Behavior Research Methods
https://doi.org/10.3758/s13428-020-01359-7

Two-part mixed-effects location scale models

Shelley A. Blozis[1] · Melissa McTernan[2] · Jeffrey R. Harring[3] · Qiwen Zheng[3]

Abstract

Longitudinal time use data afford the opportunity to study within- and between-individual differences, but can present challenges in data analysis. Often the response set includes a large number of zeros representing those who did not engage in the target behavior. Coupled with this is a continuous measure of time use for those who did engage. The latter is strictly positive and skewed to the right if relatively few individuals engage in the behavior to a greater extent. Data analysis is further complicated for repeated measures, because within-individual responses are typically correlated, and some respondents may have missing data. This combination of zeros and positive responses is characteristic of a type of semicontinuous data in which the response is equal to a discrete value and is otherwise continuous. Two-part models have been successfully applied to cross-sectional time use data when the research goals distinguish between a respondent's likelihood to engage in a behavior and the time spent conditional on any time being spent, as these models allow different covariates to relate to each distinct aspect of a behavior. Two-part mixed-effects models extend two-part models for analysis of longitudinal semicontinuous data to simultaneously address longitudinal decisions to engage in a behavior and time spent conditional on any time spent. Heterogeneity between and within individuals can be studied in unique ways. This paper presents applications of these models to daily diary data to study individual differences in time spent relaxing or engaged in leisure activities for an adult sample.

Keywords daily diary data · time use data · semicontinuous data · leisure activities

The collection of time use data is central to understanding many facets of human life. In the United States, for example, the Department of Labor supports the collection of time use data across a wide range of domains to conduct economic research, understand health, safety, and family and work-life balance, and make international comparisons. Time use data may be obtained for a single occasion in a target population, such as time devoted by students to academic study (Mucciardi, 2013), or for multiple time points to understand patterns of change in behaviors over time, such as how children's time spent with their parents changes over time (Sandberg & Hofferth, 2001).

Arguably, the manner in which time use is measured is a complex research enterprise, and the subsequent data analysis can present challenges. Here we consider longitudinal time use data that are measured using a semicontinuous scale with zero indicating that an individual did not engage in the behavior and values being positive otherwise. Importantly, we assume that there is interest in understanding correlates of the likelihood that an individual will engage in the behavior, and separate from this, interest in understanding correlates of the extent to which an individual engages in the behavior conditional on any time being spent, as well as the within-subject variation in time spent about an individual's conditional mean time across occasions. We assume that for each measurement occasion, respondents are asked to report time spent engaged in a behavior over a specified period of time, such as time spent on a task or engaged in an activity within a 24-hour

✉ Shelley A. Blozis
sablozis@ucdavis.edu

[1] Department of Psychology, University of California, Davis, CA, USA

[2] California State University Sacramento, Sacramento, CA, USA

[3] University of Maryland, College Park, MD, USA

Published online: 10 February 2020

period, and the question asked daily across multiple days. We assume a direct correspondence between each period that defines an occasion (e.g., a 24-hour period) and the period during which an individual decides to engage in the behavior (i.e., the same 24-hour period).

Daily diary of time spent relaxing, engaging in leisure activities

The National Study of Daily Experiences (NSDE) (Almeida, 2007) was designed to study time spent in various daily activities considered stressful to people. A sample of 1031 adults of the NSDE (excluding a subsample of twins and siblings) was randomly sampled from the Midlife in the United States (MIDUS) study. Interviews for the NSDE were conducted between March 1996 and April 1997. For the sample studied here, participants (54.4% women) were 47.4 years of age on average (SD = 13.2) and ranged in age from 20 to 74 years. The goal of the NSDE was to conduct telephone interviews to obtain self-reports of daily experiences over 8 consecutive days. The starting day (i.e., the day of the week) of the interview period varied between participants. The mean number of interview days for this sample was 7.0 (SD = 1.4, minimum = 1, maximum = 8).

We consider the daily measures of time spent relaxing or engaged in leisure activities, where a single question was used in the survey to pertain to these types of activities collectively. The proportions of the sample engaged in these activities by day of the week are summarized in Table 1, along with the means, standard deviations, and minimums and maximums of time spent in hours, conditional on any time being spent. Fig. 1 displays boxplots of the hours spent each day conditional on any amount of time being spent.

Key features of the data are worth noting. For each day, scores are positively skewed, with a subset of individuals

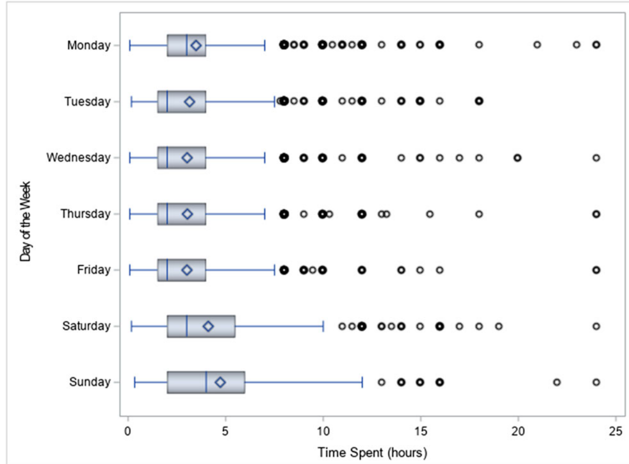

Fig. 1 Boxplots of positive reports of daily leisure time by day of week (n = 1031)

reporting no time spent engaged in leisure activities. When individuals were engaged, the reported daily time spent varied between individuals, with relatively few individuals reporting high amounts of time spent, as evidenced by the positive skew in the daily response distributions displayed in Fig. 1. In addition to the information provided in Table 1 and Fig. 1, individuals differed in their activities across days. At the individual level, the proportion of survey days that an individual reported engaging in leisure activities ranged from 0 to 1, indicating that some individuals reported no days of leisure activities and some reported leisure activities for all survey days. The daily mean time that individuals were engaged ranged from 0.33 to 23.8 hours, indicating a wide range in activity levels. Again at the individual level, the standard deviation of daily time measures when engaged ranged from 0 to 10.7, with some individuals having relative stability in their responses across days and others not. To capture these features of the data and allow for their study, a model is needed that can (1) distinguish between a response of zero and a positive

