






**Mixtures of Berkson and classical covariate measurement error in the linear mixed model: Bias analysis and application to a study on ultrafine particles**


Journal:	<i>Biometrical Journal</i>
Manuscript ID	bimj.201600188.R1
Wiley - Manuscript type:	Research Paper
Date Submitted by the Author:	28-Nov-2016
Complete List of Authors:	Deffner, Veronika; Ludwig-Maximilians-Universitat Munchen, Institute of Statistics Küchenhoff, Helmut; Ludwig-Maximilians-Universitat Munchen, Institute of Statistics Breitner, Susanne; Helmholtz Zentrum Munchen Deutsches Forschungszentrum für Umwelt und Gesundheit, Institute of Epidemiology II, German Research Center for Environmental Health (GmbH) Schneider, Alexandra; Helmholtz Zentrum Munchen Deutsches Forschungszentrum für Umwelt und Gesundheit, Institute of Epidemiology II, German Research Center for Environmental Health (GmbH) Cyrus, Josef; Helmholtz Zentrum Munchen Deutsches Forschungszentrum für Umwelt und Gesundheit, Institute of Epidemiology II, German Research Center for Environmental Health (GmbH); Universität Augsburg, Environmental Science Center Peters, Annette; Helmholtz Zentrum Munchen Deutsches Forschungszentrum für Umwelt und Gesundheit, Institute of Epidemiology II, German Research Center for Environmental Health (GmbH)
Keywords:	Longitudinal data analysis, Measurement error, Mixed model, Particulate matter
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Tab1.tex Tab2.tex Tab3.tex Tab4.tex	

## Mixtures of Berkson and classical covariate measurement error in the linear mixed model: Bias analysis and application to a study on ultrafine particles

Veronika Deffner <sup>\*,1</sup>, Helmut Küchenhoff <sup>1</sup>, Susanne Breitner , Alexandra Schneider <sup>2</sup>, Josef Cyrus  and Annette Peters <sup>2</sup>

<sup>1</sup> Statistical Consulting Unit, Department of Statistics, Ludwig-Maximilians-Universität, Akademiestr. 1, 80799 Munich, Germany

<sup>2</sup>  Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

 Environmental Science Center, Universität Augsburg, Universitätsstr. 1a, 86159 Augsburg, Germany

Received zzz, revised zzz, accepted zzz

The ultrafine particle measurements in the Augsburgger Umweltstudie, a panel study conducted in Augsburg, Germany, exhibit measurement error from various sources. Measurements of mobile devices show classical possibly individual-specific measurement error; Berkson-type error, which may also vary individually, occurs, if measurements of fixed monitoring stations are used. The combination of fixed site and individual exposure measurements results in a mixture of the two error types. We extended existing bias analysis approaches to linear mixed models with a complex error structure including individual-specific error components, autocorrelated errors and a mixture of classical and Berkson error. Theoretical considerations and simulation results show, that autocorrelation may severely change the attenuation of the effect estimations. Furthermore, unbalanced designs and the inclusion of confounding variables influence the degree of attenuation. Bias correction with the method of moments using mixture data partially yielded better results compared to the usage of incomplete data with classical error. Confidence intervals (CIs) based on the delta method achieved better coverage probabilities than those based on Bootstrap samples. Moreover, we present the application of these new methods to heart rate measurements within the Augsburgger Umweltstudie: the corrected effect estimates were slightly higher than their naive equivalents. The substantial measurement error of ultrafine particle measurements has little impact on the results. The developed methodology is generally applicable to longitudinal data with measurement error.

**Key words:** Longitudinal data analysis; Measurement error; Mixed model; Particulate matter;

Supporting Information for this article is available on the WWW under <http://dx.doi.org/10.1022/bimj.XXXXXXX>

### 1 Introduction

Measurement errors occur in varying applications, including the context of ambient particulate matter measurements (Zeger *et al.*, 2000; Dominici *et al.*, 2003; Sheppard *et al.*, 2011). Epidemiological longitudinal studies of acute health effects of outdoor air pollution often use estimated exposure of the study participants by one or several fixed outdoor monitoring sites located in an urban background of the study area (e.g. Bergen *et al.*, 2016; Alexeeff *et al.*, 2016). The ambient air pollutant concentrations measured at the fixed sites are assumed to represent the population-averaged exposure and this value usually enters the analysis. However, this “measured value” is different from individual exposure of study participants

\*Corresponding author: e-mail: Veronika.Deffner@stat.uni-muenchen.de, Phone: +49-89-2180-3197, Fax: +49-89-2180-5308

because people spend their time in different microenvironments (from high exposure microenvironments, e.g. during commuting, to low exposure microenvironments, e.g. being at home) and have different time-activity patterns, as seen in Deffner *et al.* (2016). This type of measurement error is called Berkson error. The individual exposure could be more accurately estimated by mobile measurement devices carried by each individual; these devices deliver highly resolved measurements of individual exposure. Classical measurement error occurs if the device readings vary around the true value, i.e. around the true individual exposure, due to low precision of the instruments. In addition, mobile devices are difficult to handle, costly and often record only a short period (mostly up to 24 hours) without reloading the batteries.

Statistical analyses rely on the accuracy of the applied measures for data collection. Since measurement error possibly yields distorted results, its size, relevance and impact are worth to examine even in complex but practically relevant data situations. In contrast to methods that simultaneously account for measurement error (e.g. Bayesian and maximum likelihood (ML) methods), the exact calculation of the effect of measurement error enables the easy evaluation of its relevance, transparently elucidates its impact and involves the calculation of correction formulas. General linear mixed models (e.g. Diggle *et al.*, 2002; Pinheiro and Bates, 2000) offer a widely used framework for the analysis of multiple measurements in varying statistical units with temporal autocorrelation. Therefore, this method is applied to model the relationship between ECG measurements and ultrafine particle concentrations in the Augsburg Umweltstudie. Former works extended the existing knowledge about covariate measurement errors and the corresponding methods (see e.g. Carroll *et al.*, 2006, for an overview) to models with random effects (e.g. Tosteson *et al.*, 1998; Wang *et al.*, 1998), correlated measurement error terms (Wang *et al.*, 1996; Wang, 2000) and correlated predictors with correlated measurement error (Schwartz and Coull, 2003). So far, measurement error in linear mixed models with autocorrelated error terms has not been studied. Therefore, three challenges are considered within this work to investigate measurement error in data on individual exposure to particulate matter: 1. random effects within the measurement error structure, 2. autocorrelated measurement error and 3. a mixture of classical and Berkson error.

The paper is organized as follows: The Augsburg Umweltstudie and the related validation studies are described in Section 2. Covariate measurement errors in linear mixed models are theoretically examined and the resulting biases are quantified in Section 3. The properties for a small number of observations of each individual are considered as well as allowing for additional covariates. Moreover, the distributional characteristics of the estimator for the attenuation factor based on repeated measurements are examined. The biases are further investigated by a simulation study based on the Augsburg Umweltstudie (Section 4). Finally, adequate correction factors are applied to adjust the effect estimations of the Augsburg Umweltstudie (Section 5). Derivations and other details can be found in the Supporting Information.

## 2 Data description

The Augsburg Umweltstudie (Hampel *et al.*, 2012a,b; Peters *et al.*, 2015; Ruckerl *et al.*, 2014; Guo *et al.*, 2015) was conducted between March 2007 and December 2008 in the city of Augsburg and two adjacent counties. The aim of the study was to examine the association between human health and the concentration of fine and ultrafine particles in the air, which represent two quantities to measure air pollution. Since short-term effects of particle concentrations were the focal point, their association with blood parameters was investigated. One hundred and twelve individuals were enrolled and measured up to four times for a 5–6 hour period their individual exposure to particle number concentration (PNC) in their daily life by carrying a portable particle counter. Beside the individual exposure measurements, air pollution measurements at fixed monitoring site were collected. During the same time, the individuals wore ECG devices recording their cardiac rhythm activity and documented their activities in a diary. The data were analyzed using a resolution of five minutes.

Based upon the Augsburg Umweltstudie, a validation study was conducted to quantify the sizes of the measurement errors. In the first part of the validation study, which was described in Deffner *et al.* (2016),

ten volunteers measured their individual exposure to PNC using mobile devices during three predefined scenarios in three different seasons (winter, spring summer) in 2011. The scenarios comprised commuting between the home or the city center by public transport, car or as a pedestrian and the participants kept an activity diary during their measurements. In the second part of the validation study, another person conducted mobile PNC measurements on a predefined route within the same study period. Information about the relationship of individual and fixed site measurements can be deduced from this validation study. Moreover, comparison measurements between the three mobile devices and the measurement station were accomplished before and after each measurement campaign. These data reveal information about the measurement accuracy of the devices, which were gained through statistical models presented in Section 5.

The mobile exposure measurements of the Augsburgger Umweltstudie revealed two major problems: First, the precision of the devices is quoted to be  $\pm 20\%$  by the manufacture resulting in measurements with classical measurement error. Second, about 23% of the outdoor measurements were missing due to breakdown or incorrect appliance of the device by the study participants. The problem of missing values becomes even more severe if lagged exposure effects are considered. One solution is to substitute missing personal exposure measurements through the Berkson error-prone data from the fixed monitoring site. This approach induces a mixture between classical and Berkson error, i.e. classical measurement error occurs for about 77% of the observations and Berkson error occurs for the remaining part of the observations. Classical and Berkson errors are both considered to comprise individual-specific effects and to be autocorrelated. Autocorrelated and individual-specific Berkson error may arise, because the difference of the exposure levels in the microenvironments of the individuals and of the fixed site differ with temporal and individual-specific dependencies. Individual-specific handling of the portable measurement devices as well as differing accuracy depending on the device and the environmental conditions and which is varying over time may yield individual-specific and possibly autocorrelated classical measurement error.

The biases of the effect coefficients due to classical, Berkson and mixture error-prone covariates in a general linear regression model are theoretically examined in the following section.

### 3 Bias calculations

#### 3.1 Main model and measurement error models

A simple linear mixed model is considered as the main outcome model for the association between the health–outcome  $Y_{it}$  and the true exposure  $X_{it}$  of independent individuals  $i, i = 1, \dots, n$ , at time point  $t, t = 1, \dots, T$ :

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \tau_i + \varepsilon_{it}, \quad \tau_i \sim N(0, \sigma_\tau^2), \quad \varepsilon_i \sim N(\mathbf{0}, \Sigma_\varepsilon). \quad (1)$$

$\beta_0$  is the intercept of the model and  $\beta_1$  the effect coefficient for the impact of exposure on health. The individual-specific intercept is denoted by  $\tau_i$ , and the independent random model errors by  $\varepsilon_i$ , which are assumed to follow an AR(1) process with autocorrelation coefficient  $\rho$ . In the following, the correlation matrix of an AR(1) process with autocorrelation coefficient  $\rho$  is denoted by  $\mathbf{W}_\rho$ ; thus,  $\Sigma_\varepsilon = \sigma_\varepsilon^2 \mathbf{W}_\rho$  and  $\sigma_\varepsilon^2$  is the variance of  $\varepsilon_{it}$ .

The true latent individual exposure vector is denoted by  $\mathbf{X}_i$  and the measurements of the mobile device by  $\mathbf{X}_i^{*C}$ , which are assumed to include an additive classical measurement error,  $U_i^C$ . Measurements from a monitoring station at a fixed site  $\mathbf{X}_i^{*B}$  characterize the population exposure to PNC of all included individuals. If  $\mathbf{X}_i^{*B}$  is used instead of  $\mathbf{X}_i$  as covariate in the main outcome model, Berkson error  $U_i^B$  occurs, because these measurements do not account for the individual structure of personal exposure and represent a spatially averaged value. The classical measurement error of the fixed site measurements is neglected due to the lack of data to determine the size of the error, but is assumed to be small.

