To cite this article:

Lafit, G., Ariens, S., & Artner, R. (in press). Shedding some light on the relationship between measurement error and statistical power in multilevel models applied to intensive longitudinal designs. *British Journal of Mathematical and Statistical Psychology*.

Shedding some light on the relationship between measurement error and statistical power in multilevel models applied to intensive longitudinal designs

Ginette Lafit*[1], Sigert Ariens[2], and Richard Artner[2]

[1]Methodology of Educational Sciences, KU Leuven, Leuven, Belgium
[2]Research Group of Quantitative Psychology and Individual Differences, KU Leuven, Leuven, Belgium

This manuscript has been accepted for publication at the British Journal of Mathematical and Statistical Psychology on January 2026

Abstract

We examine multilevel models applied to intensive longitudinal (IL) designs. Many measurements in IL research are influenced by measurement error, which can compromise the consistency of estimates obtained through maximum likelihood estimation (MLE). While previous research has addressed the impact of measurement error in broader multilevel contexts, its effects on statistical power—particularly in relation to cross-level interaction effects—have not been thoroughly explored. This study aims to clarify the relationship between measurement error and statistical power by deriving analytical formulas for the asymptotic bias and precision matrix of the MLE of fixed effects in multilevel models that account for additive measurement errors and autoregressive [AR(1)] within-person errors. Furthermore, we analyze how different sources of measurement error affect the standard errors of fixed effects estimates and the overall statistical power, specifically when both time-varying and time-invariant predictors are measured with error. We also investigate the estimation properties of the MLE of the fixed effects, assuming autocorrelated measurement errors in the response variable. Our findings show that measurement error in predictors leads to downward bias in the MLE of fixed effects. In contrast, the MLE estimates of fixed effects remain unbiased when there is autocorrelated measurement error in the response variable.

Keywords: statistical power; measurement error; intensive longitudinal designs; linear mixed effect models; AR(1) errors.

*Corresponding author: Ginette Lafit (ginette.lafit@kuleuven.be)

1 Introduction

Intensive longitudinal (IL) designs, such as the experience sampling method (ESM), offer the possibility of gathering real-time data on persons' feelings, thoughts, or behaviors (Mestdagh and Dejonckheere, 2021; Myin-Germeys et al., 2018). These designs have a hierarchical structure in which repeated measurements are nested within persons. A popular statistical approach to model IL data is the multilevel regression model. This modeling approach is commonly used to characterize time-varying processes within persons, and how they differ between persons. For example, it provides the basis for investigating between-person differences in the mean level of a time-varying response variable, associations between time-varying variables, and interaction effects between time-varying and time-invariant predictors (see e.g., Myin-Germeys and Kuppens, 2022).

In designing an IL study, it is crucial to determine an adequate sample size, as researchers typically need to balance statistical power with costs and feasibility. A powerful and popular method for conducting power analysis in multilevel models is the simulation-based approach (see e.g., Arend and Schäfer, 2019; Lafit et al., 2021, 2022; Lane and Hennes, 2018). Recently, Lafit et al. (2024) proposed a quick alternative that does not rely on simulations whatsoever. Instead of simulations, it uses asymptotic approximations of the precision matrix (i.e., inverse of the covariance of the maximum likelihood estimate of the fixed effects).

The previously discussed methods for conducting power analysis in multilevel models assume the time-varying and time-invariant variables are measured without error. Unfortunately, measures typically collected in IL research are often subject to measurement error (Dejonckheere et al., 2022; Brose et al., 2020; Eisele et al., 2021; Schuurman and Hamaker, 2019). These errors lead to inconsistent estimates of fixed effects and variance components when multilevel models are estimated using maximum likelihood (ML) (see e.g., Buonaccorsi et al., 2000; Tosteson et al., 1998; Wang et al., 1998; Zhong et al., 2002). Notwithstanding, the relation between measurement error and statistical power for multilevel models with AR(1) Level 1 errors has received less attention. For multilevel models not applied to IL data, Mathieu et al. (2012) conducted a simulation study and showed that the degree to which measurements are free from error has a limited impact on the statistical power to detect cross-level interaction effects. However, the authors warranted further research as the exact relationship between measurement error and statistical power for cross-level interaction effects is unknown.

This study aims to clarify how measurement error affects statistical power in multilevel models by exploring the precise mathematical relationship between the two. We will build on the approach proposed by Lafit et al. (2024) to derive exact analytical formulas for the precision matrix of the fixed effects in multilevel models that incorporate AR(1) within-person errors in the context of additive measurement errors. Using these analytical derivations, we will investigate the asymptotic properties of the ML estimates of the fixed effects when both time-varying and time-invariant variables are measured with error. Additionally, we will assess how various sources of measurement error influence the standard errors of the

ML estimates of fixed effects, and consequently, the statistical power of the model. Our focus will be on a multilevel model where both the time-varying and time-invariant predictor variables are subject to measurement errors. Furthermore, we will investigate the effect of autocorrelated measurement error in the response variables. For this model, we will derive an asymptotic approximation of the bias in the ML estimates of the fixed effects. We will demonstrate that, similar to multiple linear regression (Buonaccorsi et al., 2000), measurement error in the predictors leads to inconsistent estimates, which are asymptotically downward-biased. Conversely, we will show that autocorrelated measurement error in the response variable remains consistent and asymptotically unbiased. Finally, we will examine how measurement error affects statistical power when testing the null hypothesis regarding the cross-level interaction effect between time-varying and time-invariant predictors.

2 Measurement error in multilevel linear model

We focus on a two-level model including a continuous a Level 1 dependent variable y_{it} for person $i = (1, \ldots, N)$ and measurement occasion $t = (1, \ldots, T)$, where T denotes the total number of measurement occasions, which is assumed to be that same for all persons:

$$y_{it} = \gamma_{0i} + \gamma_{1i}\tilde{x}_{it} + \epsilon_{it}$$

where $\tilde{x}_{it}$ is a continuous Level 1, or time-varying, person-mean centered continuous predictor, such that $\tilde{x}_{it} = x_{it} - \bar{x}_i$, with $\bar{x}_i$ denotes the sample mean of the Level 1 predictor of person i. In this model, γ_{0i} is the intercept for person i, γ_{1i} is the slope of the Level 1 predictor of person i, and ϵ_{it} is the Level 1 or within-person error of person i at time-point t. We assume that the Level 1 errors follow an autoregressive [AR(1)] process (Goldstein et al., 1994):

$$\epsilon_{it} = \rho_\epsilon \epsilon_{it-1} + \varepsilon_{it}$$

where ρ_ϵ denotes the correlation between two consecutive errors and ε_{it} is a random variable that is normally distributed with mean zero and variance σ^2_ε. In this model, the variance of the Level 1 errors is given by $\sigma^2_\epsilon = \sigma^2_\varepsilon/(1 - \rho^2_\epsilon)$, and this model is guaranteed to be stationary by assuming that $|\rho_\epsilon|$ is strictly smaller than 1 (Hamilton, 1994).

To model differences between persons in their person-specific intercept and slope, we assume the following Level 2 models:

$$\gamma_{0i} = \beta_{00} + \beta_{01} w_i + \nu_{0i}$$
$$\gamma_{1i} = \beta_{10} + \beta_{11} w_i + \nu_{1i}$$

where w_i is a Level 2, or time-invariant, continuous predictor. The parameters, β_{00} and β_{10} represent the fixed intercept and slope, respectively. Meanwhile, β_{01} and β_{11} capture the differences between individuals in the fixed intercept and slope due to variations in the levels of the predictor w_i. In these models, ν_{0i} and ν_{1i} are the Level 2 errors that are assumed to follow a bivariate normal distribution with mean zero and covariance matrix $\mathbf{\Sigma}_\nu$:

$$\mathbf{\Sigma}_\nu = \begin{bmatrix} \sigma^2_{\nu_0} & \sigma_{\nu_{01}} \\ \sigma_{\nu_{01}} & \sigma^2_{\nu_1} \end{bmatrix}$$

Substituting the Level 2 models into the Level 1 equation, we obtain the so-called reduced model:

$$y_{it} = \beta_{00} + \beta_{01} w_i + \beta_{10} \tilde{x}_{it} + \beta_{11} w_i \tilde{x}_{it} + \nu_{0i} + \nu_{1i} \tilde{x}_{it} + \epsilon_{it} \tag{1}$$

The reduced model shows that the coefficient β_{11} represents the cross-level interaction effect between the Level 2 and Level 1 predictors. This allows us to examine whether the time-invariant predictor moderates the influence of the Level 1 predictor.

Following Lafit et al. (2024), we model between-person differences in the distribution of the Level 1 predictor x_{it} by specifying the following model:

$$x_{it} = \mu_x + \upsilon_i + \tau_{it}$$

In this model, μ_x represents the overall mean of the Level 1 predictor across all persons, while υ_i captures the deviation of each person's mean from μ_x. The variance of these deviations is denoted by σ^2_υ. Additionally, τ_{it} represents the within-person error of the predictor x_{it}, which is assumed to be normally distributed with mean zero and variance σ^2_τ. For each person, the mean of x_{it} can be expressed as $\mu_{x_i} = \mu_x + \upsilon_i$. The overall variance of x_{it} for person i is given by $\sigma^2_\upsilon + \sigma^2_\tau$.

We assume that the response variable is measured with errors. Specifically, we assumed the following model for the observed response variable

$$y^*_{it} = \beta_{00} + \beta_{01} w_i + \beta_{10} \tilde{x}_{it} + \beta_{11} w_i \tilde{x}_{it} + \nu_{0i} + \nu_{1i} \tilde{x}_{it} + \rho_\epsilon \epsilon^*_{it-1} + \varepsilon_{it} + \phi_{it}$$

such that ϕ_{it} is normally distributed with mean zero and variance σ^2_ϕ. In this model, the Level 1 errors follow an AR(1) process:

$$\epsilon^*_{it} = \rho_\varepsilon \epsilon^*_{it-1} + \varepsilon_{it} + \phi_{it}$$

We observe that measurement error in the Level 1 errors implies that the variance of ϵ^*_{it} is given by $(\sigma^2_\varepsilon + \sigma^2_\phi)/(1 - \rho^2_\epsilon)$[1]. As a result, measurement error will propagate over time.