Table 1 Descriptive statistics for relaxing/leisure activities (n = 1031)

		Hours spent if engaged			
Day	Proportion of sample engaged	Mean	SD	Minimum	Maximum
Monday	.90	3.5	3.0	0.08	24
Tuesday	.89	3.2	2.6	0.17	18
Wednesday	.90	3.0	2.5	0.08	24
Thursday	.89	3.0	2.5	0.08	24
Friday	.88	3.0	2.4	0.08	24
Saturday	.92	4.1	3.1	0.17	24
Sunday	.95	4.7	3.2	0.33	24

Notes: Data source: National Study of Daily Experiences (NSDE) (Almeida, 2007) report of time spent, and thus characterize the likelihood that an individual will engage in the activity, and (2) characterize time spent conditional on an individual engaging in the activity, addressing the day-to-day variation in time spent that differs between individuals as well. As two-part mixed-effects models have been successful in addressing key features of data similar to these, we build upon this modeling framework for the analysis of time use data. These models and their extensions are described next.

Two-part models for semicontinuous time use data

Two-part models (Cragg, 1971) are useful for the analysis of time use data with many zeros from cross-sectional studies, and may be preferred to other models. These models rely on the creation of two variables from the original response: A binary variable represents whether an individual engaged in the behavior, and a continuous variable represents the time spent conditional on any time being spent. Analysis of the two variables is done independently, such as by applying a logistic regression model to the binary response and a linear regression model to the conditional continuous response (Duan, Manning, Morris, & Newhouse, 1983). Tobit regression has been used for time use measures with many zeros, but the method assumes left- or right-censoring of the response. Values below zero are not possible, however, and so the assumption of left-censoring is deemed inappropriate. Further, Tobit regression is not naturally set up to allow for different predictors to relate to the likelihood that one engages in a behavior and the degree of engagement (Stewart, 2013). Standard linear regression has also been applied to time use data that include zeros (Cawley & Liu, 2012), but this approach also cannot be used to make unique predictions about the likelihood of engaging in the behavior and the conditional degree of engagement, in addition to making no particular accommodation for zeros. Indeed, the goals of the data analysis need to be carefully considered when selecting a statistical framework for analysis.

Two-part mixed-effects models

A two-part mixed-effects model extends a two-part model for application to repeated measures of a semicontinuous variable, and thus provides an appealing approach to the analysis of repeated measures of time use data. Under the model, each model part includes one or more random effects to account for within-subject dependencies in the data. Olsen and Schafer (2001) and Tooze, Grunwald, and Jones (2002) describe a model based on a joint distribution linking separate mixed-effects models for a binary and a continuous response. The two models are linked by covariances between their respective random effects, and as a result, estimation of the two model parts is simultaneous. These models are flexible in that different response distributions may be used to represent each model part (e.g., Liu, Cowen, Strawderman, & Shih, 2010). This can be particularly important for time use data, which in addition to typically including many zeros, are commonly positively skewed. Importantly, these models retain the ability to use different covariates to predict engagement in a behavior and the level of engagement conditional on any positive level of engagement.

Two-part mixed-effects models share features common to mixed-effects models more generally. For example, the models do not require complete data, and data are assumed to be missing at random. A general form of the model is assumed to apply to all members of a population with one or more of the coefficients of each model part being subject-specific. The covariance structure is separated into a within- and a between-subject component. The within-subject covariance structure describes the variation and possible covariation of scores between occasions. The between-subject covariance structure describes the variation and possible covariation of the random effects that are used to characterize the repeated response measures. Maximum likelihood (ML) (Olsen & Schafer, 2001) or Bayesian estimation (Xing et al. 2017) may be used for estimation of these models.

Using daily diary data from a large study of daily stressors in adults, we show how these models may be formulated to study between-subject differences and within-individual variability in semicontinuous data by relaxing some of the assumptions typically made when applying these models. It is common in applications of two-part mixed-effects models to assume homogeneity of variance of a random coefficient across subjects and of the variance of the occasion-level residual across occasions and subjects. To relax both of these assumptions, we incorporate features of a mixed-effects location scale model (Hedeker, Mermelstein, & Demirtas, 2008; Hedeker & Nordgren, 2013) into a two-part mixed-effects model. A mixed-effects location scale model is a model for a single normally distributed variable that includes a sub-model for the variance of a random intercept so that it can depend on within- and between-subject covariates, in addition to a sub-model for the variance of the occasion-level residual so that it can depend on within- and between-subject covariates. Additionally, the model for the variance of the occasion-level residual can include a random subject effect to allow between-subject heterogeneity of variance even after adjusting for covariates.

Analytical strategy

In analyzing data from the NSDE, we assumed that respondents had an opportunity each day to engage in relaxing or leisure activities and that a choice was made each day to engage or not engage in such activities. We began by fitting a set of unconditional models in which the response distribution of the continuous model part was assumed to follow one of three distributions: normal, log-normal or gamma distribution. The latter two distributions are positive and continuous distributions that may be well suited to addressing the positive skew observed in the positive time use responses. Using the best-fitting of these three models, covariates were added to the model under three different model formulations. First, a two-part mixed-effects model was applied in the usual way to study the log odds of engagement and the daily mean time when engaged. An extended model was then developed to include the study of possible heterogeneity in both the within- and the between-subject covariance structures. This permitted us to study any day-to-day variation in time spent about an individual's mean time when engaged and any between-subject variation in the individual log odds and daily mean time, both after adjusting for the effects of model covariates that entered the mean structure of the model. This extended model was then reduced by evaluating the need for particular covariates. Fixed effects were evaluated using likelihood ratio tests with a significance level of .05. Interpretation of the three models that include covariates is delayed until after the models are developed and estimates provided.