Thus,  $\mathbf{X}_i^{*B}$  and  $\mathbf{X}_i^{*C}$  represent the measurements with Berkson and classical measurement error, respectively. We extend the conventional error structure for both measurement error types through individual-specific components  $\nu_i^B$  and  $\nu_i^C$  and autocorrelated measurement errors  $U_i^B$  and  $U_i^C$ :

$$\begin{aligned} X_{it} &= X_{it}^{*B} + \nu_i^B U_{it}^B, & \nu_i^B &\sim N(0, \sigma_{\nu^B}^2), & U_i^B &\sim N(\mathbf{0}, \Sigma_{U^B}), & \Sigma_{U^B} &= \sigma_{U^B}^2 \mathbf{W}_{\rho^B}, & (2) \\ X_{it}^{*C} &= X_{it} + \nu_i^C U_{it}^C, & \nu_i^C &\sim N(0, \sigma_{\nu^C}^2), & U_i^C &\sim N(\mathbf{0}, \Sigma_{U^C}), & \Sigma_{U^C} &= \sigma_{U^C}^2 \mathbf{W}_{\rho^C}, & (3) \end{aligned}$$

where  $U_{it}^B$  and  $\nu_i^B$  are independent from  $X_{it}^{*B}$  and  $U_{it}^C$  and  $\nu_i^C$  are independent from  $X_{it}$ ; the measurement errors are non-differential and independent from each other and from the model errors  $\tau_i$  and  $\varepsilon_{it}$ .

Because mobile exposure measurements are missing for some observations, the information of the fixed site measurements could be used in these cases. This results in the mixture of Berkson and classical measurement error and is considered as a third type of measurement error. Formally, for a known proportion  $1 - p$  of observations, individual, but classical error-prone, PNC measurements are available, whereas the remaining fraction  $p$  of observations without individual measurements is substituted by Berkson error-prone measurements, e.g. fixed site records:

$$\begin{aligned} X_{it}^{*M} &= X_{it}^{*B} G_{it} + X_{it}^{*C} (1 - G_{it}) \\ &= \begin{cases} X_{it}^{*B} & \text{for } p \cdot 100\% \text{ of the measurements} \\ X_{it}^{*C} & \text{for } (1 - p) \cdot 100\% \text{ of the measurements} \end{cases}, \end{aligned} \quad (4)$$

with  $G_{it}$  indicating whether the observation at time point  $t$  of individual  $i$  exhibits Berkson error. This type of the mixture measurement error differs from the mixture error considered in Mallick *et al.* (2002), Carroll *et al.* (2007) or Yin *et al.* (2013), whose observed error-prone covariate consists of a latent variable with Berkson error superposed with classical measurement error.

Since distributional assumptions for  $\mathbf{X}_i^{*B}$  are necessary in some situations, we assume that

$$\mathbf{X}_i^{*B} \sim N(\mu^{X^{*B}} \mathbf{1}_T, \Sigma_{X^{*B}}) \quad \text{with} \quad \mathbb{E}(X_{it}^{*B}) = \mu^{X^{*B}} \quad \text{and} \quad \Sigma_{X^{*B}} = \sigma_{X^{*B}}^2 \mathbf{W}_{\rho^{X^{*B}}}.$$

Under these conditions,

$$\begin{aligned} \mathbf{X}_i &\sim N((\mu^{X^{*B}} + \nu_i^B) \mathbf{1}_T, \Sigma_X) & \text{with} \quad \Sigma_X &= \Sigma_{X^{*B}} + \Sigma_{U^B} \\ \mathbf{X}_i^{*C} &\sim N((\mu^{X^{*B}} + \nu_i^B + \nu_i^C) \mathbf{1}_T, \Sigma_{X^{*C}}) & \text{with} \quad \Sigma_{X^{*C}} &= \Sigma_{X^{*B}} + \Sigma_{U^B} + \Sigma_{U^C}. \end{aligned}$$

Note,  $\Sigma_X$  and  $\Sigma_{X^{*C}}$  are not the variance-covariance matrices of AR(1) processes but of ARMA processes (Lütkepohl, 1984) and  $\mathbf{X}_i^{*C}$  depends on  $\mathbf{X}_i^{*B}$ .

### 3.2 Bias and uncertainty of naive estimators

The bias of the naive estimators due to the usage of  $\mathbf{X}^* \in \{\mathbf{X}^{*B}, \mathbf{X}^{*C}, \mathbf{X}^{*M}\}$  can be calculated if the parameters describing the measurement errors are known. At the beginning of this section a general expression of the bias in a simple linear mixed model is given followed by expressions for Berkson and classical error as well as for the mixture of both error types assuming  $T \rightarrow \infty$ .

First, we derive the general form of the bias through a closer look on the estimation approach for the naive estimators.  $\beta^* = (\beta_0^*, \beta_1^*)^\top$  and  $\phi^* = (\sigma_{\varepsilon^*}^2, \rho^*, \sigma_{\tau^*}^2)^\top = (\phi_1^*, \phi_2^*, \phi_3^*)^\top$  denote the probability limits of the naive estimators for  $n \rightarrow \infty$  and are obtained as solutions from the linear mixed model score equations, if  $n \rightarrow \infty$ :

$$\mathbb{E} \left\{ \chi^{*\top} \mathbf{V}^{*-1} (\mathbf{Y} - \chi^* \beta^*) \right\} = 0, \quad (5)$$

and for  $\phi_j^*$ ,  $j = 1, 2, 3$ ,

$$\frac{1}{2} \left[ \mathbb{E} \left\{ (\mathbf{Y} - \boldsymbol{\chi}^* \boldsymbol{\beta}^*)^\top \mathbf{V}^{*-1} \frac{\partial \mathbf{V}^*}{\partial \phi_j^*} \mathbf{V}^{*-1} (\mathbf{Y} - \boldsymbol{\chi}^* \boldsymbol{\beta}^*) \right\} - \text{tr} \left( \mathbf{V}^{*-1} \frac{\partial \mathbf{V}^*}{\partial \phi_j^*} \right) \right] = 0, \quad (6)$$

with  $\boldsymbol{\chi}^* = (\mathbf{1}_T, \mathbf{X}^*)$  (see Wang *et al.*, 1998).  $\mathbf{V}^*$ ,  $\mathbf{V}^* = \boldsymbol{\Sigma}_{\varepsilon^*} + \sigma_{\tau^*}^2 \mathbf{J}_T$ , denotes the probability limit for  $n \rightarrow \infty$  of the variance–covariance matrix of the error term in the naive model and  $\mathbf{J}_T$  the  $T \times T$  matrix of ones.  $\tau_i^* \sim N(0, \sigma_{\tau^*}^2)$  is assumed and  $\varepsilon_{it}^*$  is assumed to follow an AR(1) process with autocorrelation parameter  $\rho^*$  and variance matrix  $\boldsymbol{\Sigma}_{\varepsilon^*} = \sigma_{\varepsilon^*}^2 \mathbf{W}_{\rho^*}$ ; index  $i$  is neglected in eq. (5) and eq. (6) and in the following considerations. To relate  $\beta_1^*$  to  $\beta_1$ , the first score equation (5) is rewritten as

$$\lambda = \frac{\beta_1^*}{\beta_1} = \frac{\text{tr} \{ \mathbf{V}^{*-1} \text{Cov}(\mathbf{X}, \mathbf{X}^*) \}}{\text{tr} \{ \mathbf{V}^{*-1} \text{Var}(\mathbf{X}^*) \}} \quad (7)$$

(Supporting Information, Appendix A.1.1). The attenuation factor  $\lambda$  resembles the common form of the ratio between  $\text{Cov}(\mathbf{X}, \mathbf{X}^*)$  and  $\text{Var}(\mathbf{X}^*)$ . The detailed expression of  $\lambda$  depends on the type of measurement error and is examined in the subsequent paragraphs.

The remaining model parameters  $\phi^*$  characterize the model errors and thus the uncertainty of the naive estimates. The score equations (eq. (5) and eq. (6)) do not allow to write  $\phi^*$  in terms of other unknown model parameters. Instead, only expressions for the variance–covariance structure of the model error  $\text{Var}(\mathbf{Y}|\mathbf{X}^*)$  can be derived (Supporting Information, Appendix A.2).

After the general considerations about the bias and the uncertainty of the naive estimators, we continue with a more detailed view on the impact of Berkson error. The effect estimate for a covariate with individual–specific and autocorrelated Berkson error is just as unbiased as with conventional Berkson error due to  $\mathbb{E}(\mathbf{Y}|\mathbf{X}^{*B}) = \beta_0 + \beta_1 \mathbf{X}^{*B}$ :  $\beta_1^{*B} = \beta_1$  and  $\lambda^B = 1$ .

The fraction of the variability of  $\mathbf{X}$  exceeding the variability of the Berkson error–prone measurement  $\mathbf{X}^{*B}$ , is conferred on the model error variance,  $\text{Var}(\mathbf{Y}|\mathbf{X}^{*B})$ . This implicates in conjunction with lower variability of  $\mathbf{X}^{*B}$  compared to  $\mathbf{X}$  an increased variance of the naive effect estimate,  $\widehat{\beta}_1^*$  (Supporting Information, Appendix A.2).

Classical measurement error mostly causes an attenuation of the effect estimate in linear regression analyses, i.e. the effect estimate is biased towards 0, and only a slight increase in the uncertainty of the effect (Fuller, 1987; Carroll *et al.*, 2006). If measurement error and model error are independent, the attenuation factor will be calculated through the ratio between  $\text{Cov}(\mathbf{X}, \mathbf{X}^{*C})$  and  $\text{Var}(\mathbf{X}^{*C})$  in linear regression models with the usual classical measurement error. For general linear mixed models, the attenuation factor results with some matrix algebra and with the Sherman–Morrison formula for the calculation of the inverse of  $\mathbf{V}^{*C}$  in eq. (7) as:

$$\begin{aligned} \lambda^C &= \frac{\beta_1^{*C}}{\beta_1} = \frac{\text{tr} \{ \mathbf{V}^{*C-1} \text{Cov}(\mathbf{X}, \mathbf{X}^{*C}) \}}{\text{tr} \{ \mathbf{V}^{*C-1} \text{Var}(\mathbf{X}^{*C}) \}} \\ &= \left\{ \sigma_{X^{*B}}^2 g_T^{*C}(\rho^{X^{*B}}) + \sigma_{\nu^B}^2 g_T^{*C}(1) + \sigma_{U^B}^2 g_T^{*C}(\rho^B) \right\} \\ &\quad \left\{ \sigma_{X^{*B}}^2 g_T^{*C}(\rho^{X^{*B}}) + \sigma_{\nu^B}^2 g_T^{*C}(1) + \sigma_{\nu^C}^2 g_T^{*C}(1) + \sigma_{U^B}^2 g_T^{*C}(\rho^B) + \right. \\ &\quad \left. \sigma_{U^C}^2 g_T^{*C}(\rho^C) \right\}^{-1}, \end{aligned} \quad (8)$$

with  $g_T^C(\rho) = 1 - 2\rho\rho^C + \rho^{*C2}$  if  $|\rho| < 1$  and  $g_T^C(\rho) = 0$  if  $\rho = 1$  and  $\rho^{*C} \neq 1$  for  $T \rightarrow \infty$  (Supporting Information, Appendix A.1.2 and Appendix A.1.4). Thus, for  $T \rightarrow \infty$ , the effect of  $\sigma_{\nu^B}^2$  and  $\sigma_{\nu^C}^2$  on the attenuation factor vanishes. If at least one of  $\rho^{X^{*B}}$ ,  $\rho^B$  and  $\rho^C$  differs from zero, the autocorrelation

coefficient(s) influence(s) the degree of attenuation through weighting the respective variance with the weight  $\frac{\sigma_{\nu^B}^2}{\sigma_{\nu^B}^2 + \sigma_{\nu^C}^2}$ . Although  $\text{Var}(Y|X^{*C})$  exceeds the error variance in a model without measurement error (Supporting Information, Appendix A.2), the uncertainty of the naive estimate  $\widehat{\beta_1^{*C}}$  may either increase or decrease because the variability of  $X^{*C}$  is higher compared to  $X$ .