Additionally, we assume that the Level 1 and 2 predictors are measured with error, using an additive classical measurement error model (Zhong et al., 2002). Specifically, we specify the following models for the observed variables:

$$x^*_{it} = x_{it} + \delta_{it}$$
$$w^*_i = w_i + \kappa_i$$

[1] To derive the variance of ϵ^*_{it}, we solve the following equation recursively:

$$\begin{aligned}
\epsilon^*_{it} &= \rho_\epsilon \epsilon^*_{it-1} + \varepsilon_{it} + \phi_{it} \\
&= \rho_\epsilon (\rho_\epsilon \epsilon^*_{it-2} + \varepsilon_{it-1} + \phi_{it-1}) + \varepsilon_{it} + \phi_{it} \\
&= \rho^2_\epsilon \epsilon^*_{it-2} + \varepsilon_{it} + \rho_\epsilon \varepsilon_{it-1} + \phi_{it} + \rho_\epsilon \phi_{it-1} \\
&\vdots \\
&= \rho^t_\epsilon \epsilon^*_{i0} + \sum_{s=0}^{t-1} \rho^s_\epsilon (\varepsilon_{it-s} + \phi_{it-s})
\end{aligned}$$

such that δ_{it} is normally distributed with mean zero and variance σ^2_δ, and κ_i follows also a normal distribution with mean zero and variance σ^2_κ. Finally, we assume that Level 1 errors, Level 2 errors, and measurement errors are independent.

2.1 Maximum likelihood estimator of the fixed effects

To derive the Maximum likelihood estimator (MLE, Raudenbush (1997)) of the fixed effects we use general matrix-based notation (Moerbeek, 2008) to represent the model in Equation (1) for each person i in matrix form:

$$\mathbf{y}^*_i = \mathbf{X}^*_i\boldsymbol{\beta} + \mathbf{Z}^*_i\boldsymbol{\nu}_i + \boldsymbol{\epsilon}^*_i$$

where $\mathbf{y}^*_i$ contains the T observations of the observed dependent variable for person i. $\mathbf{X}^*_i$ is the $(T \times 4)$ design matrix for the fixed part given by

$$\mathbf{X}^*_i = \begin{bmatrix} 1 & \tilde{x}^*_{i1} & w^*_i & w^*_i\tilde{x}^*_{i1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \tilde{x}^*_{iT} & w^*_i & w^*_i\tilde{x}^*_{iT} \end{bmatrix}$$

$\mathbf{Z}^*_i$ is the $(T \times 2)$ design vector for the random part

$$\mathbf{Z}^*_i = \begin{bmatrix} 1 & \tilde{x}^*_{i1} \\ \vdots & \vdots \\ 1 & \tilde{x}^*_{iT_i} \end{bmatrix}$$

The vector of fixed regression coefficients is denoted by $\boldsymbol{\beta} = (\beta_{00}, \beta_{10}, \beta_{01}, \beta_{11})^T$ and $\boldsymbol{\nu}_i$ is the vector of random effects given by $\boldsymbol{\nu}_i = (\nu_{0i}, \nu_{1i})^T$.

The $(T \times 1)$ vector $\boldsymbol{\epsilon}^*_i = (\epsilon_{i1} + \phi_{i1}, \ldots, \epsilon_{iT} + \phi_{iT})^T$ includes the Level 1 and the measurement errors in the observed dependent variable for person i. The $(T \times T)$ covariance matrix of the Level 1 errors is given by:

$$\boldsymbol{\Sigma}_{\epsilon^*_i} = \sigma^2_{\epsilon^*}\boldsymbol{\Gamma}_{\epsilon_i} = \sigma^2_{\epsilon^*}\begin{bmatrix} 1 & \rho_\epsilon & \rho^2_\epsilon & \ldots & \rho^{T-1}_\epsilon \\ \rho_\epsilon & 1 & \rho_\epsilon & \ldots & \rho^{T-2}_\epsilon \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1}_\epsilon & \rho^{T-2}_\epsilon & \rho^{T-3}_\epsilon & \ldots & 1 \end{bmatrix}$$

Under the assumption that the expected value and variance of ϵ^*_{i0} are both finite, the steady mean of ϵ^*_{it} is:

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}(\epsilon^*_{it}) &= \rho^t_\epsilon \mathbb{E}(\epsilon^*_{i0}) + \sum_{s=0}^{t-1} \rho^s_\epsilon (\mathbb{E}(\varepsilon_{it-s}) + \mathbb{E}(\phi_{it-s})) \\ &= 0 \end{aligned}$$

Similarly, the steady variance of ϵ^*_{it}, is given by:

$$\begin{aligned} \lim_{t\to\infty} \mathbb{V}(\epsilon^*_{it}) &= \rho^{2t}_\epsilon \mathbb{V}(\epsilon^*_{i0}) + \sum_{s=0}^{t-1} \rho^{2s}_\epsilon (\mathbb{V}(\varepsilon_{it-s}) + \mathbb{V}(\phi_{it-s})) \\ &= (\sigma^2_\varepsilon + \sigma^2_\phi)/(1 - \rho^2_\epsilon) \end{aligned}$$

where $\sigma^2_{\epsilon^*} = (\sigma^2_\varepsilon + \sigma^2_\phi)/(1-\rho^2_\epsilon)$.

Conditioning on the design matrices $\mathbf{X}^*_i$ and $\mathbf{Z}^*_i$, the $(T \times T)$ covariance matrix $(\mathbf{V}_i)$ of the vector $\mathbf{y}^*_i$ for person i is equal to:

$$\mathbf{V}_i := \text{cov}(\mathbf{y}^*_i | \mathbf{X}^*_i, \mathbf{Z}^*_i) = \mathbf{Z}^*_i \boldsymbol{\Sigma}_\nu \mathbf{Z}^{*T}_i + \boldsymbol{\Sigma}^*_{\epsilon_i}$$

Using the matrix representation of the two-level model, we can write the MLE of the fixed effects (i.e., $\boldsymbol{\beta}$) conditioned on the observed matrices $\mathbf{X}^*_i$ and $\mathbf{Z}^*_i$ as follows (see Laird and Ware, 1982):

$$\hat{\boldsymbol{\beta}} | \mathbf{X}^*_i, \mathbf{Z}^*_i = \left(\sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{X}^*_i \right)^{-1} \sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{y}^*_i$$

3 Asymptotic bias of the MLE of the fixed effects

In this section, we derive the asymptotic approximation of the bias of the MLE of fixed effects $\boldsymbol{\beta}$, assuming additive measurement errors in the Level 1 and Level 2 predictors and the response. First, we will calculate the expected value of the MLE estimator of $\boldsymbol{\beta}$ conditioned on the observed matrices $\mathbf{X}^*_i$ and $\mathbf{Z}^*_i$:

$$\begin{aligned}
\mathbb{E}\left(\hat{\boldsymbol{\beta}} | \mathbf{X}^*_i, \mathbf{Z}^*_i\right) &= \mathbb{E}\left[\left(\sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{X}^*_i \right)^{-1} \sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{y}^*_i \middle| \mathbf{X}^*_i, \mathbf{Z}^*_i \right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{X}^*_i \right)^{-1} \sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i (\mathbf{X}^*_i \boldsymbol{\beta}) \middle| \mathbf{X}^*_i, \mathbf{Z}^*_i \right]
\end{aligned}$$

Since we assume that the measurement errors on the predictors are additive, we can write $\mathbb{E}(\hat{\boldsymbol{\beta}})$ as[2]:

$$\begin{aligned}
\mathbb{E}(\hat{\boldsymbol{\beta}}) &= \mathbb{E}\left[\left(\sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{X}^*_i \right)^{-1} \sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{X}_i \right] \boldsymbol{\beta} \\
&= \mathbb{E}\left[\mathbf{P}^{-1}_{\hat{\boldsymbol{\beta}}} \mathbf{Q}_{\hat{\boldsymbol{\beta}}} \right] \boldsymbol{\beta}
\end{aligned} \tag{2}$$

such that $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$ and $\mathbf{Q}_{\hat{\boldsymbol{\beta}}}$ are (4×4) matrices equal to:

$$\mathbf{P}_{\hat{\boldsymbol{\beta}}} = \sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{X}^*_i \tag{3}$$

and

$$\mathbf{Q}_{\hat{\boldsymbol{\beta}}} = \sum_{i=1}^N \mathbf{X}^{*T}_i \mathbf{V}^{-1}_i \mathbf{X}_i \tag{4}$$

[2]To show this, we can represent the matrix $\mathbf{X}^*_i$ as the sum of the unobserved matrix $\mathbf{X}_i$ and a matrix

Lafit et al. (2024) showed that the matrix $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$ can be written as:

$$\mathbf{P}_{\hat{\boldsymbol{\beta}}} = \sum_{i=1}^{N} \begin{bmatrix} \mathbf{E}_i & w_i^* \mathbf{E}_i \\ w_i^* \mathbf{E}_i & w_i^{*2} \mathbf{E}_i \end{bmatrix} \tag{5}$$

where $\mathbf{E}_i$ is a (2×2) matrix:

$$\mathbf{E}_i = \mathbf{C}_i(\mathbf{I} - (\mathbf{I} + \boldsymbol{\Sigma}_v \mathbf{C}_i)^{-1} \Sigma_v \mathbf{C}_i^T)$$

such that $\mathbf{C}_i$ is a (2×2) matrix where the elements are given by:

$$\begin{aligned} c_{11} &= \frac{(T-2)(1-\rho_\epsilon)^2 + 2(1-\rho_\epsilon)}{\sigma^2_{\epsilon^*}(1-\rho^2_\epsilon)} \\ c_{12} = c_{21} &= \frac{\rho_\epsilon(1-\rho_\epsilon)(\tilde{x}^*_{i1} + \tilde{x}^*_{iT})}{\sigma^2_{\epsilon^*}(1-\rho^2_\epsilon)} \\ c_{22} &= \frac{(1+\rho^2_\epsilon)(T-1)S^2_{x_i^*} - 2\rho_\epsilon(T-1)K_{x_i^*} - \rho^2_\epsilon(\tilde{x}^{*2}_{i1} + \tilde{x}^{*2}_{iT})}{\sigma^2_{\epsilon^*}(1-\rho^2_\epsilon)} \end{aligned} \tag{6}$$

where $S^2_{x_i^*}$ and $K_{x_i^*}$ denote the sample variance and sample autocovariance of the person-mean centered observed values of the Level 1 predictor for the i-th person.