PROC NLMIXED for SAS version 9.4 software was used to carry out the analyses. The process of fitting a mixed-effects model to complex data can be a challenge. The paper by Kiernan, Tao, and Gibbs (2012) addresses issues for fitting linear and nonlinear mixed-effects models in SAS, and we relied on many of the suggestions provided in their paper—for example, providing reasonable starting values, and beginning with relatively simple models and building up a model by increasing its complexity. For the empirical analyses presented here, we began by fitting fixed effects models and gradually built up the model by increasing the complexity (e.g., adding random effects), while updating the starting values for each model based on estimates obtained by simpler versions of the model. As we added new parameters to the model, we used what we thought were reasonable starting values. With regard to estimation procedures, SAS PROC NLMIXED includes a few options. We found that using nonadaptive Gaussian quadrature, as opposed to the default method of adaptive Gaussian quadrature, was best for obtaining a solution. Taking guidance from a simulation study by Carlin, Wolfe, Brown, and Gelman (2001), we used nonadaptive Gaussian quadrature with a high number (20) of quadrature points.

Unconditional models for daily leisure time activity data

We began by fitting an unconditional two-part mixed-effects model to the data. Let y_{ij} be the observed response for individual i on day j, where $i = 1,..., N$ and $j = 1,..., n_i$, with N denoting the total number of subjects and n_i the number of survey days for individual i. Let t_{ij} be the day that y_{ij} was observed. From y_{ij} two new variables were created: $u_{ij} = 1$ if $y_{ij} > 0$ and $u_{ij} = 0$ if $y_{ij} = 0$ (if y_{ij} was missing, then u_{ij} was set to missing); $m_{ij} = y_{ij}$ if $y_{ij} > 0$ and was missing otherwise. Next, let η_{ij} be the logit of the individual engaged in the activity:

$$\eta_{ij} = \log\left[P\left(u_{ij} = 1\right) / \left(1 - P\left(u_{ij} = 1\right)\right)\right] \tag{1}$$

The logit in Eq. (1) was assume to follow a mixed-effects model:

$$\eta_{ij} = \alpha_0 + a_i \tag{2}$$

where α_0 is the mean logit and a_i is a random subject effect assumed to be independent and identically distributed (i.i.d.) across subjects as normal with mean equal to 0 and variance φ_a^2. The variance parameter φ_a^2 characterizes between-subject variation in the logit.

A positive report of time spent was modeled using a linear mixed-effects model:

$$m_{ij} = \gamma_0 + b_i + \varepsilon_{ij} \tag{3}$$

where γ_0 is the mean time spent and b_i is a random subject effect assumed to be i.i.d. normal across subjects with mean equal to 0 and variance φ_b^2. The variance φ_b^2 represents between-subject variability in the individual mean times. The residual ε_{ij} is the day- and individual-specific part of the daily response not accounted for by the subject-specific model given by $\gamma_0 + b_i$. Across days, the set of residuals $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1},..., \varepsilon_{in_i})'$ was assumed to be i.i.d. normal (with the assumption of normality relaxed later) across subjects with mean equal to $\mathbf{0}$ and covariance matrix $\boldsymbol{\Theta}_\varepsilon$. The residuals were assumed to be independent between days with constant variance: $\boldsymbol{\Theta}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_{ni}$, where σ_ε^2 was a common variance and $\mathbf{I}$ was an identity matrix of order n_i. Initially, the residual covariance structure was assumed to be homogeneous across individuals, but the dimensions of $\boldsymbol{\Theta}_\varepsilon$ could differ between individuals due to missing data.

Table 2 Indices of model fit of two-part mixed models (n = 1031)

Model	q	-2lnL	AIC	BIC	Models Compared	$\chi^2(df)$	*P*-value
Unconditional models							
A_1	6	34567	34579	34608			
A_2	6	30424	30436	30465			
A_3	6	30611	30623	30652			
Conditional models							
B_1	22	29776	29820	29929	A_2 vs. B_1	835(16)	< .0001
B_2	37	29312	29386	29568	B_1 vs. B_2	464(15)	< .0001
B_3	24	29323	29371	29489	B_2 vs. B_3	11(13)	.61

Notes: -2*lnL* is -2 times the log-likelihood. AIC = Akaike information criterion. BIC = Bayesian information criterion. χ^2 is the statistic for a deviance test with *df* equal to the difference in model parameters. With regard to the continuous model part, Model A_1 assumes a normal probability response distribution, Model A_2 assumes a log-normal probability response distribution: $f\left(m_{ij}|\mu_{\log(m_{ij})},\sigma^2_{\log(m_{ij})}\right)=\frac{1}{m_{ij}\sqrt{2\sigma^2\pi}}exp\left\{-\frac{1}{2}\left(\frac{\left(\log(m_{ij})-\mu_{\log(m_{ij})}\right)}{\sigma^2_{\log(m_{ij})}}\right)\right\}$, and Model A_3 assumes a gamma probability response distribution with shape (γ) and scale (θ) parameters: $f\left(m_{ij}|\gamma,\theta\right)=\frac{m_{ij}^{\gamma-1}exp\left(-m_{ij}/\theta\right)}{\theta^{\gamma}\Gamma(\gamma)}$, where $\Gamma(\gamma)$ is the standard gamma function of the scale parameter θ. Models A_1–A_3 assume homogeneity of the within- and between-subject covariance structures. Model B_1 assumes homogeneity of the within- and between-subject covariance structures. Model B_2 extends Model B_1 to allow heterogeneity of the within- and between-subject covariance structures. Model B_3 is a reduced model based on Model B_2 that excludes some covariates used in the different model parts

The two sub-models for η_{ij} in Eq. (2) and m_{ij} in Eq. (3) were joined at the second level by a covariance between the random effects, a_i and b_i, of the two models:

$$\boldsymbol{\Phi}=\begin{bmatrix}\varphi_a^2 & \\ \varphi_{ba} & \varphi_b^2\end{bmatrix}$$

where the covariance φ_{ba} represents the linear relationship between the individual-level logit and mean time spent across days. For instance, a positive covariance would indicate that higher log odds (or likelihood) of engaging in the activity corresponds to a higher daily mean time when engaged in the activity.

