Article

Organizational Research Methods
2022, Vol. 25(4) 673–715

DOI: 10.1177/10944281211043733
journals.sagepub.com/home/orm
SAGE

Interaction Effects in Cross-Lagged Panel Models: SEM with Latent Interactions Applied to Work-Family Conflict, Job Satisfaction, and Gender

Ozlem Ozkok[1], Manuel J. Vaulont[2], Michael J. Zyphur[3], Zhen Zhang[4], Kristopher J. Preacher[5], Peter Koval[6], and Yixia Zheng[6]

Abstract

Researchers often combine longitudinal panel data analysis with tests of interactions (i.e., moderation). A popular example is the cross-lagged panel model (CLPM). However, interaction tests in CLPMs and related models require caution because stable (i.e., between-level, *B*) and dynamic (i.e., within-level, *W*) sources of variation are present in longitudinal data, which can conflate estimates of interaction effects. We address this by integrating literature on CLPMs, multilevel moderation, and latent interactions. Distinguishing stable *B* and dynamic *W* parts, we describe three types of interactions that are of interest to researchers: 1) purely dynamic or *WxW*; 2) cross-level or *BxW*; and 3) purely stable or *BxB*. We demonstrate estimating latent interaction effects in a CLPM using a Bayesian SEM in Mplus to apply relationships among work-family conflict and job satisfaction, using gender as a stable *B* variable. We support our approach via simulations, demonstrating that our proposed CLPM approach is superior to a traditional CLPMs that conflate *B* and *W* sources of variation. We describe higher-order nonlinearities as a possible extension, and we discuss limitations and future research directions.

[1]Department of Marketing, Rennes School of Business, Rennes, France
[2]Department of Management, HKUST Business School, The Hong Kong University of Science and Technology, Hong Kong
[3]Department of Management & Marketing, Faculty of Business & Economics, University of Melbourne, Australia
[4]Edwin L. Cox School of Business, Southern Methodist University, TX, USA
[5]Department of Psychology and Human Development, Vanderbilt University, TN, USA
[6]Melbourne School of Psychological Sciences, University of Melbourne, Australia

Corresponding author:
Ozlem Ozkok, Department of Marketing, Rennes School of Business, Rennes, France.
Email: ozlem.ozkok@rennes-sb.com

Time-series and panel data methods are used to study a variety of phenomena in organization science (Bliese et al., 2020). Of these, the cross-lagged panel model (CLPM) and its variants are a popular approach (e.g., Zyphur, Allison, et al., 2020; called panel vector autoregressions or panel VARs in econometrics; see Abrigo & Love, 2016; Love & Zicchino, 2006). These models estimate lagged relationships to allow inferences that are consistent with the fact that causality unfolds over time (Zyphur, Voelkle, et al., 2020), typically estimated as structural equation models (SEM) with fewer than 10 observations over time.

The CLPM can also be used to study conditions under which lagged effects vary by testing interactions.[1] Such tests often use product terms such as xz for predictors x and z (Cortina et al., 2021; Preacher et al., 2006, 2016), which can be calculated as the products of lagged predictors. For example, Eby et al. (2015) used the product of lagged predictors to test the interaction between (past) organizational support and (past) mentoring to predict (current) organizational citizenship behaviors. Related examples include: 1) models that account for measurement error with latent two-way (e.g., Dormann & Zapf, 1999) and three-way interactions (e.g., Lian et al., 2014); 2) combining interactions and nonlinear terms (e.g., Lin et al., 2017); 3) tests for sub-group differences in lagged effects (e.g., Cieslak et al., 2007; Houkes et al., 2003; Zablah et al., 2016); and 4) lag-as-moderator models (LAM) wherein the amount of time between measurements is treated as a moderator of lagged effects (Selig et al., 2012).

However, recent work suggests caution is needed when testing interactions in panel data, because different types of interactions may be conflated. Specifically, panel data contain stable 'between-level' (B) variation (e.g., personality; Hamaker et al., 2015) and dynamic 'within-level' (W) variation due to time-varying factors (e.g., emotions; Zyphur, Allison, et al., 2020). The multilevel modeling literature notes the importance of separating these sources of variation when testing interactions to allow precisely matching theory to hypotheses while avoiding the conflation of B and W effects (Preacher et al., 2016). However, when testing interactions in CLPMs, researchers have overlooked the implications of W and B sources of variation. In turn, this may have led to interaction estimates being difficult to interpret or resulted in flawed conclusions pertaining to significance testing.

In this paper we address testing interactions in CLPMs by integrating the literatures on CLPMs, multilevel moderation, and latent interactions. We start with the multilevel nature of longitudinal data and how panel data allow testing three types of interactions: (1) purely within-level or WxW interactions among dynamic factors; (2) cross-level or BxW interactions among stable B and dynamic W factors; and (3) purely between-level or BxB interactions among stable factors (Preacher et al., 2016). By extension, we also describe how to incorporate nonlinear and higher-order interactions of various kinds, including but not limited to within-level or W^2 nonlinearity; between-level or B^2 nonlinearity; and combinations of these with cross-level interactions, which we call B^2xW, BxW^2, or B^2xW^2 cases.

We then show how CLPMs may conflate such effects, including when estimating latent interactions (e.g., Dormann & Zapf, 1999; Lian et al., 2014) or attempting to control for stable B factors (e.g., Carpenter & Fredrickson, 2001). To address this issue, we describe a computationally tractable SEM framework for testing interactions in CLPMs with a Bayes estimator and latent B and W components (Asparouhov & Muthén, 2020). We illustrate our method for estimating and interpreting CLPM interaction effects using a dataset with time-varying measures of job satisfaction and work-family conflict as well as a stable measure of gender. We also run simulations showing that our approach is less biased than alternatives. We provide code for implementing our approach in Mplus, as this software is widely known in organization research and elsewhere (see Cortina et al., 2021; Heck & Thomas, 2015; Preacher et al., 2010, 2016). We conclude by discussing how our general approach can be used in other types of longitudinal panel data models including the general cross-lagged panel model (GCLM; Zyphur, Allison, et al., 2020), and latent curve or latent growth models (LGM; Chan, 1998). In sum, by capitalizing on the multilevel nature of longitudinal data we offer guidance for testing different types of multilevel interaction effects in a familiar CLPM framework estimated as an SEM.