Following a procedure similar to that described by Lafit et al. (2024), we provide a step-by-step derivation of the matrix $\mathbf{Q}_{\hat{\boldsymbol{\beta}}}$. Using the Woodbury matrix identity[3] (Hager,

---

including the measurement error denoted by $\mathbf{U}_i$. We can, therefore, write $\mathbb{E}\left(\hat{\boldsymbol{\beta}}|\mathbf{X}_i^*, \mathbf{Z}_i^*\right)$ as:

$$\begin{aligned} \mathbb{E}\left(\hat{\boldsymbol{\beta}}|\mathbf{X}_i^*, \mathbf{Z}_i^*\right) &= \mathbb{E}\left[\left(\sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}(\mathbf{X}_i^*\boldsymbol{\beta}) \middle| \mathbf{X}_i^*, \mathbf{Z}_i^*\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}((\mathbf{X}_i + \mathbf{U}_i)\boldsymbol{\beta}) \middle| \mathbf{X}_i^*, \mathbf{Z}_i^*\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i\boldsymbol{\beta} + \left(\sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{U}_i\boldsymbol{\beta}) \middle| \mathbf{X}_i^*, \mathbf{Z}_i^*\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i\boldsymbol{\beta} \middle| \mathbf{X}_i^*, \mathbf{Z}_i^*\right] + \mathbb{E}\left[\left(\sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{U}_i\boldsymbol{\beta}) \middle| \mathbf{X}_i^*, \mathbf{Z}_i^*\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*T}\mathbf{V}_i^{-1}\mathbf{X}_i\boldsymbol{\beta} \middle| \mathbf{X}_i^*, \mathbf{Z}_i^*\right] \end{aligned}$$

since $\mathbb{E}(\mathbf{U}|\mathbf{X}_i^*, \mathbf{Z}_i^*)$ is zero.

[3]The Woodbury matrix identity provides a way to inverting the sum of a matrices, specifically it states that for an $(n \times n)$ matrix $\mathbf{A}$, $(k \times k)$ matrix $\mathbf{C}$, $(n \times k)$ matrix $\mathbf{U}$, and $(k \times n)$ matrix $\mathbf{V}$, we can write $(\mathbf{A} + \mathbf{UCV})^{-1}$ as $\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})\mathbf{VA}^{-1}$.

1989), we can write the matrix $\mathbf{Q}_{\hat{\boldsymbol{\beta}}}$ as

$$\begin{aligned}
\mathbf{Q}_{\hat{\boldsymbol{\beta}}} &= \sum_{i=1}^{N} \mathbf{X}_i^{*T} \mathbf{V}_i^{-1} \mathbf{X}_i \\
&= \sum_{i=1}^{N} \mathbf{X}_i^{*T} \left(\boldsymbol{\Sigma}_{\epsilon_i^*} + \mathbf{Z}_i^* \boldsymbol{\Sigma}_\nu \mathbf{Z}_i^{*T} \right)^{-1} \mathbf{X}_i \\
&= \sum_{i=1}^{N} \mathbf{X}_i^{*T} \left(\boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} - \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{Z}_i^* \left(\boldsymbol{\Sigma}_\nu^{-1} + \mathbf{Z}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{Z}_i^* \right)^{-1} \mathbf{Z}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \right) \mathbf{X}_i \\
&= \sum_{i=1}^{N} \underbrace{\mathbf{X}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{X}_i}_{=:\boldsymbol{A_i}} - (\underbrace{\mathbf{X}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{Z}_i^*}_{=:\boldsymbol{B_i}}) (\underbrace{\boldsymbol{\Sigma}_\nu^{-1} + \overbrace{\mathbf{Z}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{Z}_i^*}^{=:\boldsymbol{C_i}}}_{=:\boldsymbol{D_i}})^{-1} (\underbrace{\mathbf{Z}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{X}_i}_{=\boldsymbol{F_i}^T})
\end{aligned}$$

Multiplying the two terms of the matrix $\mathbf{D}_i$ by $\boldsymbol{\Sigma}_v$ and $\boldsymbol{\Sigma}_v^{-1}$, the inverse of the matrix $\mathbf{D}_i$ can be written as

$$\mathbf{D}_i^{-1} = (\mathbf{I} + \boldsymbol{\Sigma}_v \mathbf{C}_i)^{-1} \boldsymbol{\Sigma}_v$$

and, consequently, the matrix $\mathbf{Q}_{\hat{\boldsymbol{\beta}}}$ is:

$$\mathbf{Q}_{\hat{\boldsymbol{\beta}}} = \sum_{i=1}^{N} \mathbf{A}_i - \mathbf{B}_i (\mathbf{I} + \boldsymbol{\Sigma}_v \mathbf{C}_i)^{-1} \boldsymbol{\Sigma}_v \mathbf{F}_i^T$$

such that $\boldsymbol{A_i}$ is equal to

$$\mathbf{X}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{X}_i = \begin{bmatrix} \boldsymbol{Z_i}^{*T} \\ w_i^* \boldsymbol{Z_i}^{*T} \end{bmatrix} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \begin{bmatrix} \boldsymbol{Z_i} & w_i \boldsymbol{Z_i} \end{bmatrix} = \left[\begin{array}{c|c} \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \boldsymbol{Z_i} & \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} w_i \boldsymbol{Z_i} \\ \hline w_i^* \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \boldsymbol{Z_i} & w_i^* \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} w_i \boldsymbol{Z_i} \end{array} \right] = \left[\begin{array}{c|c} \boldsymbol{C_i}^* & w_i \boldsymbol{C_i}^* \\ \hline w_i^* \boldsymbol{C_i}^{*T} & w_i^* w_i \boldsymbol{C_i}^* \end{array} \right]$$

and that $\boldsymbol{B_i}$ is equal to

$$\mathbf{X}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \mathbf{Z}_i^* = \begin{bmatrix} \boldsymbol{Z_i}^{*T} \\ w_i^* \boldsymbol{Z_i}^{*T} \end{bmatrix} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \boldsymbol{Z_i}^* = \begin{bmatrix} \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \boldsymbol{Z_i}^* \\ w_i^* \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \boldsymbol{Z_i}^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{C_i} \\ \hline w_i^* \boldsymbol{C_i} \end{bmatrix}.$$

and that $\boldsymbol{F_i}$ is equal to

$$\mathbf{Z}_i^{*T} \boldsymbol{\Sigma}_{\epsilon_i}^{-1} \mathbf{X}_i = \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i}^{-1} \begin{bmatrix} \boldsymbol{Z_i} & w_i \boldsymbol{Z_i} \end{bmatrix} = \left[\boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} \boldsymbol{Z_i} \;\middle|\; \boldsymbol{Z_i}^{*T} \boldsymbol{\Sigma}_{\epsilon_i^*}^{-1} w_i \boldsymbol{Z_i} \right] = \left[\boldsymbol{C_i}^* \;\middle|\; w_i \boldsymbol{C_i}^* \right]$$

We first calculate the matrix $\boldsymbol{C_i}^* = (c_{kl})_i \in \mathbb{R}^{2\times 2}$ matrix which is equal to

$$\boldsymbol{C_i}^* = \begin{bmatrix} 1 & \dots & 1 \\ \tilde{x}_{i1}^* & \dots & \tilde{x}_{iT}^* \end{bmatrix} \frac{1}{\sigma_{\epsilon^*}^2 (1-\rho_\epsilon^2)} \begin{bmatrix} 1 & -\rho_\epsilon & 0 & 0 & \dots & 0 & 0 \\ -\rho_\epsilon & 1+\rho_\epsilon^2 & -\rho_\epsilon & 0 & \dots & 0 & 0 \\ 0 & -\rho_\epsilon & 1+\rho_\epsilon^2 & -\rho_\epsilon & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1+\rho_\epsilon^2 & -\rho_\epsilon \\ 0 & 0 & 0 & 0 & \dots & -\rho_\epsilon & 1 \end{bmatrix} \begin{bmatrix} 1 & \tilde{x}_{i1} \\ \vdots & \vdots \\ 1 & \tilde{x}_{iT} \end{bmatrix}$$

Consequently, we can calculate the four elements of the matrix $\boldsymbol{C_i}^*$ as:

$$c_{11}^* = \frac{(T-2)(1-\rho_\epsilon)^2 + 2(1-\rho_\epsilon)}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)}$$

$$\begin{aligned} c_{12}^* &= \frac{\tilde{x}_{i1} + (1+\rho_\epsilon^2)\sum_{j=2}^{T_i-1}\tilde{x}_{ij} - \rho_\epsilon\sum_{j=1}^{T}\tilde{x}_{ij} - \rho_\epsilon\sum_{j=2}^{T-1}\tilde{x}_{ij} + \tilde{x}_{iT}}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)} \\ &= \frac{(1+\rho_\epsilon)(\tilde{x}_{i1}+\tilde{x}_{iT}) - 2\rho_\epsilon\underbrace{\sum_{j=1}^{T}\tilde{x}_{ij}}_{=0} - (1+\rho_\epsilon^2)(\tilde{x}_{i1}+\tilde{x}_{iT}) + (1+\rho_\epsilon^2)\underbrace{\sum_{j=1}^{T}\tilde{x}_{ij}}_{=0}}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)} \\ &= \frac{\rho_\epsilon(1-\rho_\epsilon)(\tilde{x}_{i1}+\tilde{x}_{iT})}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)} \end{aligned}$$

$$\begin{aligned} c_{21}^* &= \frac{\tilde{x}_{i1}^* + (1+\rho_\epsilon^2)\sum_{j=2}^{T-1}\tilde{x}_{ij}^* - \rho_\epsilon\sum_{j=1}^{T}\tilde{x}_{ij}^* - \rho_\epsilon\sum_{j=2}^{T-1}\tilde{x}_{ij}^* + \tilde{x}_{iT}^*}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)} \\ &= \frac{\rho_\epsilon(1-\rho_\epsilon)(\tilde{x}_{i1}^*+\tilde{x}_{iT}^*)}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)} \end{aligned}$$

$$\begin{aligned} c_{22}^* &= \frac{\tilde{x}_{i1}^*\tilde{x}_{i1} - \rho_\epsilon\sum_{j=1}^{T-1}\tilde{x}_{i,j+1}^*\tilde{x}_{i,j} - \rho_\epsilon\sum_{j=1}^{T-1}\tilde{x}_{ij}^*\tilde{x}_{i,j+1} + (1+\rho_\epsilon^2)\sum_{k=2}^{T-2}\tilde{x}_{ij}^*\tilde{x}_{ij} + \tilde{x}_{iT}^*\tilde{x}_{iT}}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)} \\ &= \frac{(1+\rho_\epsilon^2)\sum_{j=1}^{T}\tilde{x}_{i,j}^*\tilde{x}_{i,j} - \rho_\epsilon\sum_{j=1}^{T-1}\tilde{x}_{i,j+1}^*\tilde{x}_{i,j} - \rho_\epsilon\sum_{j=1}^{T-1}\tilde{x}_{i,j}^*\tilde{x}_{i,j+1} - \rho_\epsilon^2(\tilde{x}_{i1}^*\tilde{x}_{i1} + \tilde{x}_{iT}^*\tilde{x}_{iT})}{\sigma_{\epsilon^*}^2(1-\rho_\epsilon^2)} \end{aligned}$$