To investigate the shape of the response distribution of m_{ij}, three versions of the two-part mixed-effects model were fit to the data. In Model A_1, m_{ij} was assumed to be normally distributed[1]. Although m_{ij} is strictly positive, if the responses are bell-shaped, then a normal distribution could provide a suitable approximation to the response distribution. Alternatively, if responses are positively skewed, as expected if increasingly fewer individuals engaged in the behavior to a greater extent, then a log-normal or gamma distribution, both of which are positive and continuous distributions, may approximate the distribution. Model A_2 assumed that m_{ij} was log-normally distributed and Model A_3 assumed that m_{ij} followed a gamma distribution. The upper part of Table 2 includes the -2*lnL*, AIC and BIC values for Models A_1–A_3. The model that assumed a log-normal distribution for m_{ij} (Model A_2) provided the best overall fit, based on the model having the lowest AIC (or one could use the BIC for an equivalent conclusion) value. This model was provisionally accepted as the best-fitting unconditional model. From this model, the estimated fixed effect of the logit was 3.2 (SE = 0.10), and the estimated mean log time spent was 0.93 (SE = 0.02). The estimated variance of the random intercept of the logit model was 2.98, and that of the log-normal model was 0.24. The estimated variance of the residual of the log-normal model was 0.39. Under the logistic model, the estimated intraclass correlation coefficient was $\hat{\rho}$ = 2.98/((π^2/3)+2.98) = .48, and that for the log-normal model was $\hat{\rho}$ = 0.24/(0.39+0.24) = .39, although it should be noted that both estimates assume homogeneity of variance across individuals and measurement occasions. In both model parts, an appreciable portion of the variation in responses is due to both within- and between-subject differences.

Conditional models for daily leisure time activity data

Covariates were used to study reported time spent engaged in leisure activities. The covariates were as follows: Age_i (centered to the sample mean age of 47.4 years), gender (coded as Fem_i = 1 if female and Fem_i = 0 if male), and whether the survey took place on a weekday or weekend (Day_{ij} was coded

[1] A t-distribution might also be considered if the response was expected to have heavy tails.

as 1 = Saturday or Sunday and 0 = Monday, Tuesday, Wednesday, Thursday or Friday). Day_{ij} was centered about the person's proportion of survey days that fell on a weekend, and the proportion of weekend days ($PropDay_i$) was included to statistically adjust for between-subject differences. The number of survey days that fell on a weekend ranged from 0 to 3 across respondents, with a mean of 1.7 and standard deviation of 0.6; corresponding to this, the mean proportion of survey days falling on a weekend was .23 (SD = .08, minimum = 0, maximum = .67).

The first model fit to the data, henceforth called Model B_1, represents a typical application of a two-part mixed-effects model. To include covariates, the model for the logit in Eq. (2) was extended:

$$\begin{aligned}\eta_{ij} = \alpha_0 + \alpha_1 Day_{ij} + \alpha_2 Fem_i + \alpha_3 Age_i \\ + \alpha_4 Day_{ij} {*} Fem_i + \alpha_5 Day_{ij} {*} Age_i \\ + \alpha_6 Fem_i {*} Age_i + \alpha_7 Day_i {*} Age_i {*} Fem_i \\ + \alpha_8 PropDay_i + a_i \qquad (4)\end{aligned}$$

where α_0 is the logit when all covariates are equal to 0, the coefficients α_1, α_2 and α_3 are the effects of Day_{ij}, Fem_i and Age_i, each conditional on the interacting covariates being equal to 0, the coefficients α_4, α_5, and α_6 are the two-way interaction effects, each also conditional on the interacting covariates being equal to 0, and α_7 is the three-way interaction effect between the three covariates. Finally, α_8 is the effect of $PropDay_i$ conditional on all other covariates. The model includes a random subject effect a_i assumed to be i.i.d. normal across subjects with mean equal to 0 and variance φ_a^2. The variance φ_a^2 characterizes between-subject variation in the logit after accounting for variation due to covariates and their interactions. Under a mixed-effects logistic regression model, the effects in (4) control for the random subject effect a_i (Neuhaus, Kalbfleisch, & Hauck, 1991).

The model for the log of the positive measure of time spent in Eq. (3) was extended:

$$\begin{aligned}\log(mij) = \gamma_0 + \gamma_1 Day_{ij} + \gamma_2 Fem_i + \gamma_3 Age_i \\ + \gamma_4 Day_{ij} * Fem_i + \gamma_5 Day_{ij} * Age_i \\ + \boldsymbol{\gamma_6} Fem_i * Age_i + \boldsymbol{\gamma_7} Day_i * Age_i * Fem_i \\ + \boldsymbol{\gamma_8} PropDay_i + \mathrm{b}_i + \varepsilon_{ij} \qquad (5)\end{aligned}$$

where γ_0 is the mean log time when all covariates are equal to 0, the coefficients γ_1, γ_2 and γ_3 are the effects of Day_{ij}, Fem_i and Age_i, each conditional on the interacting covariates being equal to 0, and the coefficients γ_4, γ_5 and γ_6 are the effects of the two-way interactions between the three covariates, also conditional on the interacting covariates being equal to 0. The last coefficient, γ_8, is the effect of $PropDay_i$ conditional on all other covariates. The model includes a random subject effect b_i assumed to be i.i.d. normal across subjects with mean equal to 0 and variance φ_b^2. The variance φ_b^2 represents between-subject variation in the individual means in log time after accounting for variation due to covariates and their interactions.