A Multilevel Understanding of Cross-Lagged Panel Models

Longitudinal data are inherently complex, involving multiple sources of variation. When emotions are measured over time, for example, the resulting data exhibit both stability due to time-invariant factors and instability due to occasion-specific factors (Beal, 2015; Gabriel et al., 2019), which implies a multilevel structure. From this perspective, stable factors are higher-level or between-level (e.g., between-person or merely B), whereas dynamic factors are lower-level or within-level (e.g., within-person or merely W) phenomena (Zyphur, Allison, et al., 2020). In brief, B components reflect stable factors such as time-invariant individual differences (e.g., personality), and W components reflect temporary deviations around stable levels, due to occasion-specific factors (e.g., random events). Thus, to represent W components necessitates repeated measures of a variable with W variation and based on these a stable B component in the same content domain can be treated as a latent variable reflected in the repeated measures. It is notable that for variables without W variation such as stable gender or race, a single stable measure of the B variable will suffice and drawing an analogy to multilevel discussions of 'direct-consensus' measures, it would also be possible to collect 'direct-rating' measures for B terms (e.g., participants' reporting their 'usual' or 'overall' job satisfaction across a studied timespan).[2]

We illustrate the multilevel nature of longitudinal data using two variables x and y, which might reflect work-family conflict and job satisfaction measured for a person i at an occasion t, where $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. The multilevel logic of a stable B and dynamic W component can then be shown as follows (see Figure 1 for an SEM diagram):

$$x_{it} = x_{Bi} + x_{Wit} \tag{1}$$

$$y_{it} = y_{Bi} + y_{Wit} \tag{2}$$

In multilevel parlance, x_{it} and y_{it} would be Level-1 variables measured at the lowest level of analysis and thus containing latent B and W variation. For example, if x_{it} were work-family conflict measured over time, x_{Bi} would capture time-invariant stable factors impacting an individual's average level of work-family conflict (i.e., a "trait-like" part; Hamaker et al., 2015, p. 102), whereas x_{Wit} would capture time-varying dynamic factors associated with person i's experienced conflict at a specific time t. Indeed, we could use single-level SEM notation with a stable B term η_i and a dynamic W term as ε_{it} (see Curran, 2003). However, in keeping with our multilevel conceptualization, we use B and W terms here and emphasize to the reader that these are actually latent.

In this context, there are two primary ways to decompose a variable, namely manifest group-mean centering (i.e., an 'observed-means' approach) versus latent group-mean centering (see Lüdtke et al., 2008; Preacher et al., 2010). In the observed-means approach, researchers estimate the B part for each sampled entity using an average of the observed variables over time, with the deviations from this average representing the W part. In the latent centering case, a model-based approach is used to estimate the B and W variance components without attempting to estimate the underlying B and W scores directly. We compare these approaches later with real data and in simulations, but substantial prior research suggests that latent centering is superior to observed centering because the latter fails to account for uncertainty in the B and W components (Preacher et al., 2016). This is particularly important in longitudinal panel data models with lagged effects, wherein observed-means centering is known to cause 'Nickell' bias in lagged effect estimates (see Nickell, 1981; Zyphur, Allison, et al., 2020).

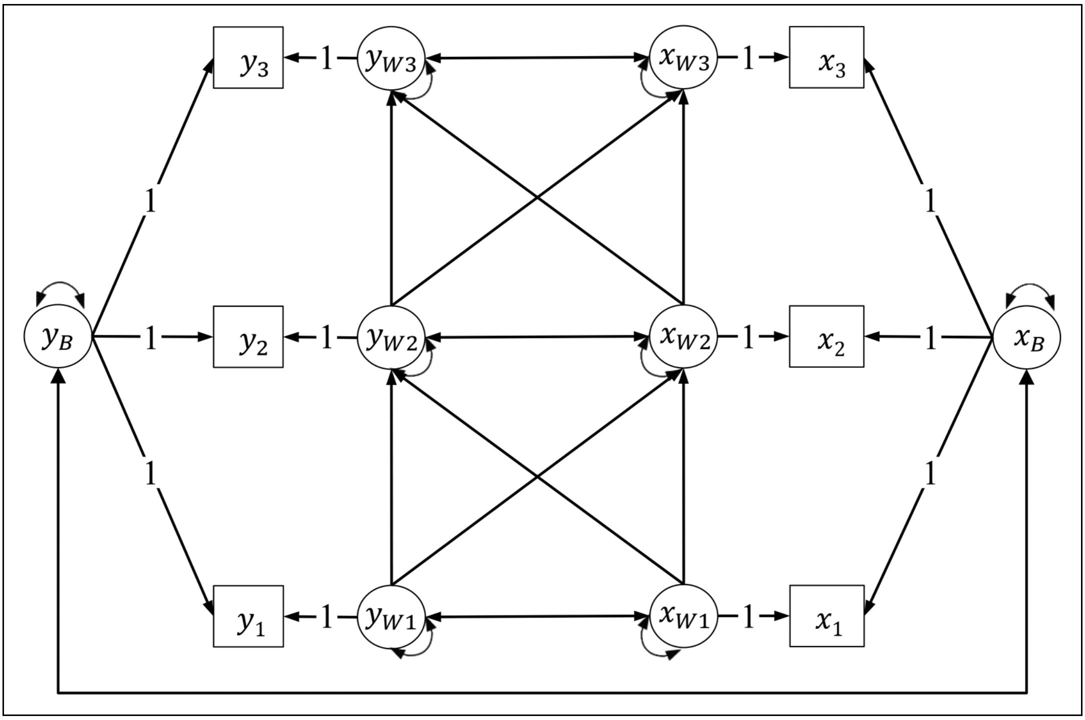

Figure 1. An SEM diagram distinguishing *W* and *B* components in a CLPM for *x* and *y* measured at three occasions.

