinal data with T=40
              TYPE = MONTECARLO;

VARIABLE:     NAMES = y x clus;
              USEVARIABLE = y x meanX;
              WITHIN = x;
              BETWEEN = meanX;
              CLUSTER = clus;

DEFINE:       meanX = CLUSTER_MEAN (x);
              CENTER x (GROUPMEAN);

ANALYSIS:     TYPE = TWOLEVEL;

MODEL:        %WITHIN%
              y ON x*1;
              x*1;
              y*4;
              %BETWEEN%
              y ON meanX*.5;
              meanX*4;
              y*4;
              [meanX*0];
              [y*0];

OUTPUT:       TECH1;
```

Mplus code method L3b

```
TITLE:        Simulation study based on 1000 replications
              Data in long format

DATA:         FILE = N100T4Lreplist.dat;     ! Panel data with T=4
              ! FILE = N100T40Lreplist.dat;  ! Intensive longitudinal data with T=40
              TYPE = MONTECARLO;

VARIABLE:     NAMES = y x clus;
              USEVARIABLE = y x;
              WITHIN = ;
              BETWEEN = ;
              CLUSTER = clus;

ANALYSIS:     TYPE = TWOLEVEL;

MODEL:        %WITHIN%
              y ON x*1;
              x*1;
              y*4;
              %BETWEEN%
              y ON x*.5;
              x*4;
              y*4;
              [x*0];
              [y*0];

OUTPUT:       TECH1;
```

Mplus code method L4

```
TITLE:        Simulation study based on 1000 replications
              Data in long format

DATA:         FILE = N100T4Lreplist.dat;      ! Panel data with T=4
              ! FILE = N100T40Lreplist.dat;  ! Intensive longitudinal data with T=40
              TYPE = MONTECARLO;

VARIABLE:     NAMES = y x clus;
              USEVARIABLE = y x meanX;
              WITHIN = x;
              BETWEEN = meanX;
              CLUSTER = clus;

DEFINE:       meanX = CLUSTER_MEAN (x);

ANALYSIS:     TYPE = TWOLEVEL;

MODEL:        %WITHIN%
              y ON x*1;
              x*1;
              y*4;
              %BETWEEN%
              y ON meanX*-.5;
              meanX*4;
              y*4;
              [meanX*0];
              [y*0];

OUTPUT:       TECH1;
```

Mplus code method W1

```
TITLE:        Simulation study based on 1000 replications
              Data in wide format

DATA:         FILE = N100T4Wreplist.dat;
              TYPE = MONTECARLO;

VARIABLE:     NAMES = y1-y4 x1-x4;

MODEL:        alpha by y1-y4@1;     ! random effect
              y1-y4 PON x1-x4*1 (b1);
              alpha WITH x1-x4@0;   ! fixing the correlations to zero

              ! Stationarity constraints (optional)
              x1-x4 WITH x1-x4 (r);
              x1-x4 (vx);
              [y1-y4] (b0);
              [x1-x4] (mu);
              y1-y4 (v);

OUTPUT:       TECH1;
```

Mplus code method W2

```
TITLE:        Simulation study based on 1000 replications
              Data in wide format

DATA:         FILE = N100T4Wreplist.dat;
              TYPE = MONTECARLO;

VARIABLE:     NAMES = y1-y4 x1-x4;

MODEL:        alpha by y1-y4@1;     ! Latent variable
              y1-y4 PON x1-x4*1 (b1);

              ! Stationarity constraints
              [y1-y4] (b0);
              [x1-x4] (mu);
              x1-x4 (vx);
              y1-y4 (v);
              x1-x4 WITH x1-x4 (r);
              x1-x4 WITH alpha (k);

OUTPUT:       TECH1;
```

Mplus code method W3

```
TITLE:       Simulation study based on 1000 replications
             Data in wide format

DATA:        FILE = N100T4Wreplist.dat;
             TYPE = MONTECARLO;

VARIABLE:    NAMES = y1-y4 x1-x4;

MODEL:       alpha by y1-y4@1;      ! person-specific factor for y
             eta BY x1-x4@1;        ! person-specific factor for x
             alpha ON eta*.5;       ! between-person regression
             alpha*4;               ! between-person residual variance on y
             eta*4;                 ! between-person variance on x

             ! Create within-person centered X variables
             cx1 BY x1;
             cx2 BY x2;
             cx3 BY x3;
             cx4 BY x4;
             x1-x4@0;               ! measurement error variances fix to zero
             eta cx1-cx4 WITH cx1-cx4@0; ! default covariances fixed to zero

             y1-y4 PON cx1-cx4*1 (b1);   ! within-person regression

             ! Stationarity constraints (optional)
             [y1-y4*0] (b0);        ! stationarity constraints on intercepts of y
             [x1-x4*0] (mu);        ! stationarity constraints on means of x
             y1-y4*4 (v);           ! stationarity constraints on residual variances of y
             cx1-cx4*1 (cvx);       ! stationarity constraints on variances of within part of x

OUTPUT:      TECH1;
```

Mplus code method W4

```
TITLE:        Simulation study based on 1000 replications
              Data in wide format

DATA:         FILE = N100T4Wreplist.dat;
              TYPE = MONTECARLO;

VARIABLE:     NAMES = y1-y4 x1-x4;

MODEL:        alpha by y1-y4@1;      ! person-specific factor for y
              eta BY x1-x4@1;        ! person-specific factor for x
              alpha ON eta*-0.5;     ! contextual effect (b^(b) - b^(w))

              y1-y4 PON x1-x4*1 (b1); ! within-person regression

              ! Stationarity constraints (optional)
              ! this allows for comparing it to the long format analyses
              [y1-y4] (b0);   ! stationarity constraints on intercepts of y
              [x1-x4] (mu);   ! stationarity constraints on means of x
              y1-y4 (v);      ! stationarity constraints on residual variances of y
              x1-x4 (cvx);    ! stationarity constraints on residual variances of x

OUTPUT:       TECH1;
```

Empirical application

The data here come from Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, . . . Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. PLoS ONE, 8 , e60188, 1-13. doi: 10.1371/journal.pone.0060188; and were obtained here:
http://journals.plos.org/plosone/article/file?type=supplementary&id=info:doi/10.1371/journal.pone.0060188.s004

The analyses are based on the first phase of that study only (prior to intervention); the dataset Bringmann1.dat only contains data from the first phase, from which we make use of momentary somberness (outcome) and event (time-varying covariate). Additionally, we use the level 2 variable neuroticism, which was measured prior to the intensive longitudinal measurements (time-invariant covariate). This variable was rescaled in Mplus prior to analysis (using the DEFINE statement, see below), for computational purposes.

Mplus input:

```
TITLE:  Multilevel model with latent mean centering for the X variable and a fixed slope

DATA:  FILE = Bringmann1.dat;

VARIABLE:      NAMES =  ID dayno beepno informat04 st_period
                             cheerful event worry fearful somber
                             relaxed neurotic;
               USEVARIABLE = somber event neurotic N;
               CLUSTER = ID;
               BETWEEN = neurotic N;
               MISSING = ALL(-99);

DEFINE:        N = neurotic/10;

ANALYSIS:      TYPE = TWOLEVEL;

MODEL:         %WITHIN%
               somber ON event;
               %BETWEEN%
               somber ON event N;
               event ON N;

MODEL INDIRECT: somber IND N;

OUTPUT: TECH1 STDYX SAMPSTAT;
```