The residual ε_{ij} in Eq. (5) is the part of the response not accounted for by the subject-specific model that now includes covariates. The residuals were assumed to be i.i.d. log-normal across subjects with mean equal to 0 and covariance matrix $\boldsymbol{\Theta}_\varepsilon$. The residuals were again assumed to be independent between days with constant variance. As was done for the unconditional model, the two sub-models for η_{ij} in Eq. (4) and m_{ij} in Eq. (5) were joined at the second level by a covariance between the random effects, a_i and b_i, with both a_i

Table 3 ML estimates of fixed effects (n = 1031)

Parameter	Model B_1	Model B_2	Model B_3
Within-subject			
α_0	3.5 (0.14)	3.6(0.18)	3.5 (0.15)
Day_{ij}, α_1	0.66(0.20)	0.67(0.20)	0.67(0.12)
Fem_i, α_2	−0.58(0.16)	−0.72(0.22)	−0.59(0.16)
Age_i, α_3	0.007(0.01)	0.015(0.012)	0.016(0.006)
$Day_{ij}*Fem_i$, α_4	0.006(0.25)	−0.013(0.25)	
$Day_{ij}*Age_i$, α_5	−0.014(0.02)	−0.013(0.017)	
Fem_i*Age_i, α_6	0.013(0.01)	0.012(0.013)	
$Day_{ij}*Age_i*Fem_i$, α_7	−0.0003 (0.020)	0.0001(0.020)	
$PropDay_i$, α_8	1.7(0.98)	1.4(0.097)	1.6(0.98)
Between-subject			
γ_0	1.0(0.025)	1.0(0.027)	1.0(0.026)
Day_{ij}, γ_1	0.40(0.026)	0.36(0.025)	0.36(0.017)
Fem_i, γ_2	−0.18(0.034)	−0.19(0.035)	−0.19(0.034)
Age_i, γ_3	0.010(0.002)	0.010(0.0020)	0.010(0.002)
$Day_{ij}*Fem_i$, γ_4	0.011(0.035)	−.0021(0.034)	
$Day_{ij}*Age_i$, γ_5	−0.007(0.002)	−.007(0.002)	−0.006(0.001)
Fem_i*Age_i, γ_6	−0.005(0.003)	−.005(0.003)	−0.005(0.002)
$Day_{ij}*Age_i*Fem_i$, γ_7	0.003(0.003)	.0024(0.0025)	
$PropDay_i$, γ_8	−0.34(0.21)	−0.32(0.22)	−0.32(0.22)

Notes: Standard errors are in parentheses. Model B_1 assumes homogeneity of the within- and between-subject covariance structures. Model B_2 assumes heterogeneity of the within- and between-subject covariance structures. Model B_3 is a reduced form of Model B_2 to exclude statistically nonsignificant covariate effects.

Table 4 ML estimates of within- and between-covariance structures (n = 1031)

Parameter	Model B_1	Model B_2	Model B_3
Within-subject			
$\ln(\sigma_\varepsilon^2)$	−1.0 (0.019)	−1.3(0.043)	−1.3(0.04)
Day_{ij}, τ_1		0.23(0.079)	0.19(0.05)
Fem_i, τ_2		0.15(0.057)	0.14 (0.06)
Age_i, τ_3		−0.008(0.003)	−0.007 (0.002)
$Day_{ij}*Fem_i$, τ_4		−0.063(0.10)	
$Day_{ij}*Age_i$, τ_5		−0.010(0.006)	
Fem_i*Age_i, τ_6		0.003(0.004)	
$Day_{ij}*Age_i*Fem_i$, τ_7		0.009(0.008)	
$PropDay_i$, τ_8		−0.48(0.40)	−0.51(0.40)
Between-subject			
$\ln(\varphi_a^2)$	1.1(0.11)	1.2(0.17)	1.1(0.11)
Fem_i, λ_1		−0.22(0.21)	
Age_i, λ_2		0.01(0.01)	
$\ln(\varphi_b^2)$	−1.5(0.06)	−1.3(0.07)	−1.3(0.07)
Fem_i κ_1		-.22(0.09)	−0.23(0.09)
Age_i, κ_2		-.0011(0.003)	
φ_{ab}	0.34(0.05)	0.38(0.05)	0.37(0.05)
φ_w^2		0.39(0.04)	0.39(0.04)
φ_{wa}		−0.51(0.09)	−0.51(0.09)
φ_{wb}		−0.18(0.02)	−0.17(0.02)
Additional parameters and their estimates			
σ_ε^2	0.35(0.01)	0.28(0.01)	0.28(0.01)
φ_a^2	2.9(0.33)	3.4(0.58)	2.9(0.34)
φ_b^2	0.23(0.013)	0.27(0.02)	0.27(0.02)
ρ_{ab}		.40	.37
ρ_{wa}		-.45	-.51
ρ_{wb}		-.54	-.54

Notes: Standard errors are in parentheses. Model B_1 assumes homogeneity of the within- and between-subject covariance structures. Model B_2 assumes heterogeneity of the within- and between-subject covariance structures. Model B_3 is a reduced form of Model B_2 to exclude statistically nonsignificant covariate effects.

and b_i now conditional on the covariates. Thus, the variances of a_i and b_i and their covariance are now the variances and covariance of the random effects conditional on covariates.

ML estimates of the fixed effects of Model B_1 are shown in the first column of results in Table 3. ML estimates of the common variance σ_ε^2 and the variances and covariance (and corresponding correlation) of the random effects a_i and b_i are in the first column of results in Table 4. Using a deviance test, the difference in fit between Models A_2 and B_1 was statistically significant ($\chi^2(16) = 835$, $P < .0001$), indicating an improvement in model fit after the addition of the covariates to predict the logit and log time spent (see Table 2). As stated previously, we delay model interpretation until all three models have been reported.

Extended model

Model B_1 represents a typical application of a two-part mixed-effects model in which the model assumes homogeneity of the within- and between-subject covariance structures. In Model B_2, the model is extended to allow for heterogeneity in both covariance structures. First, the within-subject variance σ_ε^2 is allowed to vary according to both within- and between-subject covariates. That is, the variance is a function of variables that are measured daily along with the response variable and those that are subject-specific. This is useful in situations in which heterogeneity of variance is related to covariates. Particularly for time use measures, characterizing individual differences in daily variation in a behavior may be an important element to better understanding time use. The model for the within-subject variance also includes a random subject effect so that, even after adjusting for the effects of covariates, the variance may differ between subjects. The inclusion of a random subject effect can be useful if, for example, additional sources of heterogeneity of variance are unknown. Second, the between-subject covariance structure is allowed to vary according to between-subject covariates. That is, the variance of each random effect is a function of subject-specific covariates. For instance, whereas Model B_1 assumes homogeneity of the variances of the intercepts of both the logit and log-normal models, Model B_2 permits the variances to depend on Fem_i and Age_i, as described later.