Putting the above results together, we find that the matrix $\mathbf{Q}_{\hat{\boldsymbol{\beta}}}$ is equal to

$$\begin{aligned} \mathbf{Q}_{\hat{\boldsymbol{\beta}}} &= \sum_{i=1}^{N}\begin{bmatrix}\mathbf{C}_i^* & w_i\mathbf{C}_i^* \\ w_i^*\mathbf{C}_i^* & w_i^*w_i\mathbf{C}_i^*\end{bmatrix} - \begin{bmatrix}\mathbf{C}_i \\ w_i^*\mathbf{C}_i\end{bmatrix}(\mathbf{I}+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\begin{bmatrix}\mathbf{C}_i^{*T} & w_i\mathbf{C}_i^{*T}\end{bmatrix} \\ &= \sum_{i=1}^{N}\begin{bmatrix}\mathbf{C}_i^* & w_i\mathbf{C}_i^* \\ w_i^*\mathbf{C}_i^* & w_i^*w_i\mathbf{C}_i^*\end{bmatrix} - \begin{bmatrix}\mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T} & w_i\mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T} \\ w_i^*\mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T} & w_i^*w_i\mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T}\end{bmatrix} \\ &= \sum_{i=1}^{N}\begin{bmatrix}\mathbf{C}_i^* - \mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T} & w_i(\mathbf{C}_i^* - \mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T}) \\ w_i^*(\mathbf{C}_i^* - \mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T}) & w_i^*w_i(\mathbf{C}_i^* - \mathbf{C}_i(I+\boldsymbol{\Sigma}_v\mathbf{C}_i)^{-1}\boldsymbol{\Sigma}_v\mathbf{C}_i^{*T})\end{bmatrix} \end{aligned}$$

3.1 Asymptotic bias of the MLE of the fixed effects

Since obtaining a closed-form expression of the expected value of $\hat{\boldsymbol{\beta}}$ given in Equation (2) is not possible, we obtain an asymptotic approximation of the bias of $\hat{\boldsymbol{\beta}}$ by considering the convergence in probability of

$$\text{plim } \mathbf{P}_{\hat{\boldsymbol{\beta}}}^{-1}\mathbf{Q}_{\hat{\boldsymbol{\beta}}}\boldsymbol{\beta} = \tilde{\mathbf{P}}_{\hat{\boldsymbol{\beta}}}^{-1}\tilde{\mathbf{Q}}_{\hat{\boldsymbol{\beta}}}\boldsymbol{\beta} \tag{7}$$

such that $\tilde{\mathbf{P}}_{\hat{\boldsymbol{\beta}}}$ and $\tilde{\mathbf{Q}}_{\hat{\boldsymbol{\beta}}}$ are the asymptotic approximation of the matrices $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$ and $\mathbf{Q}_{\hat{\boldsymbol{\beta}}}$.

First, we focus on obtaining the asymptotic approximation of the matrix $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$. We can show that when T and N increases, the matrix $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$ can be approximated by a (4×4) block matrix $\tilde{\mathbf{P}}_{\hat{\boldsymbol{\beta}}}$ given by the following:

$$\tilde{\mathbf{P}}_{\hat{\boldsymbol{\beta}}} = \begin{bmatrix} N\tilde{\mathbf{E}} & N\mu_{\text{w}}\tilde{\mathbf{E}} \\ N\mu_{\text{w}}\tilde{\mathbf{E}} & N(\sigma_{\text{w}}^2 + \sigma_{\kappa}^2 + \mu_{\text{w}}^2)\tilde{\mathbf{E}} \end{bmatrix} \tag{8}$$

with $\tilde{\mathbf{E}} = (\tilde{\mathbf{C}} - \tilde{\mathbf{C}}(\mathbf{I}_2 + \boldsymbol{\Sigma}_{\nu}\tilde{\mathbf{C}})^{-1}\boldsymbol{\Sigma}_{\nu}\tilde{\mathbf{C}})$ and $\tilde{\mathbf{C}}$ is the asymptotic approximation of the matrix $\mathbf{C}$ in Equation (6) and it is a (4×4) matrix with the following elements:

$$\begin{aligned} \tilde{c}_{11} &= \frac{(T-2)(1-\rho_{\epsilon})^2 + 2(1-\rho_{\epsilon})}{\sigma_{\epsilon^*}^2(1-\rho_{\epsilon}^2)} \\ \tilde{c}_{12} &= \tilde{c}_{21} = 0 \\ \tilde{c}_{22} &= \frac{(1+\rho_{\epsilon}^2)(T-1)(\sigma_{\tau}^2 + \sigma_{\delta}^2)}{\sigma_{\epsilon^*}^2(1-\rho_{\epsilon}^2)} \end{aligned} \tag{9}$$

To obtain this approximation, we write the matrix $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$ in Equation (5) as $N(\frac{1}{N}\mathbf{P}_{\hat{\boldsymbol{\beta}}})$. Next, we assume that N is sufficiently large, thus $\frac{\sum_{i=1}^{N} w_i^*}{N}$ converges in probability to μ_{w} and $\frac{\sum_{i=1}^{N} w_i^{*2}}{N}$[4] converges in probability to $(\sigma_{\text{w}}^2 + \sigma_{\kappa}^2) + \mu_{\text{w}}^2$. Assuming that T is sufficiently large, $S_{x_i^*}^2$ converges in probability to $\sigma_{\tau}^2 + \sigma_{\delta}^2$, and $K_{x_i^*}$ converges in probability to zero since we assumed the Level 1 errors τ_{it} and the measurement error δ_{it} are i.i.d. Next, we will show that as T increases, $\tilde{x}_{i1}^* + \tilde{x}_{iT}^*$ and $\tilde{x}_{i1}^{*2} + \tilde{x}_{iT}^{*2}$ are negligible and can be ignored when obtaining the asymptotic approximation of the matrix $\mathbf{C}_i$. In other words, the matrix $\mathbf{C}_i$ is asymptotically diagonal dominant. This means that when T increases, the off-diagonal elements of the matrix $\mathbf{C}_i$ become negligible compared to the diagonal ones. To show this,

[4] That can be calculated as $\frac{\sum_{i=1}^{N}(w_i^* - \frac{\sum_{i=1}^{N} w_i^*}{N})^2}{N} + \left(\frac{\sum_{i=1}^{N} w_i^*}{N}\right)^2$.

we write the matrix $\mathbf{C}_i$ as $T\left(\frac{1}{T}\mathbf{C}_i\right)$, then we can show that:

$$\begin{aligned}
\lim_{T\to\infty} \frac{\tilde{x}^*_{i1} + \tilde{x}_{iT}{}^*}{T} &= \lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T}\tilde{x}^*_{it} - \frac{1}{T}\sum_{t=2}^{T-1}\tilde{x}^*_{it}\right] \\
&= \lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\mu_x + \upsilon_i + \tau_{it} + \delta_{it} - \underbrace{\frac{\sum_{t=1}^{T} x^*_{it}}{T}}_{\to\mu_x+\upsilon_i}\right)\right. \\
&\quad \left. - \frac{1}{T}\sum_{t=2}^{T-1}\left(\mu_x + \upsilon_i + \tau_{it} + \delta_{it} - \underbrace{\frac{\sum_{t=1}^{T} x^*_{it}}{T}}_{\to\mu_x+\upsilon_i}\right)\right] \\
&\longrightarrow \frac{1}{T}\sum_{t=1}^{T}(\tau_{it} + \delta_{it}) - \frac{1}{T}\sum_{t=2}^{T-1}(\tau_{it} + \delta_{it}) \longrightarrow 0.
\end{aligned}$$

Similarly, we can show that:

$$\begin{aligned}
\lim_{T\to\infty} \frac{\tilde{x}^{*2}_{i1} + \tilde{x}^{*2}_{iT}}{T} &= \lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T}\tilde{x}^{*2}_{it} - \frac{1}{T}\sum_{t=2}^{T-1}\tilde{x}^{*2}_{it}\right] \\
&= \underbrace{\frac{1}{T}\sum_{t=1}^{T}\left(\mu_x + \upsilon_i + \tau_{it} + \delta_{it} - \frac{\sum_{t=1}^{T} x^*_{it}}{T}\right)^2}_{\longrightarrow\mathbb{E}\left(\mu_x+\upsilon_i+\tau_{it}+\delta_{it}-\frac{\sum_{t=1}^{T} x^*_{it}}{T}\right)^2} \\
&\quad - \underbrace{\left(\frac{T-2}{T}\right)}_{\longrightarrow 1}\underbrace{\frac{1}{T-2}\sum_{t=2}^{T-1}\left(\mu_x + \upsilon_i + \tau_{it} + \delta_{it} - \frac{\sum_{t=1}^{T} x^*_{it}}{T}\right)^2}_{\longrightarrow\mathbb{E}\left(\mu_x+\upsilon_i+\tau_{it}+\delta_{it}-\frac{\sum_{t=1}^{T} x^*_{it}}{T}\right)^2} = 0.
\end{aligned}$$

since $\mathbb{E}(\mu_x + \upsilon_i + \tau_{it} - \frac{\sum_{t=1}^{T} x^*_{it}}{T})^2 = (\sigma^2_\tau + \sigma^2_\delta) < \infty$. For large T, we can thus approximate $\mathbf{C}_i$ by the diagonal matrix $\tilde{\mathbf{C}}$ that does not depend on the observed values at time points $t = 1$ and $t = T$ of the Level 1 predictor.