Finally, bias and uncertainty of the naive estimator resulting from mixture error are examined. A random variable with mixture measurement error is composed by a certain percentage  $p$  of observations with Berkson error and the percentage  $(1 - p)$  of observations with classical measurement error (see eq. (4)).

In the simple linear regression case with usual classical and Berkson errors ( $\sigma_\tau^2 = \sigma_{\nu^B}^2 = \sigma_{\nu^C}^2 = 0$ ,  $\rho = \rho^B = \rho^C = 0$ ), the reliability ratio  $\lambda^M$  is calculated through

$$\lambda^M = \frac{\beta_1^{*M}}{\beta_1} = \frac{\text{Cov}(X, X^{*M})}{\text{Var}(X^{*M})} = \frac{\sigma_{X^{*B}}^2 + (1-p)\sigma_{U^B}^2}{\sigma_{X^{*B}}^2 + (1-p)(\sigma_{U^B}^2 + \sigma_{U^C}^2)}.$$

Since  $\lambda^M \leq 1$ , mixture measurement error entails an attenuation of the true effect estimate, similar to classical measurement error.

For the general framework described in Section 3.1, the attenuation factor is given by (Supporting Information, Appendix A.1.3):

$$\begin{aligned} \lambda^M &= \frac{\beta_1^{*M}}{\beta_1} = \frac{\text{tr}\{V^{*M-1}\text{Cov}(X, X^{*M})\}}{\text{tr}\{V^{*M-1}\text{Var}(X^{*M})\}} \\ &= \left[ \sigma_{X^{*B}}^2 g_T^{*M}(\rho^{X^{*B}}) + (1-p)\sigma_{\nu^B}^2 g_T^{*M}(1) + (1-p)\sigma_{U^B}^2 g_T^{*M}(\rho^B) \right] \\ &\quad \left[ \sigma_{X^{*B}}^2 g_T^{*M}(\rho^{X^{*B}}) + (\sigma_{\nu^B}^2 + \sigma_{\nu^C}^2) \{(1-p)^2 g_T^{*M}(1) + p(1-p)g_T^{*M}(0)\} + \right. \\ &\quad \left. \sigma_{U^B}^2 \{(1-p)^2 g_T^{*M}(\rho^B) + p(1-p)g_T^{*M}(0)\} + \right. \\ &\quad \left. \sigma_{U^C}^2 \{(1-p)^2 g_T^{*M}(\rho^C) + p(1-p)g_T^{*M}(0)\} \right]^{-1}. \end{aligned} \quad (9)$$

Analogously to the classical measurement error,  $g_T^{*M}(\rho) = 1 - 2\rho\rho^{*M} + \rho^{*M2}$  if  $|\rho| < 1$  and  $g_T^{*C}(\rho) = 0$  if  $\rho = 1$  and  $\rho^{*M} \neq 1$  for  $T \rightarrow \infty$  (Supporting Information, Appendix A.1.4).

Higher Berkson error reduces the bias of the effect estimate and increased classical error enlarges the bias. The random effect of the naive main model covers only an averaged individual-specific measurement error of individual  $i$ . Therefore, the variance between the individual mean of the Berkson error-prone measurements and the individual mean of the measurements with classical measurement error affects the denominator of the attenuation factor. This term strongly intensifies attenuation as it is easily seen in the case of  $\rho = \rho^B = \rho^C = 0$  and  $T \rightarrow \infty$ :

$$\lambda^M = \frac{\beta_1^{*M}}{\beta_1} = \frac{\text{Cov}(X, X^{*M})}{\text{Var}(X^{*M})} = \frac{\sigma_{X^{*B}}^2 + (1-p)\sigma_{U^B}^2}{\sigma_{X^{*B}}^2 + p(1-p)(\sigma_{\nu^B}^2 + \sigma_{\nu^C}^2) + (1-p)(\sigma_{U^B}^2 + \sigma_{U^C}^2)}. \quad (10)$$

$\lambda^M = \lambda^C$ , if  $p = 0$ , and  $\lambda^M = 1$ , if  $p = 1$ .

In general, the components of the attenuation factors  $\lambda^C$  and  $\lambda^M$  consist of the product of a weighting factor  $g_T^*(\cdot)$  and of a variance. Similar to the classical measurement error,  $\text{Var}(\widehat{\beta_1^{*M}})$  may be increased or decreased in comparison to  $\text{Var}(\widehat{\beta_1})$ , but if the fraction of observations with Berkson error is large enough, uncertainty of  $\widehat{\beta_1^{*M}}$  is higher than for  $\widehat{\beta_1}$  (Supporting Information, Appendix A.2).

### 3.3 Small number of observations per individual

The expressions for the attenuation factors considered in the previous section are valid for  $T \rightarrow \infty$ . Since asymptotic properties in the derivations of the attenuation factor do not hold for a small number of observations per individual  $T$ , the attenuation factor for autocorrelated model errors and measurement errors is calculated with the following weighting formula

$$g_T^*(\rho) = 1 - 2 \frac{T-1}{T} \rho \rho^* + \frac{T-2}{T} \rho^{*2} - \left\{ (T-2)(1-\rho^*)^4 + 2(1-\rho^*)^2(1+\rho^{T-1}) + 4(1-\rho^*)^3 \sum_{t=1}^{T-2} \rho^t + 2(1-\rho^*)^4 \sum_{k=1}^{T-3} \sum_{t=1}^k \rho^t \right\} \left[ T \left\{ T - 2(T-1)\rho^* + (T-2)\rho^{*2} + \frac{\sigma_{\varepsilon^*}^2}{\sigma_{\tau^*}^2} \rho^{*2} \right\} \right]^{-1}$$

(Supporting Information, Appendix A.1.2 and Appendix A.1.3). The influence of the weighting factor is higher for strongly differing  $\rho$  and  $\rho^*$ .

The variances of the individual-specific components of the measurement errors,  $\sigma_{\nu^B}^2$  and  $\sigma_{\nu^C}^2$ , contribute to the attenuation factors  $\lambda^C$  and  $\lambda^M$  with additional terms in the numerator and the denominator of the attenuation factor (see eq. (8) and eq. (9)). If  $\rho = \rho^{X^*B} = \rho^B = \rho^C = 0$ , individual-specific components in the measurement error affect the attenuation factor, analogically to heterogeneity in the error-prone covariate, only for small values of  $T$  (and of  $\sigma_{\tau^*}^2/\sigma_{\varepsilon^*}^2$ ) as it was shown in Wang *et al.* (1998); the weights of these components simplify to  $(1 + (T-1)\sigma_{\tau^*}^2/\sigma_{\varepsilon^*}^2)^{-1}$ , as found in Wang *et al.* (1998).

### 3.4 Unbalanced observations and missing values

So far, our bias calculation assumes a balanced sample, i.e. the observations for each individual are equidistant and have the same number  $T$ . An extension of the correction approach to a sample with missing values or with non-equidistant observation times is considered in this section. In these situations,  $\mathbf{X}_i$  and  $\mathbf{Y}_i$ ,  $i = 1, \dots, n$ , are not identically distributed and the attenuation factor is given by

$$\lambda = \frac{\beta_1^*}{\beta_1} = \frac{\text{tr} \{ \mathbf{V}^{*-1} \text{Cov}(\mathbf{X}, \mathbf{X}^*) \}}{\text{tr} \{ \mathbf{V}^{*-1} \text{Var}(\mathbf{X}^*) \}} = \frac{\sum_{i=1}^n \text{tr} \{ \mathbf{V}_i^{*-1} \text{Cov}(\mathbf{X}_i, \mathbf{X}_i^*) \}}{\sum_{i=1}^n \text{tr} \{ \mathbf{V}_i^{*-1} \text{Var}(\mathbf{X}_i^*) \}},$$

with individual-specific  $\mathbf{V}_i^*$ . In the case of individually varying measurement durations  $T_i$ , the calculation of the attenuation factors  $\lambda^C$  and  $\lambda^M$  differs in comparison to the balanced case only through individual-specific weights  $g_{T_i}^*(\cdot)$ .

If missing values occur in  $\mathbf{X}^{*M}$ , which cannot be imputed through fixed site measurements, or in  $\mathbf{X}^{*C}$ , the model errors and the measurement errors follow a continuous AR(1) process. The attenuation factor is in this case calculated with individual-specific weights (Supporting Information, Appendix A.1.5):

$$\lambda = \frac{\beta_1^*}{\beta_1} = \frac{\text{tr} \{ \mathbf{V}^{*-1} \text{Cov}(\mathbf{X}, \mathbf{X}^*) \}}{\text{tr} \{ \mathbf{V}^{*-1} \text{Var}(\mathbf{X}^*) \}} = \frac{\sum_{i=1}^n \text{tr} \{ (\mathbf{D}_i \mathbf{P}_i^\top)^{-1} \mathbf{D}_i \text{Cov}(\mathbf{X}_i, \mathbf{X}_i^*) \mathbf{D}_i^\top \}}{\sum_{i=1}^n \text{tr} \{ (\mathbf{D}_i \mathbf{P}_i^\top)^{-1} \mathbf{D}_i \text{Var}(\mathbf{X}_i^*) \mathbf{D}_i^\top \}}. \quad (11)$$

$\mathbf{D}_i$  is defined as  $\mathbf{I}_T$  where the rows of observations with missing values are deleted.

### 3.5 Calculation of $\text{Cov}(\mathbf{X}^*)$ and $\text{Var}(\mathbf{X}^*)$

As seen in eq. (7),  $\text{Cov}(\mathbf{X}^*)$  and  $\text{Var}(\mathbf{X}^*)$  are essential for calculating the attenuation factor. Two alternative calculations of  $\text{Cov}(\mathbf{X}^{*C})$  and  $\text{Cov}(\mathbf{X}^{*M})$  are presented in the following, since the



calculation of  $\text{Cov}(\mathbf{X}^*)$  can be done either based on  $\mathbf{X}^* \in \{\mathbf{X}^{*C}, \mathbf{X}^{*M}\}$  or on  $\mathbf{X}^{*B}$  (Supporting Information, Appendix A.1.6):

- Classical measurement error:

$$\begin{aligned} 1. \quad \text{Cov}(\mathbf{X}^{*C}) &= \text{Var}(\mathbf{X}^{*C}) - \sigma_{\nu^c}^2 \mathbf{J}_T - \boldsymbol{\Sigma}_{U^c} \\ 2. \quad \text{Cov}(\mathbf{X}^{*B}) &= \boldsymbol{\Sigma}_{X^{*B}} + \sigma_{\nu^B}^2 \mathbf{J}_T + \boldsymbol{\Sigma}_{U^B} \end{aligned} \quad (12)$$

- Mixture measurement error:

$$\begin{aligned} 1. \quad \text{Cov}(\mathbf{X}^{*M}) &= \text{Var}(\mathbf{X}^{*M}) + p(1-p)\sigma_{\nu^B}^2(\mathbf{J}_T - \mathbf{I}_T) + \\ &\quad p(1-p)\sigma_{\nu^B}^2 \mathbf{J}_T - \sigma_{\nu^B}^2 \mathbf{I}_T - \\ &\quad (1-p)\sigma_{\nu^c}^2 \mathbf{J}_T - p(1-p)\sigma_{\nu^c}^2 \mathbf{I}_T - \\ &\quad (1-p)^2 \boldsymbol{\Sigma}_{U^c} - p(1-p)\sigma_{U^c}^2 \mathbf{I}_T \\ 2. \quad \text{Cov}(\mathbf{X}^{*M}) &= \boldsymbol{\Sigma}_{X^{*B}} + (1-p)\boldsymbol{\Sigma}_{U^B} + (1-p)\sigma_{\nu^B}^2 \mathbf{J}_T \end{aligned} \quad (13)$$