Within-subject heterogeneity Under Model B_1, the residual variance of the continuous model part σ^2_ε was assumed to be homogeneous such that variation in an individual's observed scores about their respective average response across days was the same across days and individuals. Under Model B_2, the assumption of homogeneity of variance across days and individuals was relaxed by allowing the residual variance to be predicted by covariates. This allowed us to study whether the covariates accounted for variation in the residual variance beyond what the covariates could account for in predicting an individual's daily time spent engaged in leisure activities. Additionally, a random subject effect was added to the variance model to test whether between-subject heterogeneity of the within-subject variation remained after accounting for the effects of the covariates on the residual variance. The model for the variance $\sigma^2_{\varepsilon i}$, that now includes subscripts to show that it can vary according to both daily and subject-level covariates, is given by

$$\sigma^2_{\varepsilon i} = \exp\left(\tau_0 + \tau_1 Day_{ij} + \tau_2 Fem_i + \tau_3 Age_i + \tau_4 Day_{ij}{*}Fem_i + \tau_5 Day_{ij}{*}Age_i + \boldsymbol{\tau_6} Fem_i{*}Age_i + \tau_7 Day_i{*}Age_i{*}Fem_i + \tau_8 PropDay_i + w_i\right) \quad (6)$$

where τ_0, when exponentiated, is the common (geometric) residual variance when the covariates are equal to zero and $w_i = 0$. For any of the effects τ_1–τ_8, a positive value indicates that an increase in a covariate corresponds to an increase in the variance, and a negative value indicates that an increase in a covariate corresponds to a decrease in the variance. Additionally, the model for the variance in Eq. (6) includes a random subject effect w_i that may covary with the other subject-level random effects at the second level of the model:

$$\mathbf{\Phi} = \begin{bmatrix} \varphi^2_a & & \\ \varphi_{ba} & \varphi^2_b & \\ \varphi_{wa} & \varphi_{wb} & \varphi^2_w \end{bmatrix}$$

where φ_{wa} and φ_{wb} are the covariances between the random subject effect of the within-subject variance model in (6) and the random effects of the logistic model in Eq. (4) and linear model in Eq. (5), respectively, and φ^2_w is the variance of the random subject effect of the within-subject variance model in (6). The covariances φ_{wa} and φ_{wb} represent the linear relationships between the individual-level logits and individual mean log time measures with the random effect of the within-subject variance model in (6). Positive covariances would indicate that as either the logit or mean time spent increases, there is a corresponding increase in the daily variation in time spent about an individual's mean log time. Negative covariances would indicate that as either the logit or mean log time increases, there is a corresponding decrease in the daily variation in time about an individual's mean log time.

Between-subject heterogeneity Under Model B_1, the variance of the random effects of the logit model and the log-normal model were assumed to be homogeneous, such that variation in these random effects, conditional on the covariates of the respective models in Eqs. (4) and (5), was the same across individuals. Here the assumption of homogeneity of variance across individuals is relaxed by allowing each variance of these two random effects to be predicted by the between-subject covariates Fem_i, Age_i and their interaction (between-subject variation in these random effects was not thought to vary according to the proportion of survey days that fell on a weekend, and so $PropDay_i$ was not included in these models). This allowed us to study whether these covariates could account for variation in the random effects after controlling for these covariates as predictors of the individual's logit or mean log time. The models for the random effect variances φ^2_a and φ^2_b now can vary according to subject-level covariates as follows:

$$\varphi^2_a = \exp(\lambda_0 + \lambda_1 Fem_i + \lambda_2 Age_i + \lambda_3 Fem_i{*}Age_i) \quad (7)$$

$$\varphi^2_b = \exp(\kappa_0 + \kappa_1 Fem_i + \kappa_2 Age_i + \kappa_3 Fem_i{*}Age_i) \quad (8)$$

where λ_0 and κ_0, when exponentiated, are the (geometric) variances of the respective random effects when the covariates are equal to zero. The effects λ_1, λ_2 and λ_3, relating to Fem_i, Age_i and their interaction, respectively, are predictors of the random intercept variance of the logistic model, and κ_1, κ_2 and κ_3, relating to Fem_i, Age_i and their interaction, respectively, are predictors of the random intercept variance of the log-normal model. Positive values of these covariates would correspond to a greater degree of between-subject variation in the random effect (after adjusting for the covariates in Eqs. (4) and (5)), and negative values would correspond to a lower degree of between-subject variation in the random effect (again, after adjusting for the covariates in Eqs. (4) and (5)). As the variances of the random effects are functions of covariates, the variance-covariance matrix of the random effects is interpreted as having elements that are adjusted for the effects of covariates. Reported in Table 2, a deviance test compares the fit between Models B_1 and B_2 ($\chi^2(15) = 464$, $P < .0001$), indicating an improvement in fit after allowing for heterogeneity of the within- and between-subject covariance structures. ML estimates of the fixed effects are provided in the second column of results in Table 3, and estimates of the parameters that defined the within- and between-subject covariance structures are given in the second column of results in Table 4. Again, we delay model interpretation until later.

Reduced model

The effects of the covariates in Model B_2 were evaluated and a reduced model (henceforth called Model B_3; the Appendix provides syntax for fitting Model B_3) was formed to provide a data description using a model that was parsimonious relative to Model B_2. The three-way interaction between Fem_i, Age_i and Day_{ij}, was evaluated first, followed by the two-way interactions and then the main effects; an effect was excluded from the model if it was not statistically significant at the .05 level. For the final model obtained, a deviance test compares the fit between Models B_2 and B_3 ($\chi^2(13) = 11$, $P = .61$), indicating no appreciable decrease in fit after excluding a select subset of covariates (see Table 2).

From Table 3, the estimated effects of the covariates for the mixed-effects logistic regression model are interpreted here, noting that the estimated effects are conditional on the random subject effects. The estimated log odds that an individual engaged in leisure activities, if the respondent was male, at the sample mean age of 47.4 years, whose proportion of survey days taking place on a weekend was equal to the mean proportion of 0.23, and whose random effect for the logit intercept was equal to zero, was $\widehat{\alpha}_0 = 3.5$ (SE = 0.15). The log odds tended to be higher if the survey day fell on a weekend ($\widehat{\alpha}_1 = 0.67$, SE = 0.12) and lower if the respondent was female ($\widehat{\alpha}_2 = -0.59$, SE = 0.16). Older individuals were also more likely to engage ($\widehat{\alpha}_3 = 0.02$, SE = 0.006). These effects were each adjusted for each other and for between-subject differences in the proportion of weekend days surveyed ($\widehat{\alpha}_8 = 1.6$, SE = 0.98), as well as controlling for the random effect a_i.