To illustrate, we start by describing the dynamic *W* terms used for lagged effects in CLPMs, using notation from Zyphur, Allison, et al. (2020) where β coefficients have subscripts to indicate predictors and superscripts for outcomes:

$$x_{Wit} = \beta^x_{Wx} \cdot x_{Wit-1} + \beta^x_{Wy} \cdot y_{Wit-1} + u^x_{Wit} \tag{3}$$

$$y_{Wit} = \beta^y_{Wy} \cdot y_{Wit-1} + \beta^y_{Wx} \cdot x_{Wit-1} + u^y_{Wit} \tag{4}$$

In Eqs. 3 and 4, the autoregressive (AR) terms β^y_{Wy} and β^x_{Wx} indicate persistence in a process over time; cross-lagged (CL) terms β^x_{Wy} and β^y_{Wx} indicate unique predictions of one variable's future state by another variable's past state; and a time-specific 'impulse' u_{Wit} captures occasion-specific events or other factors that cause unpredictable (often conceptualized as random) increases or decreases in x and y at any occasion. As is common in the time-series literature, we use the term *impulse* instead of *error* (or an estimated *residual*) to emphasize the comprehensive nature of this term, encompassing not just variance unaccounted for by prior occasions but also potentially occasion-specific environmental changes (see Zyphur, Allison, et al., 2020). Using our example, if x were work-family conflict and y were job satisfaction, the AR term β^x_{Wx} would indicate persistence in work-family-conflict and the AR term β^y_{Wy} would represent persistence in job satisfaction over time. Further, for CL effects, β^x_{Wy} would indicate how past job satisfaction uniquely predicts future work-family-conflict and β^y_{Wx} would indicate the same for work-family-conflict predicting job satisfaction. Lastly, the u_{Wit} 'impulse' terms for x and y would represent occasion-specific random fluctuations in work-family conflict and job satisfaction that are not predicted by either job satisfaction or work-family conflict at the previous occasion. For example, the impulse term for job-satisfaction at a particular occasion may capture the effect of an individual unexpectedly receiving praise from their supervisor.

Equations 3 and 4 represent a familiar lagged-effects logic (for an associated logic of causality, although not the focus of our paper, see Granger, 1969; for a critical discussion see Zyphur, Allison, et al., 2020; Zyphur, Voelkle, et al., 2020). However, by substituting the lagged W terms from Eqs. 3 and 4 into the original multilevel decomposition in Eqs. 1 and 2, the model can be shown as a type of 'fixed-effects' CLPM (in econometrics terms) that implies stable B terms are held constant when examining W effects (see Figure 1):

$$x_{it} = x_{Bi} + \beta^{x}_{Wx} \cdot x_{Wit-1} + \beta^{x}_{Wy} \cdot y_{Wit-1} + u^{x}_{Wit} \tag{5}$$

$$y_{it} = y_{Bi} + \beta^{y}_{Wy} \cdot y_{Wit-1} + \beta^{y}_{Wx} \cdot x_{Wit-1} + u^{y}_{Wit} \tag{6}$$

In Eqs. 5 and 6, x_{Bi} and y_{Bi} represent the time-invariant latent means of x and y for person i. Notably, these are called 'random intercepts' in Hamaker et al. (2015) and 'unit effects' in Zyphur, Allison, et al. (2020), but we use the more common 'fixed effects' terminology from the econometrics literature to emphasize that these B terms are held constant when estimating W effects (see also Hamaker & Muthén, 2020). As we show later with an SEM, a fixed-effects specification that controls for stable B components is facilitated by allowing the latent B variables to freely covary (as in Hamaker et al., 2015; Zyphur, Allison, et al., 2020).[3]

To continue, a more general structural specification allows regression among B terms when causal effects among them can be theoretically justified, such as when studying individual differences or organizational climate and culture. For example, consider a model for y_{Bi} that includes the B part of x as a predictor and an observed B predictor variable z_{Bi} (where z_{Bi} is a Level-2 variable in terms of its measurement at the stable B level):

$$y_{Bi} = \beta^{y}_{Bx} \cdot x_{Bi} + \beta^{y}_{Bz} \cdot z_{Bi} + \zeta^{y}_{Bi} \tag{7}$$

with β^{y}_{Bx} and β^{y}_{Bz} implying, respectively, B effects of x_{Bi} and z_{Bi} on y_{Bi}, and ζ_{Bi} is a B residual or disturbance term associated with y_{Bi}. Here, the stable B part of y_{it} (e.g., the latent mean of job satisfaction) is predicted by the stable B part of x_{it} (e.g., the latent mean of work-family conflict) and a stable B predictor z_{Bi} (e.g., the observed variable gender). Therefore Equation (7) represents a regression among model-estimated latent averages for each person, which might be called a type of means-as-outcomes multilevel model (Raudenbush & Bryk, 2002).

By substituting the B model for y from Equation (7) into the larger model for y in Equation (6), we can show a full $B + W$ representation with both stable and dynamic parts as follows:

$$y_{it} = (\beta^{y}_{Bx} \cdot x_{Bi} + \beta^{y}_{Bz} \cdot z_{Bi} + \zeta^{y}_{Bi}) + (\beta^{y}_{Wy} \cdot y_{Wit-1} + \beta^{y}_{Wx} \cdot x_{Wit-1} + u^{y}_{Wit}) \tag{8}$$

In Equation (8), the first three terms represent the B model and the latter three represent the W model. When using the example of job satisfaction (y_{it}), work-family conflict (x_{it}) and gender (z_i), the first two predictors may be separated into dynamic W and stable B components, whereas gender is treated as time-invariant and therefore purely B. Such a decomposition into B and W parts is common in the multilevel literature (Preacher et al., 2010), and forms a basis for CLPMs when estimated in a multilevel framework, such as dynamic structural equation modeling (DSEM; Asparouhov et al., 2018; Zhou et al., 2021). A key point of this decomposition is that it implies fundamentally different *kinds* of effects: the W model contains lagged effects that are of primary interest to researchers who focus on AR dynamics and CL terms for causal inference, whereas the B model involves regression among stable factors. Our primary focus here is on the W model because the purpose of a CLPM is typically to estimate causal relationships by analyzing temporally ordered variables (Zyphur, Allison, et al., 2020). Yet, for the interested researcher, we formulate both B and W interaction models for completeness.

Latent Interactions in Cross-Lagged Models

Based on previous work on multilevel moderation (Preacher et al., 2016), we can also begin drawing conclusions about interactions in CLPMs. An interaction exists when the effect of a predictor x on an outcome y varies across the levels of another variable z. To test this, x and z are usually multiplied and their product xz is used as a predictor of y. As noted above, lagged-effects models that test interactions use this or conceptually similar approaches with lagged predictors. The idea behind using product terms in CLPM is that "extending this method to autoregressive models is straightforward, and no modifications to the mathematics are required to accommodate lagged values. That is, rather than (or in addition to) controlling for information available in prior lags, use lag variables as moderators" (Hayes, 2015, p. 16). This belief about the generality of methods for testing interactions has motivated the use of related interaction testing methods in longitudinal data (e.g., Selig et al., 2012; Zablah et al., 2016). However, some modifications to these approaches are needed to ensure that different sources of multilevel variation and effects are not conflated.