Using these results, we can see that the asymptotic approximation of the matrix $\mathbf{Q}_{\hat{\boldsymbol{\beta}}}$ is given by:

$$\tilde{\mathbf{Q}}_{\hat{\boldsymbol{\beta}}} = \begin{bmatrix} N(\tilde{\mathbf{C}}^* - \tilde{\mathbf{C}}(I + \boldsymbol{\Sigma}_v\tilde{\mathbf{C}})^{-1}\boldsymbol{\Sigma}_v\tilde{\mathbf{C}}^{*T}) & N\mu_{\mathrm{w}}(\tilde{\mathbf{C}}^* - \tilde{\mathbf{C}}(I + \boldsymbol{\Sigma}_v\tilde{\mathbf{C}})^{-1}\boldsymbol{\Sigma}_v\tilde{\mathbf{C}}^{*T}) \\ N\mu_{\mathrm{w}}(\tilde{\mathbf{C}}^* - \tilde{\mathbf{C}}(I + \boldsymbol{\Sigma}_v\tilde{\mathbf{C}})^{-1}\boldsymbol{\Sigma}_v\tilde{\mathbf{C}}^{*T}) & ((N-1)\sigma^2_{\mathrm{w}} + N\mu^2_{\mathrm{w}})(\tilde{\mathbf{C}}^*_i - \tilde{\mathbf{C}}(I + \boldsymbol{\Sigma}_v\tilde{\mathbf{C}})^{-1}\boldsymbol{\Sigma}_v\tilde{\mathbf{C}}^{*T}) \end{bmatrix} \tag{10}$$

such that $\tilde{\mathbf{C}}^*$ is the asymptotic approximation of the matrix $\mathbf{C}^*_i$ and is given by a (2×2)

matrix with the following elements:

$$\begin{aligned}
\tilde{c}^*_{11} &= \frac{(T-2)(1-\rho_\epsilon)^2 + 2(1-\rho_\epsilon)}{\sigma^2_{\epsilon^*}(1-\rho^2_\epsilon)} \\
\tilde{c}^*_{12} &= \tilde{c}^*_{21} = 0 \\
\tilde{c}^*_{22} &= \frac{(1+\rho^2_\epsilon)(T-1)\sigma^2_\tau}{\sigma^2_{\epsilon^*}(1-\rho^2_\epsilon)}
\end{aligned} \tag{11}$$

The results indicate that, assuming the variance components are known, measurement errors in both Level 1 and Level 2 predictors lead to bias in the estimation of the fixed effects within the multilevel model described in Equation (1). However, it is noteworthy that measurement error in the response variable does not introduce bias in estimating the coefficient vector $\boldsymbol{\beta}$.

To illustrate this point, we will demonstrate how measurement error in the Level 1 predictor leads to a downward bias in the estimation of the fixed slope β_{10} and the cross-level interaction effect β_{11}. Furthermore, we will show that measurement error in the Level 2 predictor results in a downward bias when estimating the differences in the fixed intercept associated with the Level 2 predictor (β_{01}) as well as the cross-level interaction effect.

3.2 Illustration

To illustrate the impact of measurement errors on the asymptotic bias of the maximum likelihood estimates (MLE) of the fixed effects in the multilevel model represented in Equation (1), we will refer to the example presented by Lafit et al. (2024) In this example, the Clinical ESM study described in Heininga et al. (2019) was used to investigate whether there were differences in the contemporaneous association between anhedonia (A_{it}) and negative affect (NA_{it}) based on varying levels of depression (Dep_i) in persons diagnosed with Major Depressive Disorder (MDD). To investigate this, the following two-level model was estimated:

$$\text{NA}_{it} = \beta_{00} + \beta_{10}\text{Dep}_i + \beta_{10}\tilde{\text{A}}_{it} + \beta_{11}\text{Dep}_i\tilde{\text{A}}_{it} + \nu_{0i} + \nu_{1i}\tilde{\text{A}}_{it} + \epsilon_{it} \tag{12}$$

where β_{00} is the fixed intercept, β_{01} represents the main effect of depression, β_{10} is the fixed slope of anhedonia, and β_{11} represents the cross-level interaction effect between anhedonia and depression. In this model, we further assume the Level 1 errors follow an AR(1) model.

To model between-person differences in the distribution of anhedonia, we specify a multilevel model for the Level 1 predictor Anhedonia:

$$\text{A}_{it} = \mu_\text{A} + \upsilon_i + \tau_{it} \tag{13}$$

such that μ_A represents the mean of anhedonia across all persons and all time points, υ_i represents the deviation of each person's mean anhedonia level from μ_A, and τ_{it} denotes the within-person error of A_{it}, which is assumed to follow normal distribution with mean zero and variance σ^2_τ.

To examine the impact of measurement error on the MLE of the fixed effects in the multilevel model presented in Equation (12), we conducted a simulation study. In this study, we considered scenarios varying the number of participants (N) and the number of measurement occasions (T) to 20, 50, 100, and 200, assuming that both N and T are equal. Next, we generated 1000 datasets based on the specified model. For each dataset, we estimated the model using restricted maximum likelihood (REML) with the lmer function from the lme4 package (Bates et al., 2015). We then compared the average estimates of the fixed effects across the 1000 replicates with the asymptotic results derived from Equation (7).

To obtain the values of the fixed effects of model (12), we used the Clinical ESM study. This data set includes 38 participants diagnosed with MDD who participated in a 7-day ESM study. During the study, participants responded to questionnaires 10 times a day, which included 27 items measuring various negative emotions, as well as momentary anhedonia. These items were measured using a visual analog scale that ranged from "not at all" on the left to "very much" on the right, producing scores between 0 and 100. Before starting the ESM study, the Quick Inventory of Depressive Symptomatology (QIDS; Rush et al. 2003) was administered to measure the severity of depressive symptoms. For further information about the study, please refer to Heininga et al. (2019). The data is publicly available upon request from the EMOTE database (Kalokerinos et al., 2022): https://emotedatabase.com/datasets/3. Using the ESM data set, we will determine the values of the model parameters:

$$\text{NA}_{it} = 10.37 + 1.10\text{Dep}_i + 0.16\tilde{\text{A}}_{it} + 0.02\text{Dep}_i\tilde{\text{A}}_{it} + \nu_{0i} + \nu_{1i}\tilde{\text{A}}_{it} + \epsilon_{it}$$

with $\sigma_{\nu_0} = 1$, $\sigma_{\nu_1} = 0.10$, $\sigma_{\nu_{01}} = 0.015$, $\rho_\epsilon = 0.2$, $\sigma_\epsilon = 1$, $\mu_{\text{Dep}} = 3.33$, and $\sigma_\tau = 1$. Further, we have

$$\text{A}_{it} = 17.22 + \upsilon_i + \tau_{it} \tag{14}$$

with $\sigma_\upsilon = 1$ and $\sigma_\tau = 1$. We note that we set the standard deviation values of σ_{ν_0}, σ_ϵ, and σ_τ to one to ensure that the results across the three main scenarios (measurement error in Level 1, Level 2, and response) are comparable and not affected by large differences in scale. On the GitLab repository of the project, we include the scripts to run the analyses reported in this article using R (version 4.4.1) (RStudio Team, 2015) (see https://gitlab.kuleuven.be/ppw-okpiv/researchers/u0119584/ApproxPowerIL).

To examine the effect of measurement error on the Level 1 predictor on the bias of the MLE of $\boldsymbol{\beta}$, we use the asymptotic approximation of $\mathbf{P}_{\hat{\boldsymbol{\beta}}}^{-1}\mathbf{Q}_{\hat{\boldsymbol{\beta}}}\boldsymbol{\beta}$ from Equation (7). We vary the standard deviation of the measurement error for the Level 1 predictor, σ_δ, setting it equal to $(1+a)\sigma_\tau$, where a ranges from 0 to 2. Table 1 shows the results of the plim of $\hat{\boldsymbol{\beta}}$ and the average estimates of the fixed effects using REML across 1000 replicates. We can observe that increases in σ_δ result in a downward bias in the estimates of β_{10} and β_{11}, with the values of the estimates converging to zero as the variance of the measurement error increases. In contrast, the estimates of β_{00} and β_{01} are not affected by the presence of measurement error in the Level 1 predictor. Comparing the asymptotic results with the simulation results,

we observe that the impact of measurement error in the Level 1 predictor is similar in the two approaches. We also note that increases in the variance of the measurement error in the time-varying predictors impact the estimation of the variance components, specifically, in the estimates of the standard deviation of the random slope (σ_{ν_1}). Table 2 displays the average estimate of the variance components using REML across 1000 replicates. We observe that increases in the value of σ_δ negatively impact the estimation of the standard deviation of the random slope (σ_{ν_1}), and consequently the covariance estimate between the random intercept and slope ($\sigma_{\nu_{01}}$).

Following a similar approach, we investigate the asymptotic performance of $\hat{\boldsymbol{\beta}}$ when the Level 2 predictor is measured with error. Table 3 shows the result of the plim of $\hat{\boldsymbol{\beta}}$ when $\sigma_\kappa = (1+a)\sigma_{\text{W}}$ with a ranging from 0 to 2. The results indicate that the measurement error in the Level 2 predictor causes a downward bias in the estimates of β_{01} and β_{11}. Table 4 shows the estimation of the variance components across the 1000 replicates. We observe that increases in σ_κ are associated with increases in the estimates of the standard deviation of the random intercept ($\hat{\sigma}_{\nu_0}$).