The estimated mean log time spent, conditional on any time being spent, for males at the sample mean age and at the sample mean proportion of survey days falling on a weekend was $\widehat{\gamma}_0 = 1.0$ (SE = 0.03), which when exponentiated is about 2.7 (geometric) mean hours per day. Mean log time tended to be higher for weekends at the sample mean age of 47.4 years ($\widehat{\gamma}_1 = 0.36$, SE = 0.02), with the effect of a weekend being attenuated according to an increase in a respondent's age ($\widehat{\gamma}_5 = -0.006$, SE = 0.001). For males, older age corresponded to higher mean log time ($\widehat{\gamma}_3 = 0.010$, SE = 0.002), with the relative effect of age attenuated for females ($\widehat{\gamma}_6 = -0.005$, SE = 0.002). Relative to males, females at the sample mean age of 47.4 years had a lower mean log time ($\widehat{\gamma}_2 = -0.19$, SE = 0.034). Each of these effects was adjusted for the other effects, as well as for between-subject differences in the proportion of weekend days surveyed ($\widehat{\gamma}_8 = -0.32$, SE = 0.22).

From Table 4, within-subject variance of daily log time about an individual's mean log time across days was greater if the survey took place on a weekend versus a weekday ($\widehat{\tau}_1 = 0.19$, SE = 0.05), greater if the respondent was female ($\widehat{\tau}_2 = 0.14$, SE = 0.06), and reduced for older respondents ($\widehat{\tau}_3 = -0.007$, SE = 0.002). Each of these effects was adjusted for the other effects, as well as for between-subject differences in the proportion of weekend days surveyed ($\widehat{\tau}_8 = -0.51$, SE = 0.40).

Between-subject variation in the logit (conditional on covariates) did not differ appreciably between males and females or with regard to age, indicating comparable degrees of between-subject variation in the individual logits at each level of the two covariates. Between-subject variation in the log mean time (conditional on covariates) was lower for females relative to males ($\widehat{\kappa}_1 = -0.23$, SE = 0.09), indicating greater homogeneity among females relative to males. Variation in the log mean time did not vary appreciably with regard to age.

Finally, the random effects, after adjusting for the effects of covariates, were moderately related to each other. In particular, a higher likelihood to engage in leisure activities was positively related to the daily mean log time ($\widehat{\rho}_{ab} = .37$, 95% CI: [.32, .42]), but likelihood to engage and daily mean log time were both negatively related to the within-subject daily variation in time spent ($\widehat{\rho}_{wa} = -.51$, 95% CI: [−.55, −.46]; $\widehat{\rho}_{wb} = -.54$, 95% CI: [−.50, −.58], respectively), indicating that greater stability in time spent tended to correspond to a lower likelihood to engage and a lower daily mean log time across days. It is important to mention as a final note that a deviance test done to evaluate the need for the random effect in the within-subject variance model indicated that between-subject heterogeneity of the within-subject variance remained after adjusting for the effects of covariates ($\chi^2(3) = 416$, $P < .0001$).

Comparing the different models

The key difference between the typical application of a two-part mixed-effects model (here, Model B_1) and an extended version (here, Model B_2), is that the latter allows for heterogeneity of both the within- and between-subject covariance structures. Thus, greater insight can be gained about a behavior by allowing covariates to predict the different sources of variation. For the leisure time data, covariates were particularly informative about heterogeneity of variation in daily time about a person's mean time, and heterogeneity of variance remained even after accounting for these covariates. With regard to heterogeneity of variance in the random effects at the subject level, between-subject variation in the mean log time (adjusting for covariates) was greater for males relative to females, suggesting that males differ from each other to a greater extent than females with regard to daily mean time.

Discussion

Two-part models are useful in addressing research questions that relate to the likelihood that individuals will engage in a behavior, and separate from this, the extent of engagement given that any time was spent. In this way, two-part models may be preferred over other methods that do not make this distinction in a measured response. This is particularly relevant for time use data if a segment of a population does not engage in the target behavior and it is important to understand not only whether people engage in the behavior, but if they do, to what extent. A two-part model can be extended to a two-part mixed-effects model given repeated-measures data. A two-part mixed-effects model addresses dependencies in repeated measures by including random effects in each of the two model parts. The model permits one to study of the trajectories of the likelihood that individuals will engage in a behavior over time and the trajectories of the extent of engagement when engaged.

In a typical application of a two-part mixed-effects model, homogeneity of the within-subject and the between-subject covariance structures is assumed. For the within-subject covariance structure, an assumption of homogeneity of variance for the continuous model part implies that the degree of variation in an individual's observed responses about their fitted mean response across the repeated measures does not differ appreciably between individuals. For the between-subject covariance structure, an assumption of homogeneity of variance for both the binary and the continuous model parts implies that the degree of between-subject variation in the random effects (e.g., random intercepts) is constant at all levels of any between-subject covariates.

An important feature of two-part mixed-effects models is their flexibility in how these models can be formulated. As illustrated in this paper, the distributions of the binary and the continuous response can be assumed to follow a distribution of one's choosing. Here, we considered a logistic distribution for the binary response, and a normal, log-normal and gamma distribution for the continuous response. Other distributions for either model part are possible. This paper serves to highlight a particularly interesting aspect of this modeling framework, namely that it is possible to model the within-subject variation in scores across repeated measures, either by including covariates to study how or why individuals may differ in their variation in responses over time, or by including a random effect to document that there is heterogeneity of variation. It is also possible to model variation in the random effects that may be attributable to covariates. As illustrated using the empirical example presented in this paper, there was a notable difference between males and females in the variance of the random intercept for the continuous model part, suggesting a greater degree of individual differences between males than between females in time spent engaged in leisure activities. Not only did the extension of the model to allow for heterogeneity of variance in both the within-subject and between-subject components improve model fit, but interesting aspects of time spent in leisure activities were also brought to light.