First, product terms among observed variables will confound different types of multilevel interactions when observed variables have B and W parts—consider the implications of expanding the product term $x_{it}y_{it} = (x_{Bi} + x_{Wit})(y_{Bi} + y_{Wit}) = x_{Bi}y_{Bi} + x_{Bi}y_{Wit} + y_{Bi}x_{Wit} + x_{Wit}y_{Wit}$; where there are now purely B, purely W, and two cross-level interactions implied in what would be a single predictor $x_{it}y_{it}$ (we return to this point in Equation. (24), below). Second, B and W components are not observed and thus are better conceptualized as latent variables (consider interactions among η_i and ε_{it} terms in SEM; see Ozkok et al., 2019).

In sum, multilevel research shows that testing interactions with multilevel data requires (1) decomposing observed variables into B and W parts, and then (2) specifying latent interactions separately with these parts (Preacher et al., 2016; Zyphur et al., 2018). Complementing this, recent work on autoregressive confirmatory factor analysis (AR-CFA) shows how such effects can be tested with latent interactions (see Ozkok et al., 2019). To adapt this research for CLPMs, we start by describing three types of interactions with panel data: purely within-level (WxW) interactions among dynamic factors; cross-level (BxW) interactions among stable B and dynamic W factors; and purely between-level (BxB) interactions among stable factors. We then elaborate on other cases.

Purely Within-Level or WxW Interactions. Many studies that test interactions using CLPMs seek to examine how CL effects vary across the levels of another dynamic variable. For example, Zablah et al. (2016) investigate whether the effect of past customer satisfaction on future job satisfaction is moderated by past customer engagement. In such cases, lagged predictors and moderators are often described as *time-varying* (alluding to W effects) and discussions of the effects imply that time-varying moderating variables are of interest. This focus on the dynamic W parts of variables also exists in studies of time-varying moderators of AR terms, such as investigations of how time-varying stress moderates the persistence (i.e., AR term) of emotional states (see Koval & Kuppens, 2012). Thus, the literature has various examples that analyze W interactions between lagged predictors (x_{Wit-1} and z_{Wit-1}) as these interact to predict an outcome (y_{Wit}). However, Koval and Kuppens (2012) findings illustrate that—although it is often overlooked—the lagged predictors in a typical CLPM may interact to predict their own future values, such as the interaction among x_{Wit-1} and y_{Wit-1} predicting x_{Wit} in Eq. 3 (along with predicting y_{Wit} in Equation (4)). Therefore, for simplicity and novelty reasons, as well as for illustration, we use this bivariate case to present WxW interactions here. It is easy to extend this to trivariate CLPM models with another time-varying variable z for testing xz, yz, or xyz interactions, as needed.

With an interest in forming latent interactions among W factors we first present purely W moderation effects. We show the simplest WxW case that does not require any external moderators and instead forms latent interactions among the lagged W components of x and y from Eqs. 3 and 4,

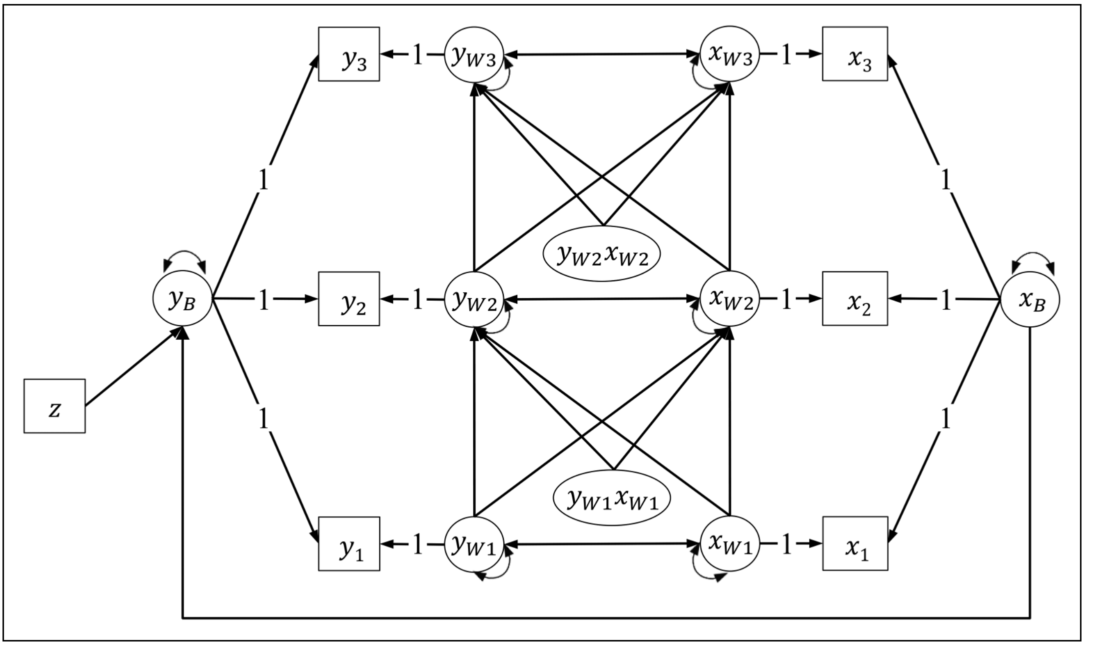

Figure 2a. CLPM for *x* and *y* at three occasions with latent *W*x*W* interactions.

which are always available in CLPMs (see Figure 2a):

$$x_{Wit} = \beta^x_{Wx} \cdot x_{Wit-1} + \beta^x_{Wy} \cdot y_{Wit-1} + \beta^x_{Wxy} \cdot x_{Wit-1} \cdot y_{Wit-1} + u^x_{Wit} \tag{9}$$

$$y_{Wit} = \beta^y_{Wy} \cdot y_{Wit-1} + \beta^y_{Wx} \cdot x_{Wit-1} + \beta^y_{Wxy} \cdot x_{Wit-1} \cdot y_{Wit-1} + u^y_{Wit} \tag{10}$$

In Equations (9) and (10), the purely *W* interaction effects among the dynamic parts of *x* and *y* are represented by β^x_{Wxy} and β^y_{Wxy} (we show both predictors in subscripts to signify an interaction). Uniquely, these effects imply that the lagged relationship for each predictor-outcome pair is conditional on the lagged value of the other predictor, which has not been explored in most CLPMs. To explain, β^x_{Wxy} and β^y_{Wxy} imply that (i) an outcome variable's past moderates the CL effect of a predictor, such as in the case of past job satisfaction impacting the effect of work-family conflict on future job satisfaction; or (ii) that a predictor's past moderates the persistence (or 'inertia') of an outcome variable, such as previous work-family conflict moderating the AR effect of job satisfaction. These effects may be quite common, as in Koval and Kuppens (2012) who found that inducing stress caused a decrease in emotional inertia (i.e., a smaller AR effect). In terms of our example, consider that the effect of past work-family conflict on future job satisfaction may be weaker when past job satisfaction is high. This would imply that high job satisfaction provides a buffer, which mitigates the effects of work-family conflict on future satisfaction. We examine this case later, but again note that additional predictors can easily be included for more typical two-way *W*x*W* interactions.