Table 5 shows that the measurement error in the response variable, when measurement error is autocorrelated, does not cause bias in the $\boldsymbol{\beta}$ estimates. In this case, we set the values of $\sigma_\phi = (1+a)\sigma_\epsilon$ and investigate the performance of the plim of $\hat{\boldsymbol{\beta}}$. As evident from the above results, increases in the values of σ_ϕ do not introduce bias in the MLE of the fixed effects. However, as expected, we observe a positive bias in the estimation of σ_ϵ with higher values of σ_ϕ (see Table 6).

Finally, we investigated the asymptotic bias when σ_δ, σ_κ and σ_ϕ are equal to one. The asymptotic approximations of $(\hat{\beta}_{00}, \hat{\beta}_{01}, \hat{\beta}_{10}, \hat{\beta}_{11})$ are $(10.370, 0.550, 0.080, 0.005)$. The simulation results also yield comparable results when T and N are equal to 200 (i.e., $(\hat{\beta}_{00}, \hat{\beta}_{01}, \hat{\beta}_{10}, \hat{\beta}_{11})$ are $(10.374, 0.547, 0.080, 0.005)$ across 1000 replicates). These results indicate that measurement error in the Level 1 and Level 2 predictors negatively impacts the estimation of β_{01}, β_{10}, and β_{11}.

4 The impact of measurement error on statistical power

We now turn our attention to the effect of measurement error on statistical power, particularly when our objective is to test hypotheses related to fixed effects $\boldsymbol{\beta}$. We specifically examine the impact of measurement error while testing the hypothesis concerning the cross-level interaction effect within the multilevel model represented in Equation (1). In this context, we focus on the hypothesis $H_0 : \beta_{11} \leq 0$ versus the alternative hypothesis $H_1 : \beta_{11} > 0$. To assess this hypothesis, we conduct a t-test, where the test statistic is calculated as the ratio of the fixed effect estimate $\hat{\beta}_{11}$ to the standard error estimate $\widehat{\text{SE}}(\hat{\beta}_{11})$:

$$T = \frac{\hat{\beta}_{11}}{\widehat{\text{SE}}(\hat{\beta}_{11})}$$

the t-test approximately follows a t-distribution with $NT - N - 2$ degrees of freedom (Raudenbush and Bryk, 2002).

In this model, the standard error in the denominator of the t-test can be estimated using the square root of the diagonal elements of the covariance of $\hat{\boldsymbol{\beta}}$ conditioned on $\mathbf{X}^*$ and $\mathbf{Z}^*$:

$$\mathbb{V}(\hat{\boldsymbol{\beta}}|\mathbf{X}^*, \mathbf{Z}^*) = \left(\sum_{i=1}^{N} \mathbf{X}_i^{*T} \mathbf{V}_i^{-1} \mathbf{X}_i^* \right)^{-1}$$

which can be calculated as the inverse of the matrix $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$. Using the t-test, H_0 is rejected if it is larger than the $1-\alpha$ quantile of the t-distribution with $NT-N-2$ degrees of freedom.

For this one-sided test, the statistical power is defined as:

$$\mathcal{P}(\alpha, \beta^*) := \Pr\left(\frac{\hat{\beta}_{11}}{\widehat{\mathrm{SE}}(\hat{\beta}_{11})} \geq \mathcal{T}_{\mathrm{r}}^{-1}(1-\alpha) \middle| \alpha, \beta_{11} = \beta_{11}^* \right)$$

with $\mathcal{T}_r^{-1}(\cdot)$ denoting the quantile function of a Student's t-distribution with $r = TN-N-2$ degrees of freedom. Since we are assuming that the number of repeated measurements is the same among all persons and equal to T, we can use the following asymptotic result (see Snijders and Bosker, 1993, for a proof):

$$\mathcal{P}(\alpha, \beta^*) \approx 1 - \mathcal{T}_r \left(\mathcal{T}_r^{-1}(1-\alpha) - \frac{\beta_{11}^*}{\mathrm{SE}(\hat{\beta}_{11})} \right)$$

where $\mathrm{SE}(\hat{\beta}_{11})$ is the asymptotic approximation of the standard error of the MLE of $\hat{\beta}_{11}$. Following the results in Lafit et al. (2024) the asymptotic approximation of the matrix $\mathbf{P}_{\hat{\boldsymbol{\beta}}}$ is given by the matrix $\tilde{\mathbf{P}}_{\hat{\boldsymbol{\beta}}}$ in Equation (8).

To calculate the statistical power for testing the hypothesis that β_{11} is positive, we need to specify the value of the cross-level interaction effect under the alternative hypothesis (β_{11}^*) and determine the values of all other model parameters required for computing the standard error of $\hat{\beta}_{11}$. As discussed in the previous section, measurement errors in both Level 1 and Level 2 predictors can lead to a downward bias in the estimation of the cross-level interaction effect. Additionally, measurement error affects the asymptotic approximation of the standard error of $\hat{\beta}_{11}$. Therefore, it is crucial to consider this when determining the β_{11}^* value for calculating statistical power.

To assess the impact of measurement error on statistical power, we adjust the expected value of the effect of interest under the alternative hypothesis, β_{11}^*, based on the assumption that variables are measured without error. Consequently, the expected value of the cross-level interaction effect that we anticipate finding in the population, while accounting for measurement error, can be computed using the asymptotic approximation of the bias of the MLE of the fixed effects, as defined in Equation (7):

$$\tilde{\boldsymbol{\beta}}^* = \tilde{\mathbf{P}}_{\hat{\boldsymbol{\beta}}} \tilde{\mathbf{Q}}_{\hat{\boldsymbol{\beta}}} \boldsymbol{\beta}^* \tag{15}$$

where $\boldsymbol{\beta}^* = (\beta_{00}^*, \beta_{01}^*, \beta_{10}^*, \beta_{11}^*)^T$ is the vector of fixed effects that is expected if the alternative hypothesis is true and variables are measured without error. Thus, statistical power to test

if the cross-level interaction effect is positive when taking into account measurement error in the variables can be calculated using the following formula:

$$\mathcal{P}(\alpha, \beta^*) \approx 1 - \mathcal{T}_r \left(\mathcal{T}_r^{-1}(1-\alpha) - \frac{\tilde{\beta}_{11}^*}{\text{SE}(\hat{\beta}_{11})} \right)$$

with $\tilde{\beta}_{11}^*$ calculated using Equation (15).

4.1 Illustration

We revisit the example from the Clinical ESM study to demonstrate the impact of measurement error on statistical power when conducting hypothesis testing of the cross-level interaction effect. We will calculate the statistical power to evaluate the hypothesis that the cross-level interaction effect between anhedonia and depression is positive (i.e., $H_1 : \beta_{11} > 0$). For this test, we set the significance level (α) at 5% and N and T to 20, 50, 100, and 200 (i.e., assuming N and T are equal). We compare the performance of the analytic approximation with that of the simulation study, which includes 1000 replicates. The R script for the analyses presented in this article is available on the project's GitLab repository. The results on the impact of statistical power on the rest of the fixed effects are included in the supplementary material.

Table 7 presents the estimates of statistical power for testing the hypothesis $H_1 : \beta_{11} > 0$ as well as the standard error of $\hat{\beta}_{11}$. These estimates are derived using both the asymptotic approximation and Monte Carlo simulation, under the assumption that only the Level 1 predictor contains measurement error. Our findings indicate that measurement error in the Level 1 predictor has a significant impact on statistical power. In particular, larger values of σ_δ lead to a reduction in statistical power for the one-sided hypothesis related to the cross-level interaction effect. Furthermore, increasing either the sample size N or the number of time points T does not seem to enhance statistical power when σ_δ is large. When comparing the asymptotic approximation with the simulation results, we notice that the statistical power is slightly higher in the Monte Carlo simulation. This discrepancy can be explained by the increase in the variance of the observed person-mean centered Level 1 predictor (i.e., $\sigma_\tau^2 + \sigma_\delta^2$), which decreases $\text{SE}(\hat{\beta}_{11})$ in the simulation study compared to the asymptotic approximation. The simulation results show that the differences in statistical power arising from measurement errors in the Level 1 predictor are primarily due to the downward bias in estimating $\tilde{\beta}_{11}^*$ and the downward bias in calculating $\text{SE}(\hat{\beta}_{11})$.

Next, we examine the impact of measurement error in the Level 2 predictor on statistical power. Table 8 illustrates how the statistical power to test the hypothesis $H_1 : \beta_{11} > 0$ varies as a function of σ_κ, T, and N. The results of the simulation study align with the asymptotic results. Furthermore, results indicate that statistical power is negatively associated with the variance of the measurement error in the Level 2 predictor. By analyzing the performance of $\text{SE}(\hat{\beta}_{11})$, we observe that measurement error produces a negative bias in $\tilde{\beta}_{11}^*$, which is explained by increasing in the variance of the observed Level 2 predictor (i.e., $\sigma_{\text{w}}^2 + \sigma_\kappa^2$).

Finally, Table 9 presents the estimates of statistical power for testing the hypothesis $H_1 : \beta_{11} > 0$ as well as the estimates of the standard error of $\hat{\beta}_{11}$ under the assumption of autocorrelated measurement error in the response. We observe that the asymptotic approximation and simulation study provide similar results, with statistical power decreasing as a function of σ_ϕ. As demonstrated in the previous section, measurement error in the dependent variable does not affect $\tilde{\beta}^*_{11}$. However, SE($\hat{\beta}_{11}$) is positively related to σ_ϕ, negatively impacting statistical power.

5 Discussion

In this study, we examined the asymptotic performance of the MLE of the fixed effect in a two-level model where variables are measured with error and include autocorrelated Level 1 errors. This model incorporates both time-varying and time-invariant predictors, as well as their interaction effect. Starting from a classical additive measurement error model, we found that the MLE of the fixed effects is negatively asymptotically biased when both Level 1 and Level 2 variables are measured with errors. We also showed that the MLE of the fixed effect is not consistent, as it does not converge in probability to the true parameter value. This finding aligns with previous results observed in generalized linear mixed-effects models Zhong et al. (2002). Conversely, we demonstrated that when the response variable is measured with error, the MLE of the fixed effect is asymptotically unbiased. However, it is important to note that the standard error of the MLE of the fixed effect increases with the variance of the measurement error.