Another alternative for the calculation of  $\text{Cov}(\mathbf{X}^*)$  and  $\text{Var}(\mathbf{X}^*)$  arises from known breakdown times of the mobile devices, i.e. the entries of  $\mathbf{G}$  are known. Since the breakdown times are known in the Augsburg-Umweltstudie, they can be involved in the estimation of  $\lambda^M$  (Supporting Information, Appendix A.1.6):

$$\begin{aligned} \lambda^M &= \frac{\beta_1^{*M}}{\beta_1} = \frac{\text{tr}\{\mathbf{V}^{*M^{-1}} \text{Cov}(\mathbf{X}, \mathbf{X}^{*M})\}}{\text{tr}\{\mathbf{V}^{*M^{-1}} \text{Var}(\mathbf{X}^{*M})\}} \\ &= \left[ \sum_{i=1}^n \text{tr}\left\{ \mathbf{V}^{*M^{-1}} \left( \boldsymbol{\Sigma}_{X^{*B}} + \left( \mathbf{1}_T - \mathbf{G}_i \right)^\top \circ \left( \sigma_{\nu^B}^2 \mathbf{J}_T + \boldsymbol{\Sigma}_{U^B} \right) \right) \right\} \right] \\ &\quad \left[ \sum_{i=1}^n \text{tr}\left\{ \mathbf{V}^{*M^{-1}} \left( \boldsymbol{\Sigma}_{X^{*B}} + \left( \mathbf{1}_T - \mathbf{G}_i \right) \left( \mathbf{1}_T - \mathbf{G}_i \right)^\top \circ \right. \right. \right. \\ &\quad \left. \left. \left. \left( \sigma_{\nu^B}^2 \mathbf{J}_T + \sigma_{\nu^c}^2 \mathbf{J}_T + \boldsymbol{\Sigma}_{U^B} + \boldsymbol{\Sigma}_{U^c} \right) \right) \right\} \right]^{-1}. \end{aligned} \quad (14)$$

### 3.6 Bias correction and confidence intervals

Knowing the formula of the attenuation factor enables us to directly correct bias due to classical or mixture measurement error; bias correction is unnecessary in the case of Berkson error. Therefore, the components of the attenuation factor  $\lambda \in \{\lambda^C, \lambda^M\}$  are estimated through their empirical equivalents resulting e.g. from validation data or repeated measurements using empirical estimators. Subsequently, the estimated naive effect coefficient is corrected with the estimated attenuation factor  $\hat{\lambda}$ :

$$\hat{\beta}_1 = \frac{\hat{\beta}_1^*}{\hat{\lambda}}.$$

This approach is known as the method of moments (e.g. Carroll *et al.*, 2006).

The variance of the corrected effect coefficient is approximated using the delta method:

$$\text{Var}\left(\hat{\beta}_1\right) = \text{Var}\left(\frac{\hat{\beta}_1^*}{\hat{\lambda}}\right) \approx \frac{\mathbb{E}\left(\hat{\beta}_1^*\right)^2}{\mathbb{E}\left(\hat{\lambda}\right)^4} \text{Var}\left(\hat{\lambda}\right) + \frac{1}{\mathbb{E}\left(\hat{\lambda}\right)^2} \text{Var}\left(\hat{\beta}_1^*\right). \quad (15)$$

The variance of the corrected estimator is constantly higher than the naive variance estimation, because  $0 < \hat{\lambda} \leq 1$ . Since the variability of the estimated attenuation factor  $\hat{\lambda}$  influences the variability of the bias corrected effect coefficient  $\hat{\beta}_1$ ,  $\text{Var}(\hat{\lambda})$  has to be considered for the calculation of confidence intervals for  $\hat{\beta}_1$ . Whereas the calculation of  $\text{Var}(\hat{\lambda})$  is possible in simple settings (Supporting Information, Appendix A.3), bootstrap replicates of the validation data may be used to estimate  $\text{Var}(\hat{\lambda})$  in more complex settings.

Bootstrap techniques provide confidence intervals with the advantage of relaxing the assumption of approximate normality of the estimator. For the longitudinal setting in our study this was realized by drawing  $b, b = 1, \dots, B$ , samples with replacement from the independent individuals of the main study and of the independent units of the validation study resulting in  $B$  bootstrap replications of  $\hat{\lambda}$ . Bootstrap percentile intervals were derived through the required percentiles of the  $B$  bootstrap replications.

### 3.7 Further covariates

The previous considerations are restricted to a single covariate  $X$ . The inclusion of additional precisely measured covariates  $Z$  does not affect the bias of the effect coefficient induced by the deficient covariate  $X^*$ , if  $Z$  and  $X^*$  are independent. However, the attenuation of the effect may change and biased effect estimates  $\beta_Z^*$  of  $Z$  possibly occur if  $X^*$  and  $Z$  are dependent. Attenuation even could be inverted if the measurement errors of two covariates are correlated (Carroll *et al.*, 2006) or if several covariates are measured with error (Buzas *et al.*, 2005).

Analogously to the effect estimator for multiple error-prone covariates in linear regression models of Carroll *et al.* (2006), the score equation for  $\beta_X^*$  and  $\beta_Z^*$  of the linear mixed model with one further precisely measured covariate  $Z$ , i.e.  $Y_{it} = \beta_0 + \beta_X X_{it} + Z_{it}\beta_Z + \tau_i + \varepsilon_{it}$ , is

$$\begin{pmatrix} \beta_0 \\ \beta_X \\ \beta_Z \end{pmatrix} = \begin{pmatrix} \text{tr}(\mathbf{V}^{*-1}\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\mu_X\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\mu_Z\mathbf{J}_T) \\ \text{tr}(\mathbf{V}^{*-1}\mu_X\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\text{Cov}(\mathbf{X}, \mathbf{X}^*)) & \text{tr}(\mathbf{V}^{*-1}\text{Cov}(\mathbf{X}, \mathbf{Z})) \\ \text{tr}(\mathbf{V}^{*-1}\mu_Z\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\text{Cov}(\mathbf{X}, \mathbf{Z})) & \text{tr}(\mathbf{V}^{*-1}\text{Var}(\mathbf{Z})) \end{pmatrix}^{-1} \begin{pmatrix} \text{tr}(\mathbf{V}^{*-1}\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\mu_X\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\mu_Z\mathbf{J}_T) \\ \text{tr}(\mathbf{V}^{*-1}\mu_X\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\text{Var}(\mathbf{X}^*)) & \text{tr}(\mathbf{V}^{*-1}\text{Cov}(\mathbf{X}^*, \mathbf{Z})) \\ \text{tr}(\mathbf{V}^{*-1}\mu_Z\mathbf{J}_T) & \text{tr}(\mathbf{V}^{*-1}\text{Cov}(\mathbf{X}^*, \mathbf{Z})) & \text{tr}(\mathbf{V}^{*-1}\text{Var}(\mathbf{Z})) \end{pmatrix} \begin{pmatrix} \beta_0^* \\ \beta_X^* \\ \beta_Z^* \end{pmatrix}, \quad (16)$$

with  $\mu_X, \mu_{X^*}$  and  $\mu_Z$  denoting the expected values of the respective random variable. Eq. (16) describes the bias correction exemplary for one further covariate, but the principle also holds for more than one additional covariate. Two items cause the differences to the simple case: (I) The probability limit of the variance–covariance matrix of the naive model errors  $\mathbf{V}^*$  changes due to additional explanatory power of the additional covariates and (II)  $Z$  directly influences the estimation of  $\beta_X^*$  through  $\text{Cov}(\mathbf{X}^*, \mathbf{Z})$ ,  $\text{Cov}(\mathbf{Z})$  and  $\beta_Z$ .

Cross–covariances between the covariates contribute to eq. (16), which will be difficult to estimate in real situations: in the Augsburgger Umweltstudie, these parameters strongly varied between the individuals. Therefore, empirical, individual–specific equivalents of the expectations, variances and cross–covariances were used and the equation system (16) was numerically solved to estimate the effect coefficients  $\beta_0, \beta_X$  and  $\beta_Z$ .

$\text{Cov}(\mathbf{X}, \mathbf{X}^*)$  and  $\text{Cov}(\mathbf{X}, \mathbf{Z})$  cannot be estimated from the data.  $\text{Cov}(\mathbf{X}, \mathbf{X}^*)$  has to be calculated with additional information (see Section 3.5), e.g. through a validation study.  $\text{Cov}(\mathbf{X}, \mathbf{Z})$  equals  $\text{Cov}(\mathbf{X}^*, \mathbf{Z})$  because the classical measurement errors and the covariate  $Z$  are assumed to be independent.

## 4 Simulation study

### 4.1 Simulation design

A simulation study was performed to visualize the theoretical results and to examine the properties of the estimators. The data was generated according to the main model (eq. (1)) and the measurement error models (eq. (2)-(4)) defined in Section 3.1. The parameters are specified as described in Table 1, inspired by the Augsburgger Umweltstudie. The regression coefficients  $\beta_0$  and  $\beta_1$  were chosen to be 73 and 1, respectively.

**Table 1** Parameter choices for simulations.

Variable	Mean	Variance	Autocorrelation
$X_{it}^{X^*B}$	$\mu^{X^*B} = 9.5$	$\sigma_{X^*B}^2 = 0.3$	$\rho^{X^*B} = 0.99$
$\nu_i^B$		$\sigma_{\nu^B}^2 = 0.04$	
$U_{it}^B$		$\sigma_{U^B}^2 = 0.2$	$\rho^B = 0.7$
$\nu_i^C$		$\sigma_{\nu^C}^2 = 0.03$	
$U_{it}^C$		$\sigma_{U^C}^2 = 0.06$	$\rho^C = 0.2$
$\tau_i$		$\sigma_{\tau}^2 = 140$	
$\varepsilon_{it}$		$\sigma_{\varepsilon}^2 = 70$	$\rho = 0.7$

Three scenarios were considered for the structure of the measurement errors: (I) uncorrelated errors without individual-specific structure ( $\nu^B = \mathbf{0}, \nu^C = \mathbf{0}, \rho^B = 0, \rho^C = 0, \rho = 0$ ), (II) uncorrelated random errors and individual-specific measurement error components ( $\rho^B = 0, \rho^C = 0, \rho = 0$ ), (III) autocorrelated errors and individual-specific measurement error components.

These scenarios were evaluated for data with the three examined types of measurement error (Berkson, classical, mixture) as well as for the proper data without measurement error using linear models with a random intercept and an AR(1) error term. The size of measurement error was varied by scaling the variances of the random errors ( $\sigma_{U^B}^2, \sigma_{U^C}^2$ ) with the scaling factors (0.5, 1, 2 and 5; the other parameters remained equal within each scenario. The percentage of missing values in  $X^{*C}$  and the percentage of Berkson error-prone measurements in  $X^{*M}$  was, accordingly to the percentage of missing values in the Augsburgger Umweltstudie, chosen to be 23 %. Within each of the  $N = 100$  iterations, measurement series of length  $T = 75$  were generated for  $n = 100$  individuals.