The point is that only the *W* part of each variable should form product terms to estimate a purely *W* interaction. This means that the stable *B* components are not only controlled, they also do not bias estimates of the *interaction* among *x* and *y* (a concern in the multilevel literature; see Preacher et al., 2016). Indeed, this would apply even when assessing *W*x*W* effects among variables that are relatively stable over time (e.g., personality), because these might still exhibit some *W* variation (e.g., across contexts;

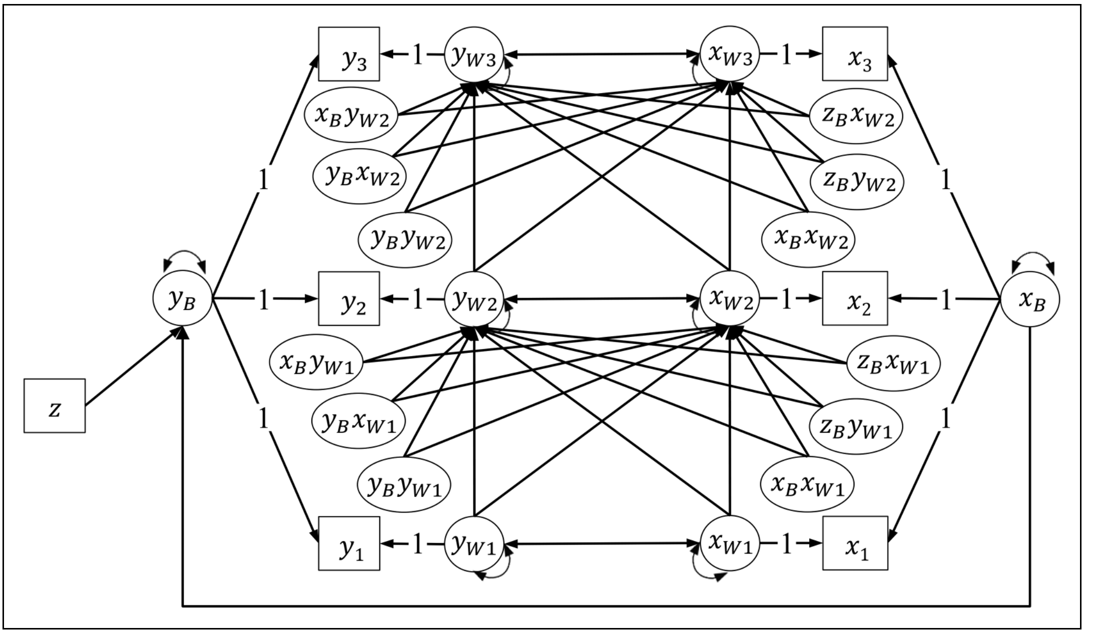

Figure 2b. CLPM for *x* and *y* at three occasions with latent *B*x*W* interactions.

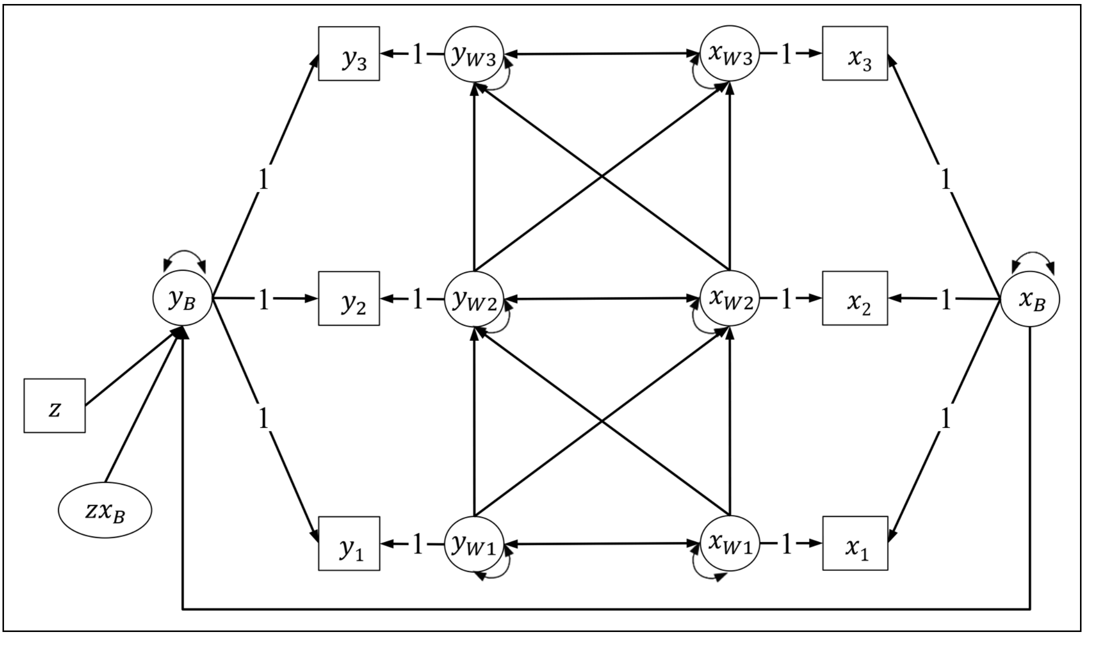

Figure 2c. CLPM for *x* and *y* at three occasions with latent *B*x*B* interactions.

Fleeson & Gallagher, 2009; Tett & Burnett, 2003) and tests of any dynamic *W* parts should not be biased by their stable *B* parts. Finally, our exploration of *W*x*W* effects shows how estimating these has always been possible in CLPMs because, by design, these models always include at least two lagged predictors. This realization opens up new opportunities to use CLPMs to test interactions—consider that *every* published CLPM can be re-estimated to test *W*x*W* interactions.