We also examined how measurement error affects statistical power during null hypothesis testing of cross-level interaction effects. Our findings indicate that power decreases when both predictors and the response variable are subject to measurement error. Specifically, when the predictors are measured inaccurately, this leads to a negative bias, causing the estimated fixed cross-level interaction effect to approach zero as the measurement error variance increases. Additionally, when there is measurement error in the response variable, an increase in the variance of this error negatively impacts statistical power, particularly when the number of repeated measurements and the sample size are small.

The current study examines the impact of measurement error on the statistical power of multilevel models applied to IL data. However, several limitations must be considered when interpreting the findings. First, we derived the asymptotic approximations for the bias and standard errors, based on the assumption of additive measurement error and a model that includes both Level 1 and Level 2 continuous predictors. We treated measurement error on the response variable y_{it} as an additive effect on the Level 1 error term ϵ_{it}. The Level 1 measurement errors ϵ_{it} are modeled with an AR(1) structure. This means that these measurement errors influence subsequent time points in the same manner as the within-person errors that arise from sampling error. If we consider measurement errors to have strictly transient effects (for instance, as discussed in Schuurman et al. (2015)), we would need to adopt a different model structure, which is beyond the scope of this paper. The AR(1) assumption for Level 1 errors further indicates that the effects of the Level 1 predictor

x_{it} do not persist over time (see Ariens et al. (2023)). In contrast, more general models allow the effects of Level 1 predictors to carry over time by including, for example, the lagged values of time-varying predictors (i.e., x_{it-1}). Such models are explored in Ariens et al. (2025). In these models, measurement error in x_{it} can also propagate over time, leading to biases in estimating lagged effects. While it is theoretically possible to investigate the consequences of measurement errors in Level 1 predictors using the methods outlined in this manuscript, such an analysis falls outside the scope of the current paper.

Furthermore, our focus was solely on the impact of measurement error on fixed effect estimates. It is essential to note that measurement error also affects the estimation of variance components in multilevel models (Wang et al., 1998; Zare et al., 2012; Zhong et al., 2002). To fully interpret the results found in the simulation studies, we need to understand the asymptotic performance of the MLE for the variance of Level 1 errors and random effects. Lastly, we concentrated exclusively on multilevel models with only one Level 1 predictor. In deriving the asymptotic approximation of the bias, we started with the results from Lafit et al. (2024), which provided the asymptotic approximation of the standard errors in multilevel models without measurement error. Lafit et al. (2024) also derived the precision matrix of the MLE of the fixed effects for a model with p predictors. Moreover, we assumed that the Level 1 predictor follows a linear mixed-effect model with i.i.d. normally distributed within-person errors. Future studies can build on these results to derive the asymptotic performance of the MLE of fixed effects for multilevel models with multiple time-varying predictors or by relaxing the distribution assumptions of the time-varying predictor to accommodate different types of distributions (e.g., categorical predictors, right-skewed within-person errors).

6 Conclusion

This study shows that the MLE of fixed effects in multilevel models with AR(1) within-person errors is inconsistent and negatively biased when both Level 1 and 2 variables contain measurement errors. Conversely, the MLE remains asymptotically unbiased when only the response variable is measured with error. In both scenarios, measurement error adversely affects statistical power when testing cross-level interaction effects.

Author Contributions

The authors made the following contributions. Ginette Lafit: Conceptualization, Formal Analysis, Methodology, Software, Visualization, Writing - Original Draft Preparation, Writing -Review & Editing; Sigert Ariens: Formal Analysis, Methodology, Writing -Review & Editing & Editing; Richard Artner: Formal Analysis, Methodology, Writing -Review & Editing.

Conflicts of Interest

The authors declare no conflicts of interest concerning the authorship or the publication of this article.

Supplemental Material

The supplemental material for this paper is available at the GitLab repository in the folder *Measurement Error in Multilevel Models*: https://gitlab.kuleuven.be/ppw-okpiv/researchers/u0119584/ApproxPowerIL.

Funding

The research presented in this article was supported by research grants from the Research Council of KU Leuven (STG/23/050, C+/24/002, C16/23/011) awarded to G. Lafit and the Fund for Scientific Research - Flanders (FWO; 1278525N) awarded to Sigert Ariens.

References

Arend, M. G. and Schäfer, T. (2019). Statistical power in two-level models: A tutorial based on Monte Carlo simulation. *Psychological Methods*, 24(1):1–19.

Ariens, S., Adolf, J. K., and Ceulemans, E. (2023). One does not simply correct for serial dependence. *Structural Equation Modeling: A Multidisciplinary Journal*, 30(5):807–821.

Ariens, S., Vanhasbroeck, N., Lafit, G., and Ceulemans, E. (2025). Interpreting parameters of dynamic regression models. Preprint.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Brose, A., Schmiedek, F., Gerstorf, D., and Voelkle, M. C. (2020). The measurement of within-person affect variation. *Emotion*, 20(4):677.

Buonaccorsi, J., Demidenko, E., and Tosteson, T. (2000). Estimation in longitudinal random effects models with measurement error. *Statistica Sinica*, pages 885–903.

Dejonckheere, E., Demeyer, F., Geusens, B., Piot, M., Tuerlinckx, F., Verdonck, S., and Mestdagh, M. (2022). Assessing the reliability of single-item momentary affective measurements in experience sampling. *Psychological assessment*, 34(12):1138.

Eisele, G., Lafit, G., Vachon, H., Kuppens, P., Houben, M., Myin-Germeys, I., and Viechtbauer, W. (2021). Affective structure, measurement invariance, and reliability across different experience sampling protocols. *Journal of Research in Personality*, 92:104094.

Goldstein, H., Healy, M. J., and Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, 13(16):1643–1655.

Hager, W. W. (1989). Updating the inverse of a matrix. *SIAM Review*, 31(2):221–239.

Hamilton, J. D. (1994). *Time series analysis*, volume 2. Princeton New Jersey.

Heininga, V. E., Dejonckheere, E., Houben, M., Obbels, J., Sienaert, P., Leroy, B., van Roy, J., and Kuppens, P. (2019). The dynamical signature of anhedonia in major depressive disorder: positive emotion dynamics, reactivity, and recovery. *BMC Psychiatry*, 19(1):59.

Kalokerinos, E., Russo-Batterham, D., Koval, P., Moeck, E. K., Shrestha, K. M., Garrett, P., Greenaway, Katharine Michalewicz, A., van Os, D., Garber, J., and Kuppens, P. (2022). The emote database: An open, searchable database of experience sampling data mapping everyday life. unpublished.

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., and Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, 4(1):2515245920978738.

Lafit, G., Artner, R., and Ceulemans, E. (2024). Enabling analytical power calculations for multilevel models with autocorrelated errors through deriving and approximating the precision matrix. *Behavior Research Methods*, pages 1–27.

Lafit, G., Sels, L., Adolf, J. K., Loeys, T., and Ceulemans, E. (2022). Powerlapim: An application to conduct power analysis for linear and quadratic longitudinal actor–partner interdependence models in intensive longitudinal dyadic designs. *Journal of Social and Personal Relationships*, 39(10):3085–3115.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, pages 963–974.

Lane, S. P. and Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships*, 35(1):7–31.

Mathieu, J. E., Aguinis, H., Culpepper, S. A., and Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, 97(5):951–966.

Mestdagh, M. and Dejonckheere, E. (2021). Ambulatory assessment in psychopathology research: Current achievements and future ambitions. *Current Opinion in Psychology*.

Moerbeek, M. (2008). Powerful and cost-efficient designs for longitudinal intervention studies with two treatment groups. *Journal of Educational and Behavioral Statistics*, 33(1):41–61.

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., and Reininghaus, U. (2018). Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry*, 17(2):123–132.

Myin-Germeys, I. and Kuppens, P. (2022). *The open handbook of experience sampling methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies*. Myin-Germeys, Inez and Kuppens, Peter.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*, 2(2):173–185.

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

RStudio Team (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., et al. (2003). The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-c), and self-report (QIDS-sr): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, 54(5):573–583.

Schuurman, N. K. and Hamaker, E. L. (2019). Measurement error and person-specific reliability in multilevel autoregressive modeling. *Psychological methods*, 24(1):70.

Schuurman, N. K., Houtveen, J. H., and Hamaker, E. L. (2015). Incorporating measurement error in n= 1 psychological autoregressive modeling. *Frontiers in psychology*, 6:1038.

Snijders, T. A. and Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational Statistics*, 18(3):237–259.

Tosteson, T. D., Buonaccorsi, J. P., and Demidenko, E. (1998). Covariate measurement error and the estimation of random effect parameters in a mixed model for longitudinal data. *Statistics in Medicine*, 17(17):1959–1971.

Wang, N., Lin, X., Gutierrez, R. G., and Carroll, R. J. (1998). Bias analysis and SIMEX approach in generalized linear mixed measurement error models. *Journal of the American Statistical Association*, 93(441):249–261.

Zare, K., Rasekh, A., and Rasekhi, A. A. (2012). Estimation of variance components in linear mixed measurement error models. *Statistical Papers*, 53:849–863.

Zhong, X.-P., Fung, W.-K., and Wei, B.-C. (2002). Estimation in linear models with random effects and errors-in-variables. *Annals of the Institute of Statistical Mathematics*, 54:595–606.

Table 1: Plim of $\hat{\boldsymbol{\beta}}$ and average estimates using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the Level 1 predictor.