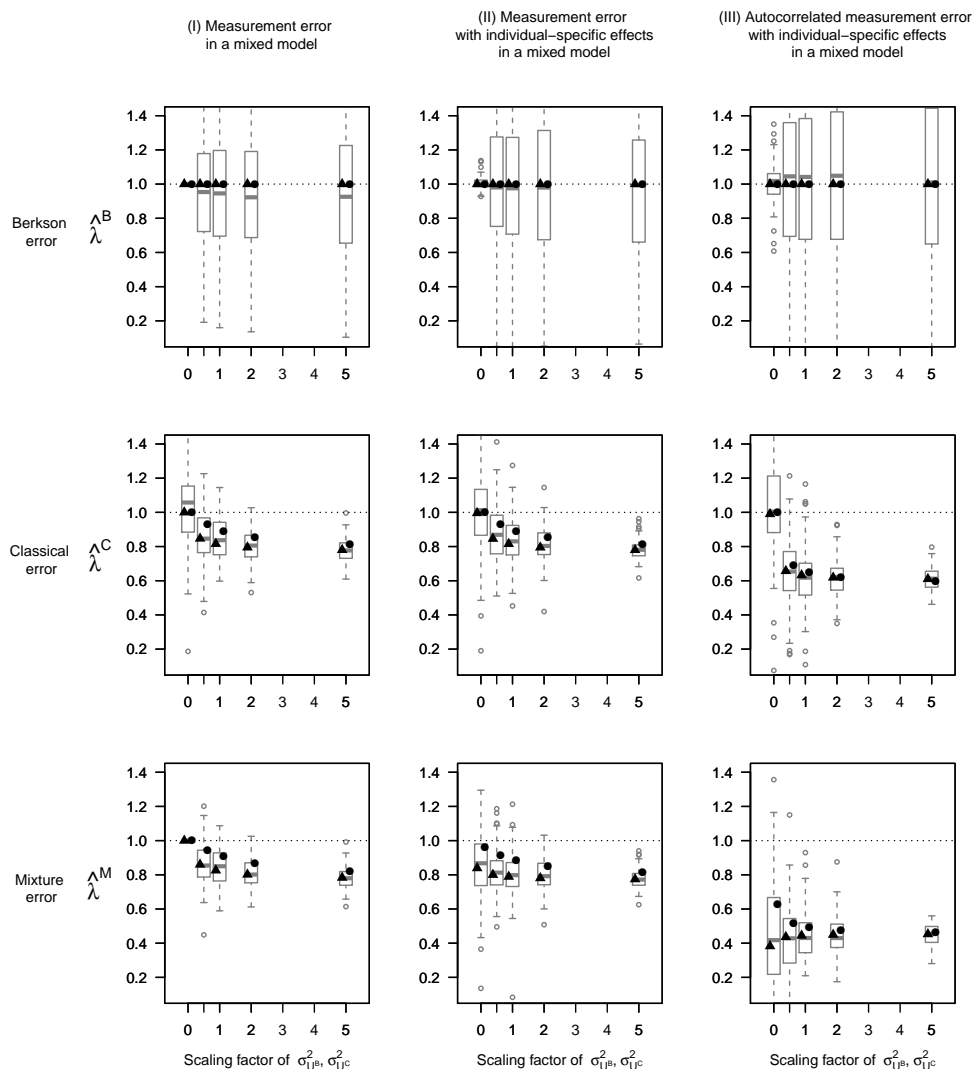
### 4.2 Attenuation in the general linear mixed model

The simulation results for the empirical attenuation factors depending on the varying measurement error sizes are presented in Figure 1. The empirical attenuation factor was calculated as the ratio between the estimated effect using the error-prone data  $\mathbf{X}^*, \widehat{\beta}_1^*$ , and the estimated effect using the proper data  $\mathbf{X}, \widehat{\beta}_1$ .

On average, measurements with Berkson error provided unbiased effect estimates, also in the presence of individual-specific measurement error or autocorrelated measurement error. AR(1) Berkson errors (scenario (III)) intensified the known accuracy reduction of the effect estimates.

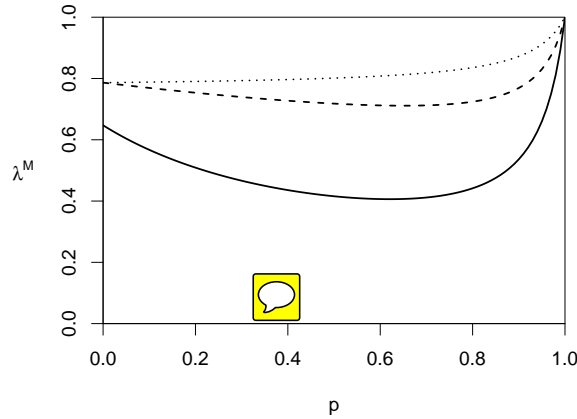
In contrast to individual-specific classical measurement error, which did not affect attenuation for  $T \rightarrow \infty$ , autocorrelated errors strengthened the degree of attenuation in particular regarding small measurement error variances. Note, that the theoretical estimations of the attenuation factor allow for the presence of missing values in the measurements with classical measurement error (see Section 3.4).

Mixture error attenuated the effect estimates as it is known for classical error-prone data, but to a different extent. Individual-specific components in the classical and Berkson error induced bias in the effect estimates even without random measurement errors ( $U^B, U^C$ ). The bias resulting from mixtures of autocorrelated measurement errors was high and strongly depended on both the size of the measurement errors and



**Figure 1** Simulation results for the empirical attenuation factor depicted in boxplots for increasing size of classical and Berkson measurement error; the theoretical estimations of  $\lambda$  with the exact value of  $T$  (using eq. (11) for  $\lambda^C$  and the combination of eq. (11) and eq. (14) for  $\lambda^M$ ) are marked with triangles and with circles for calculations assuming  $T \rightarrow \infty$  according to eq. (8) and eq. (9). Please note,  $\sigma_{\nu_B}^2 = 0.04$  and  $\sigma_{\nu_C}^2 = 0.03$ .

the size of the autocorrelation coefficients. Increasing mixture measurement error reduced the attenuation of the effect in scenario (III); this is possible to occur because the numerator of the attenuation factor is not completely contained in its denominator. In general, attenuation resulting from mixture measurement



**Figure 2** The role of the proportion  $p$  of observations affected by Berkson error for the attenuation factor  $\lambda^M$  in the case of mixture measurement error; dotted line: scenario (I), dashed line: scenario (II), solid line: scenario (III).

error strongly depends on the proportion  $p$  of observations affected by Berkson error (Figure 2, eq. (9) and eq. (10)). Note,  $\lambda^M = \lambda^C$ , if  $p = 0$ , and  $\lambda^M = 1$ , if  $p = 1$ .

**Table 2** Bias and RMSE for the bias corrected effect estimations for scenarios (I)-(III); bold: lowest absolute bias and lowest RMSE of each row (disregarding the results based on the true data).

Scen.	SF of $\sigma_{UB}^2, \sigma_{UC}^2$	Truth		Berkson		Classical		Mixture	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
(I)	0	0.015	0.360	<b>0.015</b>	<b>0.360</b>	0.020	0.421	<b>0.015</b>	<b>0.360</b>
	0.5	0.024	0.242	<b>-0.014</b>	0.407	0.016	0.295	0.019	<b>0.275</b>
	1	0.002	0.187	<b>-0.008</b>	0.378	0.022	0.246	0.014	<b>0.234</b>
	2	0.003	0.137	0.003	0.370	<b>-0.002</b>	0.182	-0.003	<b>0.177</b>
	5	-0.012	0.095	0.013	0.390	<b>-0.009</b>	<b>0.118</b>	-0.011	0.119
(II)	0	0.062	0.387	0.062	<b>0.389</b>	0.074	0.409	<b>0.033</b>	0.413
	0.5	-0.039	0.237	-0.067	0.364	<b>-0.035</b>	0.282	-0.050	<b>0.280</b>
	1	0.015	0.179	0.015	0.397	-0.010	0.243	<b>0.000</b>	<b>0.238</b>
	2	0.031	0.151	<b>0.000</b>	0.370	0.027	0.185	0.029	<b>0.180</b>
	5	-0.006	0.092	-0.034	0.387	-0.011	0.117	<b>-0.008</b>	<b>0.117</b>
(III)	0	-0.015	0.658	-0.018	<b>0.670</b>	-0.048	0.759	<b>0.005</b>	0.953
	0.5	0.036	0.272	<b>-0.007</b>	0.606	0.053	<b>0.402</b>	0.088	0.533
	1	-0.012	0.198	-0.002	0.672	0.017	<b>0.299</b>	<b>-0.001</b>	0.348
	2	-0.009	0.165	0.013	0.677	<b>-0.002</b>	<b>0.204</b>	0.013	0.255
	5	0.009	0.097	0.095	0.644	<b>0.031</b>	<b>0.150</b>	0.038	0.188

Comparing the correction methods in terms of bias and RMSE (Table 2) shows that the corrected estimations for classical or mixture error-prone data exceeded the accuracy of the estimations for Berkson

error-prone data. In scenarios (I) and (II) the correction based on mixture data is superior or comparable to the correction based on incomplete data with classical error, whereas in scenario (III) the correction of classical error-prone data is more precise.

The RMSE values of corrected effect estimations for data with mixture error are not generally higher than for incomplete data with classical measurement error. For example, for lower  $\rho^{X^*B}$ , the RMSE is better or comparably good for the estimated coefficient based on data with mixture error (Supporting Information, Table 1). Hence, using incomplete data with classical error outperforms the usage of data with mixture error only under certain parameter specifications.

Weights  $g_T^*(\cdot)$  depending on  $T$ , as described in the Sections 3.2, 3.3 and 3.4, were used for the estimation of the attenuation factors to account for the finite number of observations for each individual and unbalanced or not equidistant data. Figure 1 and Supporting Information, Figure 1 indicate that the finite sample correction is indeed relevant, because the weights  $g_T^*(\cdot)$ , and thus also the attenuation factor, strongly depend on the number of observations per individuals  $T$ , especially in scenario (III).

### 4.3 Confidence intervals

Three types of 95 % confidence intervals (CIs) for the corrected effect estimations were compared regarding their empirical coverage probabilities with a further simulation: bootstrap percentile intervals using 100 bootstrap iterations (“bootstrap”) and CIs assuming normality of  $\widehat{\beta}_1$  with the variance of the naive estimate (“naive”),  $\widehat{\beta}_1^*$ , and with an approximation of the variance using the delta method (eq. (15), “delta method”). The comparison was based on 1000 CIs. Varying values of  $\lambda^C$  and  $\lambda^M$  were investigated, generated through varying sizes of classical measurement error  $\sigma_{V_C}^2$ ; the other components of  $\lambda^C$  and  $\lambda^M$  were fixed according to Table 1 and were assumed to be known. Variability of  $\widehat{\lambda}$  was generated through random shifting of  $\widehat{\lambda}$  with a variance of 0.0004, which was approximately the bootstrapped variance of  $\widehat{\lambda}$  in the Augsburg Umweltstudie.

Table 3 presents the simulation results. Indeed, the naive estimations of the confidence intervals exhibited the most narrow intervals, but their coverage probability was in nearly all scenarios below 95 %. Only for strong attenuation, bootstrap percentile intervals provided inadequate coverage probabilities. In comparison to confidence intervals based on the normality assumption of the estimator and on variance calculations using the delta method, bootstrap confidence intervals were slightly narrower. The coverage probabilities of the delta method CIs were also for large measurement errors high. Uncertainty in the estimation of  $\lambda$  was adequately incorporated in CIs based on the delta method, whereas the coverage of bootstrap CIs diminished for strong measurement error.

The simulations for the complex situation with autocorrelation showed: confidence intervals based on the delta method and also on the bootstrap method tended to be too conservative, especially for small measurement error.

### 4.4 Further covariates

For the Augsburg Umweltstudie we assumed that the individual PNC level was the only parameter measured with error. Besides exposure to PNC, variable selection within the main study identified some further covariates with a significant concurrent or lagged impact on the heart rate of the individuals (Hampel *et al.*, 2012b). The following additional covariates affected the effect estimate of PNC, because they were correlated (according to the correlation coefficient of Spearman) with the fixed site and thus also with the individual PNC measurements: temperature at lag 2 ( $r = -0.399$ ), quadratic temperature at lag 2 ( $r = -0.376$ ) and relative humidity at lag 1 ( $r = 0.17$ ).

We examined the influence of correlated covariates on the attenuation of the effect coefficient in the framework of classical and mixture measurement error using bootstrap samples (of the individuals) from the original data set. Realizations of  $X$ ,  $X^{*C}$  and  $X^{*M}$  were generated according to the settings in Section 4.1 and scenario (III) based on the log-transformed samples of the fixed site measurements as realizations

**Table 3** Comparison of naive CIs, bootstrapped CIs and CIs based on the delta method regarding their empirical coverage probability and interval width; bold: empirical coverage probabilities greater than 0.95.