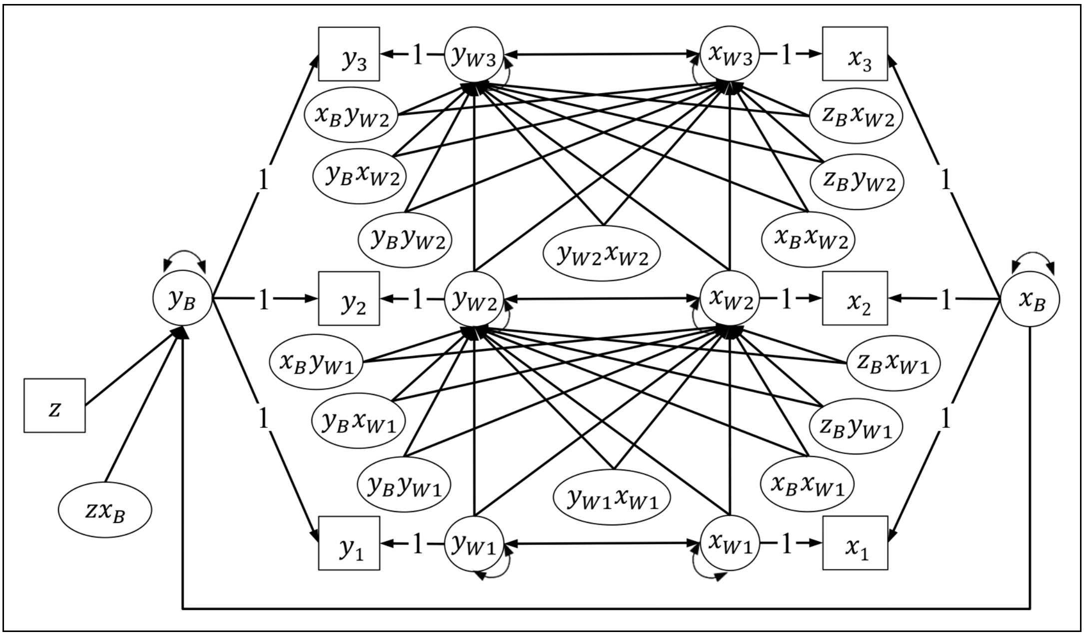

Figure 2d. CLPM for *x* and *y* at three occasions with latent *W*x*W*, *B*x*W*, and *B*x*B* interactions.

Cross-Level or BxW Interactions. Next, consider that time-invariant *B* variables, such as the stable components of personality, gender, or the environment may moderate lagged relationships in the *W* model. This is a familiar type of cross-level interaction that in CLPMs involves a *B* variable moderating a lagged *W* effect. An example from the emotions literature is that stable trait neuroticism (*B*) is associated with stronger persistence of negative emotions over time (i.e., slower reversion to the mean), such that the *W* AR terms for negative emotions are larger at higher levels of the stable *B* variable neuroticism (Koval et al., 2016; Suls et al., 1998). Similarly, the stable *B* part of negative emotions measured over time may moderate its own *W* AR effect (for insight see Figure 3 in Hamaker et al., 2018). Cross-level interactions may be present elsewhere, such as in Zablah et al. (2016), if the lagged *W* effect of customer satisfaction (x_{Wit-1}) on employee job satisfaction (y_{Wit}) were moderated by the stable *B* part of customer engagement (z_{Bi}), which these authors tested as a moderator.

For this, the *B* and *W* parts of observed variables must be decomposed or else the product terms will conflate *W* and *B* variation when estimating interactions (Preacher et al., 2016). To illustrate a solution to this problem, below we study how stable *B* gender (z_{Bi}) moderates the *W* cross-lagged effect of work-family conflict (x_{Wit-1}) on job satisfaction (y_{Wit}), an effect which was previously supported with models that likely conflated multilevel effects (e.g., Grandey et al., 2005). Furthermore, we show how the stable *B* parts of job satisfaction and work-family conflict can also moderate the *W* lagged effects among the same variables in a CLPM—again, this represents new opportunities as all published CLPMs can be re-analyzed to test for cross-level interactions among the modeled variables.

To formalize this familiar type of cross-level moderation we start with Eqs. 3 and 4, adding latent interactions among *B* and *W* terms (see Figure 2b). The equations are rather cumbersome, so for concision we show only the model for *y*. For clarity, we separate the purely *W* effects (including the *W*x*W* term) from the cross-level terms wherein a *B* variable moderates AR and CL effects in the *W* model. For the cross-level terms, Equation (11) describes a single *B* variable z moderating both *W* effects:

$$y_{Wit} = \beta^{y}_{Wy} \cdot y_{Wit-1} + \beta^{y}_{Wx} \cdot x_{Wit-1} + u^{y}_{Wit} + \ z_{Bi}(\beta^{y}_{BzWy} \cdot y_{Wit-1} + \beta^{y}_{BzWx} \cdot x_{Wit-1}) \tag{11}$$