Open Practices Statement: The data set used in the example is available for download at https://www.icpsr.umich.edu/icpsrweb/. This study was not preregistered.

Appendix

/ * SAS PROC NLMIXED may be used to fit a two-part mixed-effects model to repeated measures of semicontinuous data. The GENERAL model statement allows for the mixture of response distributions for the binary and continuous outcomes. For the continuous model part, the variance of the residuals may include within- and between-subject covariates and a random subject effect to allow for heterogeneity of variance. For both the binary and the continuous model parts, the variance of the random effects may include between-subject covariates to allow for heterogeneity variance. Different sets of covariates may be used in each of the three model parts.

Below is syntax for fitting a model to daily reports of time spent in leisure activities using data from the National Study of Daily Experiences (NSDE) (Almeida, 2007). Two new daily response variables were created prior to fitting the model:

1. binary response: $u = 1$ if $y > 0$ and $u = 0$ if $y = 0$; otherwise u = missing.
2. continuous response: $m = y$ if y > 0 and m = missing otherwise).

*/

* Syntax corresponds to Model B3 that includes a reduced covariate set;

```
proc nlmixed maxiter=10000 gconv=0 absconv=0 noad qpoints=20;

*user-provided starting values;
parms
a0    3.5065 a1a  0.6727 a1b  1.6357 a2   -0.593 a3   0.01566
b0    1.022  b1a  0.3575 b1b  -0.3163 b2  -0.1871 b3  0.01043 b4  -0.0054
b8    -0.0057
tau0  -1.2532 tau1a 0.1876 tau1b -0.5082 tau2 0.1413 tau3 -0.00649
alp0  1.0768
covab 0.3706
gam0  -1.3168 gam1     -0.2256
covwa -0.5054 covwb    -0.1743
varw  0.3869;

pi = arcos(-1);

*model variance of random intercepts for binary and continuous model parts;

vara = exp(alp0); varb = exp(gam0 + gam1*fem);

*LL1: Binary model part where ai is a random subject effect;

ueta = a0 + a1a*dayc + a1b*(propdayc) + a2*fem + a3*agec + ai;
expeta = exp(ueta); p=expeta/(1+expeta);

LL1 = log((1-p)**(1-u)) + log(p**(u));

*LL2: Continuous model part;

if leisureT > 0 then do;

*Daily time spent where bi is a random subject effect;
mu = b0 + b1a*dayc + b1b*(propdayc) + b2*fem + b3*agec + b4*fem*agec +
b8*dayc*agec + bi;

*Variance of daily time spent where wi is a random subject effect;
s2e = EXP(tau0 + tau1a*dayc + tau1b*(propdayc) + tau2*fem + tau3*agec +
wi);

*lognormal loglikelihood function;
LL2=(1/(sqrt(2*pi*s2e)*m))*exp(-(((log(m)-mu)**2)/(2*s2e))); end;

if leisureT = 0      then Loglik=LL1;
else if leisureT > 0 then Loglik=LL1+log(LL2);

*MVN distribution of random effects corresponding to model parts;
random ai bi wi ~ normal([0,0,0],[vara,
                                  covab,varb,
                                  covwa,covwb,varw]) subject=ID;
model leisureT ~ general(Loglik);
estimate 'conditional level 2 intercept variance, u' exp(alp0);
estimate 'conditional level 2 intercept variance, m' exp(gam0);
estimate 'conditional level 1 variance, s2e' exp(tau0);
run;
```

References

Almeida, D. M. (2007). *National Study of Daily Experiences (NSDE), March 1996-April 1997 [Computer file]*. ICPSR03725-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-12-13.

Carlin, J. B., Wolfe, R., Brown, C. H., & Gelman, A. (2001). A case study on the choice, interpretation, and checking of multilevel models for longitudinal, binary outcomes. *Biostatistics*, *2*(4), 397–416.

Cawley, J. & Liu, F. (2012). Maternal employment and childhood obesity: A search for mechanisms in time use data. *Economics & Human Biology*, *10*(4), 352-364.

Cragg, J.G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, *39*(5), 829-844.

Duan, N., Manning, W.G., Morris, C. N., & Newhouse, J. P. (1983). A comparison of alternative models for the demand for medical care. *Journal of Business & Economic Statistics*, *1*(2), 115-126.

Hedeker, D., & Nordgren, R. (2013). MIXREGLS: A Program for Mixed-Effects Location Scale Analysis. *Journal of statistical software*, *52*(12), 1–38.

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics*, *64*(2), 627-634.

Kiernan, K., Tao, J., & Gibbs, P. (2012). Tips and strategies for mixed modeling with SAS/STAT® Procedures. SAS Institute Inc., Cary, NC.

Liu, L., Cowen, M. E., Strawderman, R. L., & Shih, Y.-C. T. (2010). A flexible two-part random effects model for correlated medical costs. *Journal of Health Economics*, *29*(1), 110-123.

Mucciardi, M. (2013). Student time allocation and self-rated performance – Evidence from a sample survey in Sicily (Italy). *International Journal of Time Use Research*, *10*(1), 1-8.

Neuhaus, J. M., Kalbfleisch, J. D., & Hauck, W. W. (1991). A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *International Statistical Review*, *59*(1), 25-35.

Olsen, M. K., & Schafer, J. L. (2001). A two-part random effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association*, *96*(454), 730-745.

Sandberg, J. F., & Hofferth, S. L. (2001). Changes in children's time with parents: United States, 1981-1997. *Demography*, *38*(3), 423-436.

Stewart, J. (2013). Tobit or not Tobit? *Journal of Economic and Social Measurement*, *3*, 263-290.

Tooze, J. A., Grunwald, G. K., & Jones, R. H. (2002). Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research*, *11*(4), 341-355.

Xing, D., Huang, Y., Chen, H., Zhu, Y., Dagne, G. A., & Baldwin, J. (2017). Bayesian inference for two-part mixed-effects model using skew distributions, with application to longitudinal semicontinuous alcohol data. *Statistical Methods in Medical Research*, *26*(4), 1838-1853.