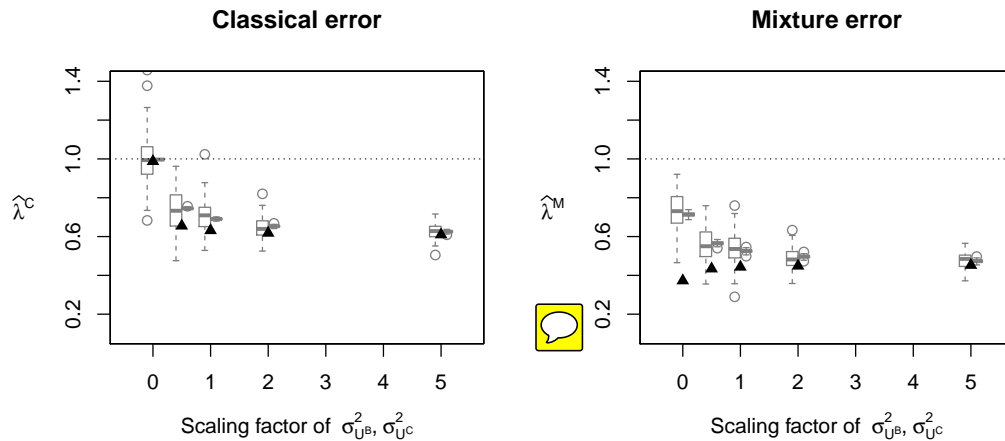
Error type	Var( $\hat{\lambda}$ )	$\lambda$	Empirical coverage probability			Interval width		
			Naive	Bootstrap	Delta method	Naive	Bootstrap	Delta method
Classical error	0	0.050	0.085	0.920	0.941	0.205	3.899	4.099
		0.200	0.329	<b>0.954</b>	<b>0.970</b>	0.410	1.949	2.049
		0.400	0.634	<b>0.955</b>	<b>0.979</b>	0.579	1.377	1.448
		0.600	0.895	<b>0.987</b>	<b>0.992</b>	0.709	1.126	1.182
	0.0004	0.800	<b>0.988</b>	<b>0.995</b>	<b>0.999</b>	0.818	0.975	1.023
		0.950	<b>1.000</b>	<b>0.998</b>	<b>1.000</b>	0.892	0.895	0.938
		0.200	0.344	0.933	<b>0.968</b>	0.410	1.998	2.119
		0.400	0.642	<b>0.950</b>	<b>0.980</b>	0.579	1.382	1.464
Mixture error	0	0.600	0.882	<b>0.986</b>	<b>0.994</b>	0.709	1.129	1.192
		0.800	<b>0.986</b>	<b>0.992</b>	<b>0.998</b>	0.818	0.977	1.030
		0.050	0.098	0.928	<b>0.953</b>	0.203	3.872	4.062
		0.200	0.343	<b>0.958</b>	<b>0.974</b>	0.406	1.929	2.031
	0.0004	0.400	0.683	<b>0.968</b>	<b>0.988</b>	0.574	1.361	1.436
		0.600	0.891	<b>0.988</b>	<b>0.994</b>	0.703	1.118	1.172
		0.200	0.343	<b>0.950</b>	<b>0.974</b>	0.406	1.968	2.091
		0.400	0.656	<b>0.969</b>	<b>0.987</b>	0.574	1.368	1.456
	0.600	0.890	<b>0.986</b>	<b>0.994</b>	0.703	1.120	1.181	

of  $X^{*B}$ . The main health outcome was generated with the simulated individual PNC measurements and the corresponding measurements of temperature, relative humidity, time of a day and the trend variable of the bootstrap sample using the naive parameter estimates of the precisely measured parameters in the main model of the Augsburg Umweltstudie. The attenuation effect is estimated with  $\hat{\beta}_X^*/\hat{\beta}_X$  after the calculation of  $\hat{\beta}_X$  using eq. (16).

The results in Figure 3 show that the degree of attenuation due to classical and mixture measurement error was reduced when further confounder variables were included in the model (Figure 3) in comparison to the simple regression model (triangles in Figure 3). The attenuation factor of the effect in multiple regression would have been underestimated if the correlation between the covariates was neglected and the correction formulas for the simple regression model were applied. Correcting the naive estimations according to eq. (16) yielded more precise estimations. Bias and RMSE of the estimators are listed in Supporting Information, Table 1.

## 5 Application to Augsburg Umweltstudie

The method of moments correction for measurements with Berkson, classical and mixture measurement error in a mixed model framework was applied to the models of the Augsburg Umweltstudie by using the data of a validation study described in Section 2. For the naive main model of the Augsburg Umweltstudie, the 5-min. resolved heart rate was the health outcome and the corresponding (log-transformed) mobile PNC measurements were regarded as the deficient covariate (Hampel *et al.*, 2012b). The confounder model established for the analysis of the Augsburg Umweltstudie was used, which included a linear time trend, linear and quadratic effects of 2-h lagged temperature at the measurement station, a linear effect of 1-h lagged relative humidity and a binary variable for time of the day (before/after noon).



**Figure 3** Simulation results of scenario (III) for the empirical attenuation factor of  $\hat{\lambda}^C$  for increasing size of classical and Berkson measurement error in the presence of further covariates; left boxplots: empirical attenuation factor; right boxplots: estimated attenuation factors according to eq. (16) and the first versions of eq. (12) and Supporting Information, Appendix A.1.6 (with known break down times  $\mathbf{G}$ ); the triangles represent the estimated attenuation factors according to eq. (11) and eq. (14) assuming independence between the deficient and precisely measured covariates.

The analysis was solely based on outdoor measurements, because indoor measurements differ in the size of Berkson error in comparison to outdoor measurements. The log-transformed fixed site measurements were adapted to the mean of the log-transformed individual measurements by adding a constant to fulfill the assumptions for the bias correction defined in Section 3.1. Varying degrees of lagged effects of the log-transformed PNC concentrations were examined up to a lag of 60 minutes.

The parameters for the correction of the measurement errors were estimated with validation data as well as with data from the Augsburgger Umweltstudie and are listed in Table 4. Variance and autocorrelation coefficient of  $X^{*B}$  were directly estimated from the stationary PNC measurements of the Augsburgger Umweltstudie through the sample variance and through a general linear intercept model for fixed site measurements assuming a continuous AR(1) process for the error term, respectively. The remaining parameters were calculated in three steps:

First, comparison measurements of the mobile devices and the devices at the fixed site were conducted before and after each measurement campaign of the validation study characterizing the classical measurement error. Several battery changes were necessary during each period of comparison measurements. This data was used to estimate  $\sigma_{\nu^C}^2$ ,  $\rho^C$  and  $\alpha_j^{\text{device}}$  with a regression model for the differences between (log-transformed) values of the mobile and stationary measurement devices:

$$\log(PNC_{jkt}^{\text{mobile}}) - \log(PNC_{kt}^{\text{station}}) = \alpha_0 + \alpha_j^{\text{device}} + \alpha_{jk}^{\text{bat. change}} + \varepsilon_{jkt}^C.$$

$j$  is the index for the three mobile devices, which measure PNC levels and  $k$  for the periods between battery changes. The classical measurement error in this experimental setup consists of three components: 1) a random error  $\varepsilon_{jkt}^C$ , 2) an error defined by periods between the battery changes  $\alpha_{jk}^{\text{bat. change}}$  and 3) a device error  $\alpha_j^{\text{device}}$ . This error structure was modeled with a linear model including categorical effects for device and battery change period. The classical measurement error  $\sigma_{\nu^C}^2$  was approximated through  $\widehat{\text{Var}}(\hat{\varepsilon}_{jkt}^C)$  and the individual specific classical measurement error  $\sigma_{\nu^C}^2$  through  $\widehat{\text{Var}}[\log(PNC_{jkt}^{\text{mobile}}) - \log(PNC_{kt}^{\text{station}})] -$



$\widehat{\text{Var}}(\widehat{\varepsilon}_{jkt}^C)$ .  $\widehat{\rho}^C$  was estimated by a general linear intercept model (using restricted maximum likelihood (REML) estimation) for  $\widehat{\varepsilon}_{jk}^C$  assuming a continuous AR(1) process for the error term.

Second, a regression model with the differences between mobile and stationary PNC measurements of the Augsburgger Umweltstudie while the individual was staying outdoors as response and the individual as covariate provided information about  $\sigma_{UC}^2 + \sigma_{UB}^2$ ,  $\sigma_{\nu^C}^2 + \sigma_{\nu^B}^2$  and  $\rho^B$  for individual  $i$  and time point  $t$ :

$$\log(PNC_{it}^{\text{mobile}}) - \log(PNC_{it}^{\text{station}}) = \nu_0 + \nu_i^{\text{BC}} + \varepsilon_{it}^{\text{BC}}.$$

$\nu_i^{\text{BC}}$  denotes a categorical effect for individual  $i$ .  $\widehat{\rho}^B$  was estimated by a general linear intercept model (using REML estimation) for  $\widehat{\varepsilon}_i^{\text{BC}}$  assuming a continuous AR(1) process for the error term and is only an approximation, because actually, the sum of classical and Berkson error is considered.

Third, the size of the Berkson error was directly estimated from the difference between the size of the sum of measurement errors (second step) and the classical measurement error (first step).  $\sigma_{UB}^2$  was estimated through  $\widehat{\text{Var}}(\widehat{\varepsilon}_{it}^{\text{BC}}) - \widehat{\sigma}_{UC}^2$ . The size of individual specific Berkson error  $\sigma_{\nu^B}^2$  was estimated through  $\widehat{\text{Var}}[\log(PNC_{it}^{\text{mobile}}) - \log(PNC_{it}^{\text{station}})] - \widehat{\text{Var}}(\widehat{\varepsilon}_{it}^{\text{BC}}) - \widehat{\sigma}_{\nu^C}^2$ .

The reason for this approach is that the Berkson error prone measurements were likely to underly sample-specific, unadjusted variations; therefore, external validation data may not be appropriate.

**Table 4** Parameter specifications for application (concurrent personal and fixed site measurements).

Variable	Variance	Autocorrelation
$X^{\text{AB}}$	$\sigma_{X^{\text{AB}}}^2 = 0.34$	$\rho^{X^{\text{AB}}} = 0.932$
$\nu_i^B$	$\sigma_{\nu^B}^2 = 0.21$	
$U_{it}^B$	$\sigma_{U^B}^2 = 0.3$	$\rho^B = 0.582$
$\nu_i^C$	$\sigma_{\nu^C}^2 = 0.03$	
$U_{it}^C$	$\sigma_{U^C}^2 = 0.03$	$\rho^C = 0.696$

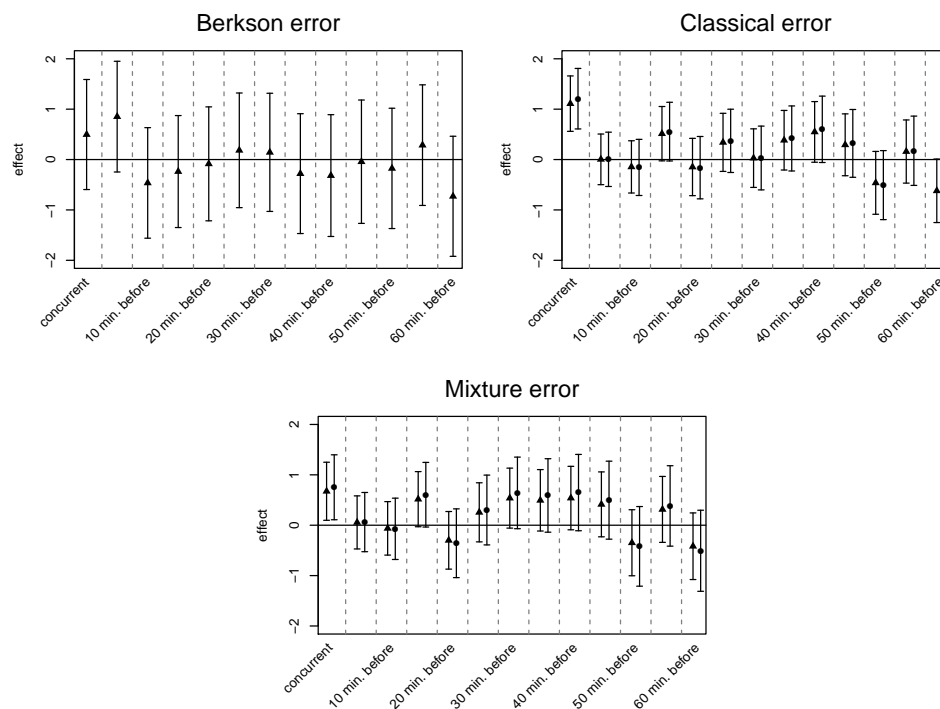
Three models were considered: Fixed site PNC levels were used as covariate in the first model. For the second model, only observations with individual measurements were used and for the third model, missing individual PNC observations were substituted with the values from the fixed measurement station. Equation system (16) was used for bias correction and  $\text{Cov}(X^{\text{AB}*})$  was estimated adjusting for the missing values in  $X^*$  and the known breakdown times in  $X^{\text{M}}$ .