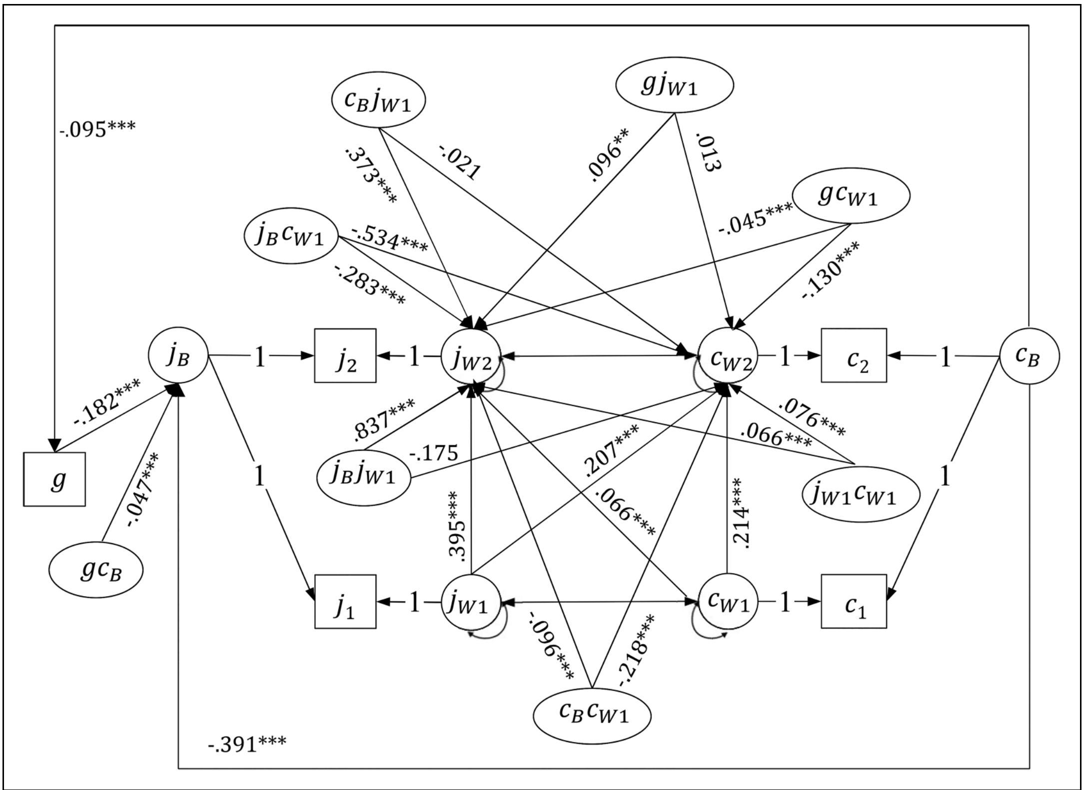

Figure 3. Graphic representation of the full model with estimates.

In Equation (11), we use subscripts B and W on the cross-level interaction effect β^y_{BzWx} to indicate the stable B and dynamic W predictors involved. In this equation, the B variable z moderates: (1) the AR term, with the level of persistence in y differing across values of the B variable (i.e., β^y_{BzWy}); and (2) a CL term, with the lagged effect of one W variable on the other differing across values of the B variable (i.e., β^y_{BzWx}). Therefore, Equation (11) shows the logic of a stable B variable measured at Level-2, such as gender (z_{Bi}), moderating the W effect of one dynamic variable, such as work-family conflict (x_{Wit-1}), on another dynamic variable, such as job satisfaction (y_{Wit}). However, because every CLPM involves at least two variables with repeated measurements, Equation. (11) can be expanded to include BxW interactions involving the latent B components (i.e., stable parts of x and y) in addition to an observed z_{Bi} with:

$$\begin{aligned} y_{Wit} = \beta^y_{Wy} \cdot y_{Wit-1} + \beta^y_{Wx} \cdot x_{Wit-1} + u^y_{Wit} \\ + \; y_{Bi}(\beta^y_{ByWy} \cdot y_{Wit-1} + \beta^y_{ByWx} \cdot x_{Wit-1}) \\ + \; x_{Bi}(\beta^y_{BxWy} \cdot y_{Wit-1} + \beta^y_{BxWx} \cdot x_{Wit-1}) \\ + \; z_{Bi}(\beta^y_{BzWy} \cdot y_{Wit-1} + \beta^y_{BzWx} \cdot x_{Wit-1}) \end{aligned} \tag{12}$$

Here, Equation (12) shows how the stable B parts of all variables can, in theory, be used to specify and test cross-level moderation (e.g., higher stable levels of work-family conflict x_{Bi} might sensitize people to the same variable's dynamic changes over time, thus increasing the CL effect β^y_{Wx} of time-varying work-family conflict x_{Wit-1} on job satisfaction y_{Wit} via the cross-level interaction β^y_{BxWx}). Of course, the WxW interaction effect in Equation (10) could also be moderated by a stable B variable, such as $x_{Bi} \cdot x_{Wit-1} \cdot y_{Wit-1}$ with a coefficient β^y_{BxWxy}, but this type of three-way BxWxW interaction is beyond the scope of our current paper.

Here, we focus on interpreting the effects in Equation (12) for readers who prefer either a traditional interaction/moderation depiction of coefficients or a multilevel random-slope depiction of cross-level interaction effects (as in Raudenbush & Bryk, 2002). We begin with a traditional depiction wherein 'main effects' and interactions are grouped as a compound coefficient on each primary lagged W predictor in the model, showing interactions separately for AR and CL terms as follows:

$$\begin{aligned} y_{Wit} = (\beta^y_{Wy} + \beta^y_{ByWy} \cdot y_{Bi} + \beta^y_{BxWy} \cdot x_{Bi} + \beta^y_{BzWy} \cdot z_{Bi}) y_{Wit-1} \\ + \ (\beta^y_{Wx} + \beta^y_{ByWx} \cdot y_{Bi} + \beta^y_{BxWx} \cdot x_{Bi} + \beta^y_{BzWx} \cdot z_{Bi}) x_{Wit-1} + \ u^y_{Wit} \end{aligned} \tag{13}$$

The model in Equation (13) is identical to the one in Equation (12), but the implied modification of the AR and CL terms is made clearer with a familiar moderated regression style of presentation, wherein moderation by B terms is shown as modifying the coefficients within parentheses on lagged predictors.

Alternatively, the same BxW terms can be shown in a multilevel random-slope style as follows, with arbitrary but conceptually useful equations for the AR and CL coefficients:

$$y_{Wit} = AR_i \cdot y_{Wit-1} + CL_i \cdot x_{Wit-1} + u^y_{Wit} \tag{14a}$$

$$AR_i = \beta^y_{Wy} + \beta^y_{ByWy} \cdot y_{Bi} + \beta^y_{BxWy} \cdot x_{Bi} + \beta^y_{BzWy} \cdot z_{Bi} \tag{14b}$$

$$CL_i = \beta^y_{Wx} + \beta^y_{ByWx} \cdot y_{Bi} + \beta^y_{BxWx} \cdot x_{Bi} + \beta^y_{BzWx} \cdot z_{Bi} \tag{14c}$$

In Equations (14*a*) to (14*c*), each slope is treated as varying (as in Raudenbush & Bryk, 2002; Yuan et al., 2014), but its variation is treated as fixed in the predictors rather than being given a random residual term. In other words, the 'random' part of a random slope is not present. This reduces model complexity and facilitates estimation particularly with small-T cases, like the CLPM wherein such a random variance would be estimated with low levels of precision. Instead, only the relevant fixed effects are estimated so that the AR and CL terms in the W model are treated as varying across the B variables, but *not* 'randomly' varying as in a random-slope model—note that we avoid the 'Level-1' and 'Level-2' nomenclature here to stay focused on the W and B effects (as in Preacher et al., 2010, 2016; Zhang et al., 2009).