			$\beta_{00} = 10.37$		$\beta_{01} = 1.1$		$\beta_{10} = 0.16$		$\beta_{11} = 0.02$	
N	T	σ_δ	Plim	REML	Plim	REML	Plim	REML	Plim	REML
20	20	0	10.370	10.380	1.100	1.100	0.160	0.158	0.020	0.018
50	50	0	10.370	10.374	1.100	1.100	0.160	0.159	0.020	0.020
100	100	0	10.370	10.362	1.100	1.099	0.160	0.160	0.020	0.020
200	200	0	10.370	10.369	1.100	1.095	0.160	0.160	0.020	0.020
20	20	0.5	10.370	10.366	1.100	1.101	0.128	0.127	0.016	0.017
50	50	0.5	10.370	10.372	1.100	1.097	0.128	0.128	0.016	0.016
100	100	0.5	10.370	10.370	1.100	1.098	0.128	0.128	0.016	0.017
200	200	0.5	10.370	10.369	1.100	1.102	0.128	0.128	0.016	0.016
20	20	2	10.370	10.379	1.100	1.100	0.032	0.033	0.004	0.004
50	50	2	10.370	10.370	1.100	1.105	0.032	0.032	0.004	0.004
100	100	2	10.370	10.368	1.100	1.099	0.032	0.032	0.004	0.004
200	200	2	10.370	10.368	1.100	1.105	0.032	0.032	0.004	0.004

Table 2: Average estimates of variance components using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the Level 1 predictor.

N	T	σ_δ	σ_ϵ=1	ρ_ϵ=0.20	$\sigma_{\nu_0} = 1$	$\sigma_{\nu_1} = 0.10$	$\sigma_{\nu_{01}} = 0.015$
20	20	0	1.001	0.199	0.989	0.102	0.009
50	50	0	0.999	0.200	0.996	0.097	0.014
100	100	0	1.000	0.200	0.999	0.098	0.015
200	200	0	1.000	0.200	1.000	0.099	0.015
20	20	0.5	1.000	0.197	0.988	0.092	0.010
50	50	0.5	1.002	0.198	1.000	0.076	0.010
100	100	0.5	1.004	0.199	0.998	0.080	0.012
200	200	0.5	1.004	0.199	0.998	0.080	0.012
20	20	2	1.008	0.192	0.975	0.038	0.001
50	50	2	1.012	0.193	0.994	0.023	0.002
100	100	2	1.014	0.195	0.998	0.018	0.002
200	200	2	1.015	0.194	0.997	0.017	0.002

Table 3: Plim of $\hat{\boldsymbol{\beta}}$ and average estimates using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the Level 2 predictor.

			$\beta_{00} = 10.37$		$\beta_{01} = 1.1$		$\beta_{10} = 0.16$		$\beta_{11} = 0.02$	
N	T	σ_κ	Plim	REML	Plim	REML	Plim	REML	Plim	REML
20	20	0	10.370	10.369	1.100	1.105	0.160	0.161	0.020	0.016
50	50	0	10.370	10.371	1.100	1.097	0.160	0.161	0.020	0.020
100	100	0	10.370	10.369	1.100	1.095	0.160	0.160	0.020	0.020
200	200	0	10.370	10.371	1.100	1.101	0.160	0.160	0.020	0.020
20	20	0.5	10.370	10.372	0.880	0.883	0.160	0.160	0.016	0.017
50	50	0.5	10.370	10.374	0.880	0.879	0.160	0.160	0.016	0.016
100	100	0.5	10.370	10.369	0.880	0.873	0.160	0.161	0.016	0.016
200	200	0.5	10.370	10.367	0.880	0.883	0.160	0.160	0.016	0.016
20	20	2	10.370	10.374	0.220	0.215	0.160	0.160	0.004	0.003
50	50	2	10.370	10.373	0.220	0.218	0.160	0.160	0.004	0.004
100	100	2	10.370	10.375	0.220	0.223	0.160	0.161	0.004	0.004
200	200	2	10.370	10.370	0.220	0.221	0.160	0.160	0.004	0.004

Table 4: Average estimates of variance components using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the Level 2 predictor.

N	T	σ_κ	$\sigma_\epsilon = 1$	$\rho_\epsilon = 0.20$	$\sigma_{\nu_0} = 1$	$\sigma_{\nu_1} = 0.10$	$\sigma_{\nu_{01}} = 0.015$
20	20	0	0.996	0.199	0.984	0.106	0.011
50	50	0	0.998	0.200	0.994	0.095	0.015
100	100	0	1.000	0.200	0.996	0.099	0.015
200	200	0	1.000	0.200	1.001	0.100	0.015
20	20	0.5	0.999	0.202	1.098	0.103	0.012
50	50	0.5	1.000	0.200	1.110	0.096	0.018
100	100	0.5	1.000	0.200	1.109	0.099	0.018
200	200	0.5	1.000	0.200	1.114	0.100	0.019
20	20	2	0.998	0.200	1.382	0.110	0.025
50	50	2	1.001	0.202	1.396	0.099	0.032
100	100	2	1.000	0.199	1.396	0.100	0.032
200	200	2	1.000	0.200	1.400	0.102	0.033

Table 5: Plim of $\hat{\boldsymbol{\beta}}$ and average estimates using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the response.

			$\beta_{00} = 10.37$		$\beta_{01} = 1.1$		$\beta_{10} = 0.16$		$\beta_{11} = 0.02$	
N	T	σ_ϕ	Plim	REML	Plim	REML	Plim	REML	Plim	REML
20	20	0	10.370	10.364	1.100	1.098	0.160	0.162	0.020	0.021
50	50	0	10.370	10.370	1.100	1.096	0.160	0.158	0.020	0.019
100	100	0	10.370	10.371	1.100	1.098	0.160	0.160	0.020	0.019
200	200	0	10.370	10.367	1.100	1.099	0.160	0.160	0.020	0.020
20	20	0.5	10.370	10.354	1.100	1.091	0.160	0.162	0.020	0.019
50	50	0.5	10.370	10.373	1.100	1.098	0.160	0.159	0.020	0.021
100	100	0.5	10.370	10.372	1.100	1.103	0.160	0.160	0.020	0.020
200	200	0.5	10.370	10.370	1.100	1.099	0.160	0.159	0.020	0.020
20	20	2	10.370	10.363	1.100	1.098	0.160	0.161	0.020	0.019
50	50	2	10.370	10.363	1.100	1.101	0.160	0.158	0.020	0.016
100	100	2	10.370	10.369	1.100	1.096	0.160	0.160	0.020	0.020
200	200	2	10.370	10.374	1.100	1.102	0.160	0.160	0.020	0.020

Table 6: Average estimates of variance components using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the response.

N	T	σ_ϕ	$\sigma_\epsilon = 1$	$\rho_\epsilon = 0.20$	$\sigma_{\nu_0} = 1$	$\sigma_{\nu_1} = 0.10$	$\sigma_{\nu_{01}} = 0.015$
20	20	0	0.999	0.198	0.987	0.106	0.012
50	50	0	1.000	0.199	0.995	0.097	0.014
100	100	0	1.000	0.200	0.995	0.099	0.015
200	200	0	1.000	0.200	0.999	0.100	0.015
20	20	0.5	1.118	0.197	0.988	0.111	0.010
50	50	0.5	1.122	0.199	0.989	0.097	0.016
100	100	0.5	1.122	0.199	0.993	0.099	0.014
200	200	0.5	1.123	0.200	0.997	0.099	0.014
20	20	2	2.262	0.197	0.977	0.187	0.007
50	50	2	2.270	0.199	0.984	0.113	0.007
100	100	2	2.274	0.199	0.997	0.085	0.008
200	200	2	2.273	0.200	0.999	0.092	0.012

Table 7: Statistical power for testing $H_1 : \beta_{11} > 0$ and standard error of $\hat{\beta}_{11}$ for different values of σ_δ, N, and T using the assymptotic approximations (i.e., Plim) and average estimates using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the Level 1 predictor.

			Power		Standard Error	
N	T	σ_δ	Plim	REML	Plim	REML
20	20	0	0.099	0.096	0.056	0.060
50	50	0	0.206	0.215	0.024	0.025
100	100	0	0.416	0.410	0.014	0.014
200	200	0	0.754	0.755	0.009	0.009
20	20	0.5	0.092	0.088	0.051	0.053
50	50	0.5	0.174	0.209	0.023	0.021
100	100	0.5	0.330	0.392	0.013	0.012
200	200	0.5	0.611	0.712	0.008	0.007
20	20	2	0.064	0.060	0.032	0.026
50	50	2	0.080	0.094	0.017	0.010
100	100	2	0.100	0.229	0.011	0.005
200	200	2	0.135	0.500	0.007	0.003

Table 8: Statistical power for testing $H_1 : \beta_{11} > 0$ and standard error of $\hat{\beta}_{11}$ for different values of σ_κ, N, and T using the assymptotic approximations (i.e., Plim) and average estimates using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the Level 2 predictor.

			Power		Standard Error	
N	T	σ_κ	Plim	REML	Plim	REML
20	20	0	0.099	0.083	0.056	0.060
50	50	0	0.206	0.211	0.024	0.025
100	100	0	0.416	0.417	0.014	0.014
200	200	0	0.754	0.762	0.009	0.009
20	20	0.5	0.093	0.088	0.050	0.054
50	50	0.5	0.182	0.195	0.022	0.022
100	100	0.5	0.358	0.362	0.012	0.013
200	200	0.5	0.670	0.658	0.008	0.008
20	20	2	0.069	0.053	0.025	0.027
50	50	2	0.101	0.105	0.011	0.011
100	100	2	0.158	0.174	0.006	0.006
200	200	2	0.274	0.270	0.004	0.004

Table 9: Statistical power for testing $H_1 : \beta_{11} > 0$ and standard error of $\hat{\beta}_{11}$ for different values of σ_ϕ, N, and T using the assymptotic approximations (i.e., Plim) and average estimates using restricted maximum likelihood (REML) across 1000 replicates assuming measurement error in the response.

			Power		Standard Error	
N	T	σ_ϕ	Plim	REML	Plim	REML
20	20	0	0.099	0.091	0.056	0.060
50	50	0	0.206	0.194	0.024	0.025
100	100	0	0.416	0.398	0.014	0.014
200	200	0	0.754	0.731	0.009	0.009
20	20	0.5	0.094	0.082	0.061	0.067
50	50	0.5	0.189	0.213	0.026	0.027
100	100	0.5	0.384	0.399	0.015	0.015
200	200	0.5	0.725	0.699	0.009	0.009
20	20	2	0.070	0.044	0.117	0.132
50	50	2	0.112	0.096	0.047	0.049
100	100	2	0.206	0.209	0.024	0.025
200	200	2	0.455	0.441	0.013	0.013