Confidence intervals for the effect coefficients were based on the more conservative delta method (Section 4.3). The variance of the attenuation factor,  $\text{Var}(\widehat{\lambda})$ , was calculated using bootstrap samples of the appropriate data sets. The results are presented in Figure 4.

The application of the developed bias correction on the data of the Augsburgger Umweltstudie revealed a slight increase of the estimated effect coefficient considering the models for lag zero. We prefer the model based on mixture data because more information is used; this argument would be of special interest in situations with higher proportions of missing values. To further illustrate our theoretical results, also simple regression models were estimated. The regression coefficients and their bias corrected equivalents only slightly differ from the effect estimates presented in Figure 4 and are not depicted here.

## 6 Discussion

We examined Berkson error, classical measurement error and a special mixture error of a covariate in general linear mixed models with autocorrelated errors and permitted a wide range of error structures. The structure of the mixture error was motivated by the Augsburgger Umweltstudie as a mixture distribution



**Figure 4** Multiple regression analysis: Naive (triangles) and corrected (circles) effect estimates for the association between heart rate and individual PNC levels with Berkson, classical and mixture measurement error.

of Berkson and classical error. The complex model structure enables valuable insights into the role of exposure measurement error during the analysis of the association between exposure and human health.

We quantified the biases in the linear mixed models with AR(1) error term when error-prone measurements are used. Analogously to the linear model, Berkson error-prone measurements yield unbiased effect estimations but the error variance of the model and the variance of the random effects are overestimated involving wider confidence intervals.

The size of attenuation of the effect estimation resulting for a covariate with classical measurement error may change with autocorrelated errors in both directions. Each variance included in the attenuation factor is therefore modified with a multiplicative factor depending on the correlation coefficients. We also found that individual-specific components in the error structure can be neglected in the bias calculation for classical measurement error if the number of observations for each cluster is high enough.

Mixture measurement error has similar characteristics as classical measurement error. However, the main difference is that the individual-specific components of the error strengthen attenuation. Furthermore, autocorrelated or individual-specific Berkson error affects the estimation of the effect using data with mixture error.

Classical measurement error of the fixed site measurements was neglected in our considerations. This data situation would result in a mixture of Berkson and classical measurement error, which is defined as a mixture of errors (different to our definition) by Mallick *et al.* (2002) as follows:

$$X = X^{*B} + U^B$$

$$X^{*BC} = X^{*B} + U^{BC}.$$

Instead of the Berkson error-prone measurements  $X^{*B}$ , only  $X^{*BC}$  can be observed, which represents the latent fixed site exposure with an additive measurement error. The attenuation factor in this case can be obtained along the lines of the derivations in Supporting Information, Appendix A.1 and is given by

$$\frac{\sigma_{X^{*B}}^2}{\sigma_{X^{*B}}^2 + \sigma_{U^{BC}}^2}$$

in the situation without autocorrelation and individual-specific effects.


Moreover, the concept of the mixture error can be generalized to observations with heteroscedastic errors depending on a categorical covariate, like differing measurement errors for indoor and outdoor observations or for observations from different microenvironments, as considered in Muff and Keller (2015). Thus,  $\mathbf{G}$ , a vector indicating the breakdown of the devices, can be generally seen as an arbitrary categorical covariate.


The developed bias correction formulas are generally applicable to the considered types of covariate measurement errors in linear mixed models with an AR(1) error term. Extensions of the basic method to common practical data situations with unbalanced design, missing values, and additional covariates are provided. Hence, the theoretical, general considerations are transferable to other longitudinal studies. For this purpose, the method of moments is a simple, effective and practically established method for the correction of measurement error (Carroll *et al.*, 2006) requiring only little computational demand.

A further strength of the analysis is the usage of individual exposure measurements involving more reliable results for individual health outcomes than data from one or several fixed monitoring sites. However, only a single pollutant measure is collected and multipollutant analyses, like in Bergen *et al.* (2016) are not feasible.

The methods are restricted to linear mixed models, i.e. normally distributed model errors are assumed. Moreover, the measurement errors are supposed to be non-differential, which is fulfilled in the Augsburg Umweltstudie, and to exhibit a certain structure, which indeed extends the usual structure, but still may be too simplistic. Error corrections in hierarchical multi-level analyses are not considered within this work. The presented approaches are extensible to more complex correlation structures. In these cases, the variances in the attenuation factor are also weighted, but with different weights. Since these weights possibly cannot be simplified to general explicit formulas, the impact of the parameters describing the measurement error does not become obvious in contrast to the presented data situation. Error correction through the method of moments provides only an approximation, because the estimation of the correlation structure of the model error term cannot be adapted.

Alternative estimation approaches include the use of instrumental variables (e.g. Schennach, 2013) the ML method or also Bayesian methods as described and applied in e.g. Rosner *et al.* (1989), Spiegelman *et al.* (2000), Thoresen and Laake (2000) and Carroll *et al.* (2006). The ML approach results in a more or less pronounced increase in efficiency (Carroll *et al.*, 2006); the advantage of Bayesian methods is the possibility to combine various sources of information (Carroll *et al.*, 2006). However, these approaches do not allow a detailed understanding of the impact of measurement error in linear mixed models. Additionally, the ML and the Bayes approach are known to be sensitive to distributional assumptions; this is particularly relevant for the specification of an exposure model for personal exposure to PNC.

Applying the correction methods to the Augsburg Umweltstudie revealed slight differences in the effect coefficients compared to the naive estimations. The examined complex error structure had little impact on the investigated effect estimations in the Augsburg Umweltstudie. The major reason was, that the measurement error of the devices, which was quoted by  $\pm 20\%$  for 1-min. resolved data, was substantially lower for the analysis of data with a 5-min. resolution. Thus, the conclusions deduced from the original naive analyses persist. This is an interesting and important result because unexpectedly, the measurement error  only a slight impact on the effect estimates.

Deriving information  about the measurement error from repeated measurements in a linear mixed model setting is elaborate and hardly realizable, because many repetitions and time points are necessary. Information from validation studies or comparison measurements seem to be more promising as shown with the

practical example of the Augsburger Umweltstudie. The combination of the information from the main study and the validation studies was accomplished with a two-stage procedure: (I) modeling with deficient data and (II) error correction with validation data. Future Bayesian analyses will combine these steps in one model.

**Acknowledgements** The contribution of Prof. Dr. Annette Peters has been supported by the US Environmental Protection Agency STAR center grant RD 832415 (EPA Particulate Matter Centre).

Data collection for the validation study of the Augsburger Umweltstudie as well as the contributions of Susanne Breitner and Veronika Deffner has been funded by the German Research Foundation (DFG grant number KU1359/2-1).

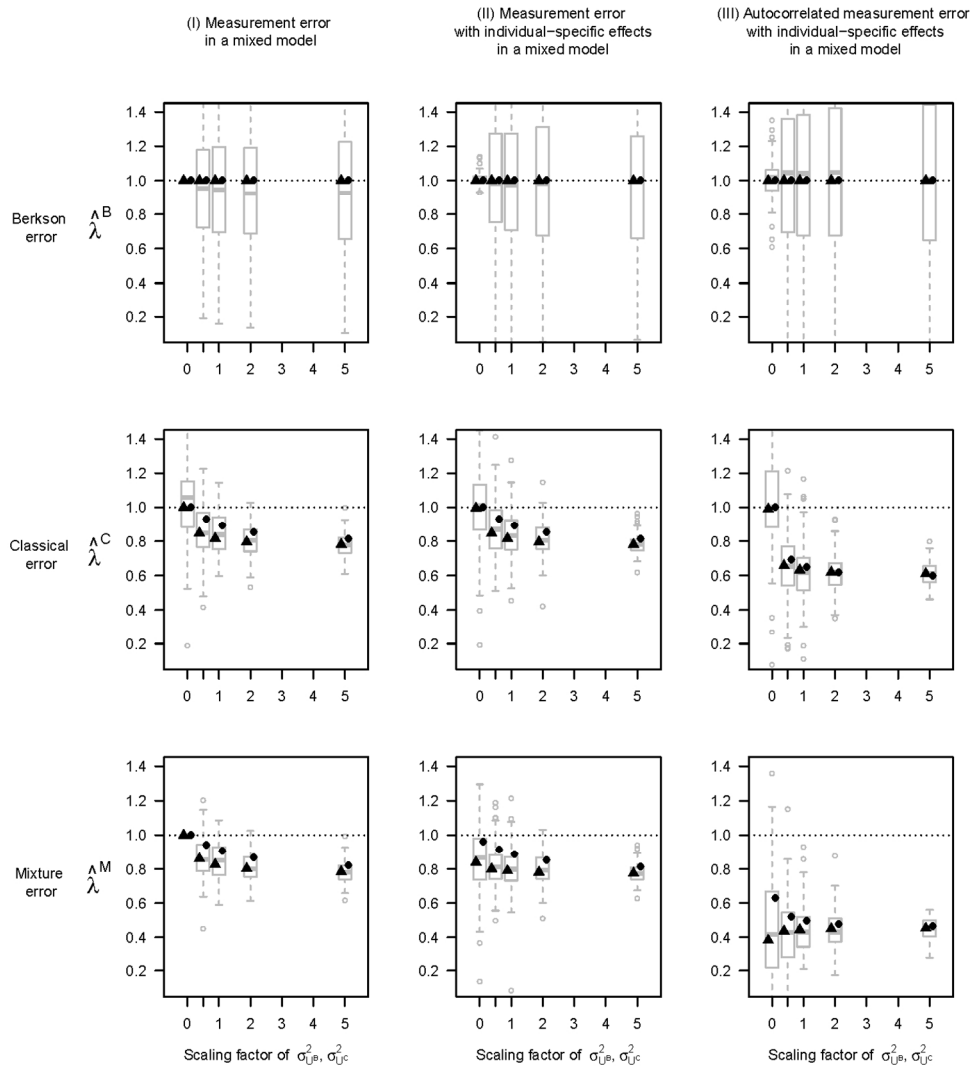
### Conflict of Interest

The authors have declared no conflict of interest.

### References

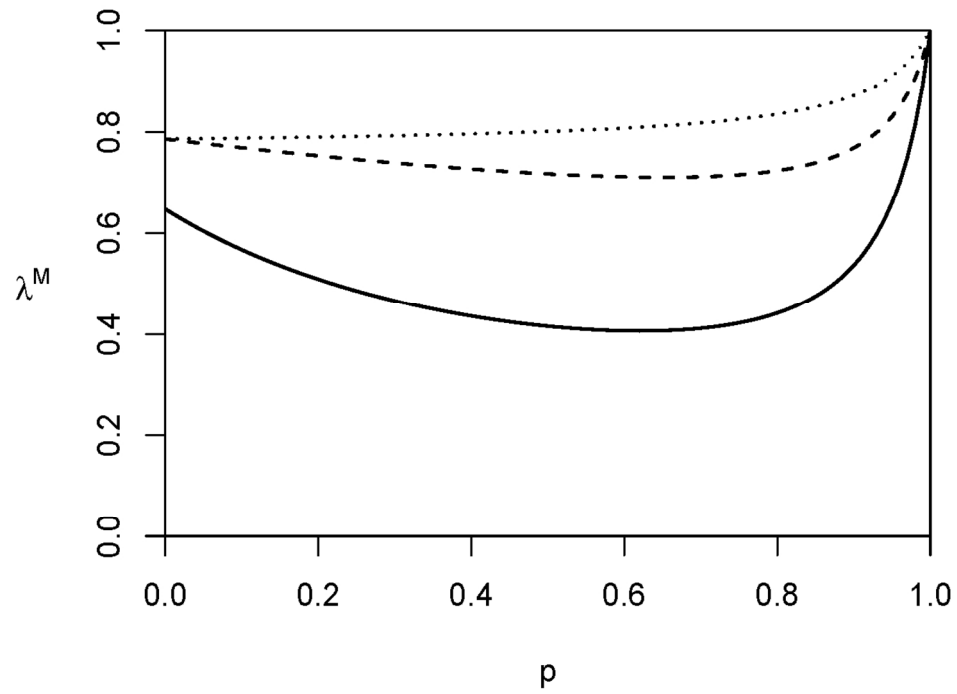
- Alexeeff, S., Carroll, R. J. and Coull, B. (2016). Spatial measurement error and correction by spatial SIMEX in linear regression models when using predicted air pollution exposures *Biostatistics* **17**, 377–389.
- Bergen, S., Sheppard L., Kaufman, J. D. and Szpiro, A. A. (2016). Multipollutant measurement error in air pollution epidemiology studies arising from predicting exposures with penalized regression splines *Journal of the Royal Statistical Society: Series C (Applied Statistics)*.
- Buzas, J. S., Stefanski, L. A. and Tosteson, T. D. (2005). Measurement error. In: *Handbook of Epidemiology* (eds. W. Ahrens and I. Pigeot), Springer, Berlin, Heidelberg, 729–765.
- Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models - A Modern Perspective*. Chapman & Hall/CRC, Boca Raton, 2. edition.
- Carroll, R. J., Delaigle, A. and Hall, P. (2007). Non-parametric regression estimation from data contaminated by a mixture of Berkson and classical errors: Non-parametric regression estimation *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 859–878.
- Deffner, V., Küchenhoff, H., Maier, V., Pitz, M., Cyrys, J., Breitner, S., Schneider, A., Gu, J., Gerschkat, U. and Peters, A. (2016). Personal exposure to ultrafine particles: Two-level statistical modeling of background exposure and time-activity patterns during three seasons. *Journal of Exposure Science and Environmental Epidemiology* **26**, 17–25.
- Diggle, P. J., Heagerty, P., Liang, K.-Y. and Zeger, S. L. (2002). *Analysis of Longitudinal Data*. Oxford University Press Inc., New York.
- Dominici, F., Sheppard, L. and Clyde, M. (2003). Health effects of air pollution: a statistical review. *International Statistical Review* **71**, 243–276.
- Fuller, W. A. (1987). *Measurement Error Models*. Wiley, New York.
- Gu, J., Kraus, U., Schneider, A., Hampel, R., Pitz, M., Breitner, S., Wolf, K., Hänninen, O., Peters, A. and Cyrys, J. (2015). Personal day-time exposure to ultrafine particles in different microenvironments. *International journal of hygiene and environmental health* **218**, 188–195.
- Hampel, R., Breitner, S., Schneider, A., Zareba, W., Kraus, U., Cyrys, J., Gerschkat, U., Belcredi, P., Müller, M., Wichmann, H. and Peters, A. (2012). Acute air pollution effects on heart rate variability are modified by SNPs involved in cardiac rhythm in individuals with diabetes or impaired glucose tolerance. *Environmental Research* **112**, 177–185.
- Hampel, R., Breitner, S., Zareba, W., Kraus, U., Pitz, M., Gerschkat, U., Belcredi, P., Peters, A. and Schneider, A. (2012). Immediate ozone effects on heart rate and repolarisation parameters in potentially susceptible individuals. *Occupational and Environmental Medicine* **69**, 428–436.
- Lütkepohl, H. (1984). Linear transformations of vector ARMA processes. *Journal of Econometrics* **26**, 283–293.
- Mallik, B., Hoffman, F. O. and Carroll, R. J. (2002). Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada Test Site. *Biometrics* **58**, 13–20.
- Muff, S. and Keller, L. F. (2015). Reverse attenuation in interaction terms due to covariate measurement error. *Biometrical Journal* **57**, 1068–1083.

- Peters, A., Hampel, R., Cyrys, J., Breitner, S., Geruschkat, U., Kraus, U., Zareba, W., and Schneider, A. (2015). Elevated particle number concentrations induce immediate changes in heart rate variability: a panel study in individuals with impaired glucose metabolism or diabetes. *Particle and Fibre Toxicology* **12**, 7.
- Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-Effects Models in S and S-Plus*. Springer, New York,
- Rosner, B., Willett, W. C. and Spiegelman, D. (1989). Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Statistics in Medicine* **8**, 1051–1069.
- Rückerl, R., Hampel, R., Breitner, S., Cyrys, J., Kraus, U., Carter, J., Dailey, L., Devlin, R. B., Diaz-Sanchez, D., Koenig, W., Phipps, R., Silbajoris, R., Soentgen, J., Soukup, J., Peters, A. and Schneider, A. (2014). Associations between ambient air pollution and blood markers of inflammation and coagulation/fibrinolysis in susceptible populations. *Environment International* **70**, 32–49.
- Schennach, S. (2013). Regressions with Berkson errors in covariates – A non-parametric approach. *The Annals of Statistics* **41**, 1642–1668.
- Schwartz, J. and Coull, B. A. (2003). Control for confounding in the presence of measurement error in hierarchical models. *Biostatistics* **4**, 539–553.
- Sheppard, L., Burnett, R. T., Szpiro, A. A., Kim, S. Y., Jerrett, M., Pope, C. A. and Brunekreef, B. (2011). Confounding and exposure measurement error in air pollution epidemiology. *Air Quality, Atmosphere & Health* **5**, 1–14.
- Spiegelman, D., Rosner, B. and Logan, R. (2000). Estimation and inference for logistic regression with covariate misclassification and measurement error in main study/validation study designs. *Journal of the American Statistical Association* **95**, 51–61.
- Thoresen, M. and Laake, P. (2000). A simulation study of measurement error correction methods in logistic regression. *Biometrics* **56**, 868–872.
- 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**, 1959–1971.
- Wang, C. Y. (2000). Flexible regression calibration for covariate measurement error with longitudinal surrogate variables. *Statistica Sinica* **10**, 905–921.
- Wang, N., Carroll, R. J. and Liang, K.-Y. (1996). Estimation in measurement error models with correlated replicates. *Biometrics* **52**, 401–411.
- 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**, 249–261.
- Yin, Z., Gao, W., Tang, M.-L. and Tian, G.-L. (2013). Estimation of nonparametric regression models with a mixture of Berkson and classical errors. *Statistics & Probability Letters* **83**, 1151–1162.
- Zeger, S. L., Thomas, D., Dominici, F., Samet, J. M., Schwartz, J., Dockery, D. and Cohen, A. (2000). Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environmental Health Perspectives* **108**, 419–426.



Simulation results for the empirical attenuation factor depicted in boxplots for increasing size of classical and Berkson measurement error; the theoretical estimations of  $\lambda$  with the exact value of  $T$  (using eq. (11) for  $\lambda^C$  and the combination of eq. (11) and eq. (14) for  $\lambda^M$ ) are marked with triangles and with circles for calculations assuming  $T \rightarrow \infty$  according to eq. (8) and eq. (9). Please note,  $\delta_v^2_B = 0.04$  and  $\delta_v^2_C = 0.03$ .

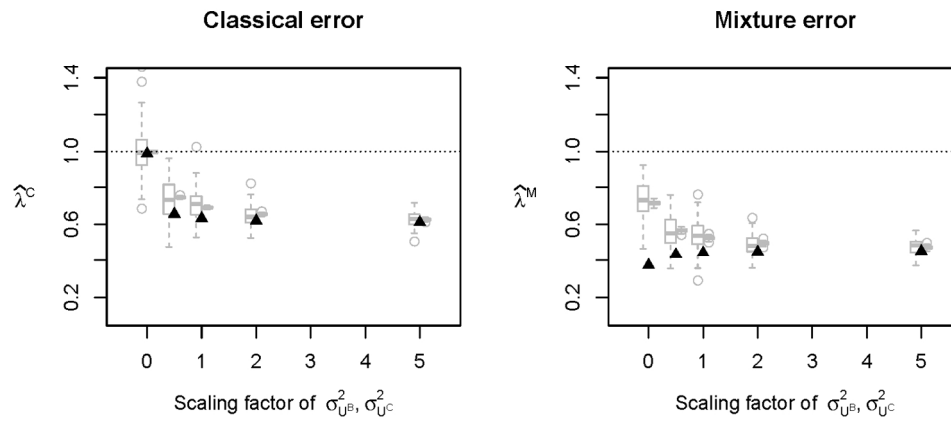
Figure 1  
185x228mm (200 x 200 DPI)



The role of the proportion  $p$  of observations affected by Berkson error for the attenuation factor  $\lambda^M$  in the case of mixture measurement error; dotted line: scenario (I), dashed line: scenario (II), solid line: scenario (III).

Figure 2

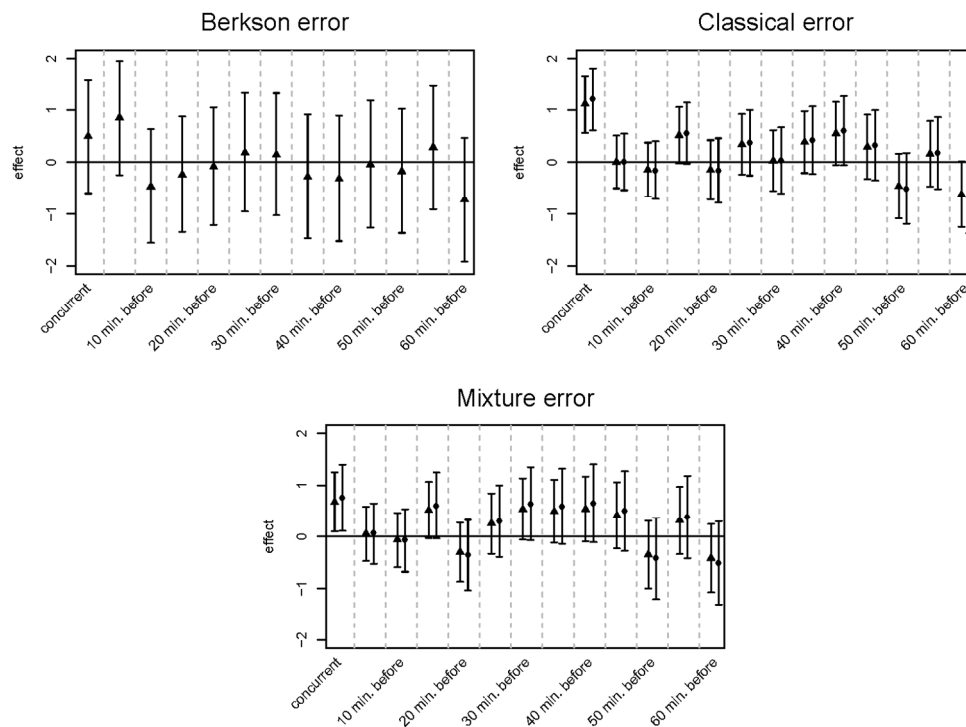
177x152mm (200 x 200 DPI)



Simulation results of scenario (III) for the empirical attenuation factor of  $\beta_1$  for increasing size of classical and Berkson measurement error in the presence of further covariates; left boxplots: empirical attenuation factor; right boxplots: estimated attenuation factors according to eq. (16) and the first versions of eq. (12) and Supporting Information, Appendix A.1.6 (with known break down times  $\mathbf{G}$ ); the triangles represent the estimated attenuation factors according to eq. (11) and eq. (14) assuming independence between the deficient and precisely measured covariates.

213x101mm (200 x 200 DPI)





Multiple regression analysis: Naive (triangles) and corrected (circles) effect estimates for the association between heart rate and individual PNC levels with Berkson, classical and mixture measurement error.

Figure 4

203x152mm (200 x 200 DPI)