This formulation clarifies the typical 'slopes-as-outcome' interpretation of multilevel interactions wherein the AR and CL coefficients β^y_{Wy} and β^y_{Wx} reflect the average effects across all N units of analysis when the interacting B predictors are mean-centered (i.e., y_{Bi}, x_{Bi}, z_{Bi}). However, when the interacting B predictors are not mean-centered, the AR and CL coefficients are treated like random-slope intercepts, such that an average β^y_{Wy} is equal to $\beta^y_{Wy} - \beta^y_{ByWy} \cdot M_{y_{Bi}} - \beta^y_{BxWy} \cdot M_{x_{Bi}} - \beta^y_{BzWy} \cdot M_{z_{Bi}}$; which shows why mean-centered predictors are often useful. Conveniently, in our SEM approach all latent B predictors are mean-centered by default because the mean structure will be accounted for by observed-variable intercepts a_t. When using observed-variable B predictors such as gender, these can either be manually mean-centered or the implied averages can be derived using the equation for M as noted (assuming that the overall averages of AR and CL terms are of interest).

As our approach thus far makes clear, not only is it possible to conceptualize W interactions in CLPMs, but it is also possible to formulate cross-level BxW interactions because panel data imply B components that may serve as moderators of W lagged effects. Seen this way, the bulk of the literature on multilevel moderation appears relevant to CLPMs including for interpreting effects. For example, when B moderators are mean-centered the 'main effects' β^y_{Wy} and β^y_{Wx} are the average

AR and CL effects across all N panels as in Equations (3) and (4). Yet, these may vary across the levels of stable B variables—here y_{Bi}, x_{Bi}, and z_{Bi} —which can be easily estimated by latent interactions among B and lagged W variables.

Purely Between-Level or BxB Interactions. Many researchers propose effects among variables that are conceptualized as having strong stable components, such as personality and job performance (see Barrick & Mount, 1991; Tett et al., 1991). In these cases, it is possible to test interactions among B variables. For example, Zablah et al. (2016) could have focused their theorizing on B effects, such that the stable B part of customer satisfaction predicts the stable B part of job satisfaction, and this effect is moderated by the stable B part of customer engagement. This model would describe a purely B form of moderation, which we refer to as a BxB interaction. This type of interaction can involve observed or latent variables at the B level. For example, gender could be treated as an observed B moderator (z_{Bi}) of the relationship between the stable B part of work-family conflict (x_{Bi}) and the stable B part of job satisfaction (y_{Bi}).

With this kind of approach, the B model for y in Equation (7) could be shown as involving a B moderator z_{Bi} to represent a BxB interaction as follows (see Figure 2c):

$$y_{Bi} = \beta^{y}_{Bx} \cdot x_{Bi} + \beta^{y}_{Bz} \cdot z_{Bi} + \beta^{y}_{Bxz} \cdot x_{Bi} \cdot z_{Bi} + \zeta^{y}_{Bi} \tag{15}$$

In Equation (15), we use the term β^{x}_{Bxz} to indicate a purely B interaction effect among the stable components of x and z. In our example, the stable B part of work-family conflict x_{Bi} may interact with gender z_{Bi} to predict the stable B part of job satisfaction y_{Bi}. Notably, even if z_{Bi} is observed (such as gender, self-reported without any measurement error assumed), a latent interaction needs to be specified if one component (in this case x_{Bi}) is latent.

Nonlinearity and Higher-Order Interactions. With these three types of interactions, more complex and interesting pictures of dynamic systems are available for theory testing (see Figure 2d for all interactions in one model). Yet, our approach also allows incorporating nonlinearity and higher-order interactions. Consider Lian et al.'s (2014) three-way interaction wherein abusive supervision, self-control capacity, and intention to quit at $t = 1$ were modeled as interacting to predict organizational deviance at $t = 2$. Their model focused on W relationships but did not account for stable B components. With our approach, building on Eqs. 9 and 10, it is possible to model pure WxWxW interactions of three variables by forming latent W interactions in a CLPM. A WxWxB interaction could also be modeled by using the B part of a variable and forming a three-way interaction with two W lagged predictors.

As another example, consider Lin et al. (2017) who studied a type of nonlinear moderation using squared predictors multiplied by a moderator. Specifically, perceived underemployment at $t = 1$ and task crafting at $t = 2$ had an inverted U-shaped relationship (modelled by including a squared term for underemployment), which was moderated by organizational identification at $t = 1$. Using our approach (assuming a sufficient number of measurement occasions) it would be possible to model the purely W nonlinear effect of underemployment on task crafting as a latent WxW interaction of underemployment with itself (what we call a W^2 model term), which could then be moderated by organizational identification's W or B part using a higher-order latent interaction—all while controlling for the stable B parts of the variables modeled.

Although the possibilities for testing nonlinear and higher-order interactions are essentially unlimited, we point out five potential cases that should be of primary interest and rather intuitive for most researchers. Notably, our examples here contain a variety of different types of interactions and nonlinear effects, not all of which may be of interest in any given application, but which we show for completeness. The first is a purely W quadratic effects or W^2 case that may include AR and CL terms using squared W lagged predictors as follows:

$$\begin{aligned} x_{Wit} = \beta^x_{Wx} \cdot x_{Wit-1} + \beta^x_{Wx^2} \cdot x^2_{Wit-1} \\ + \beta^x_{Wy} \cdot y_{Wit-1} + \beta^x_{Wy^2} \cdot y^2_{Wit-1} + u^x_{Wit} \end{aligned} \tag{16}$$

$$\begin{aligned} y_{Wit} = \beta^y_{Wy} \cdot y_{Wit-1} + \beta^y_{Wy^2} \cdot y^2_{Wit-1} \\ + \beta^y_{Wx} \cdot x_{Wit-1} + \beta^y_{Wx^2} \cdot x^2_{Wit-1} + u^y_{Wit} \end{aligned} \tag{17}$$

In Equations. (16) and (17), the coefficients $\beta^x_{Wx^2}$ and $\beta^y_{Wy^2}$ imply nonlinear AR effects, and coefficients $\beta^x_{Wy^2}$ and $\beta^y_{Wx^2}$ imply nonlinear CL effects. In the context of AR effects, this implies nonlinear persistence in a variable over time. For example, intense negative emotions may show higher inertia (i.e., greater persistence in the form of larger AR effects) because they are more difficult to down-regulate than mild negative emotions. Alternatively, consider the CL case where a W^2 term implies nonlinear lagged effects of a predictor. As we show below, these can be investigated by forming latent interactions among lagged W terms with themselves.

Next is the case of a quadratic B effect, which we refer to as B^2, that can be shown by elaborating on the B model for y from Equation (7) as follows:

$$y_{Bi} = \beta^y_{Bx} \cdot x_{Bi} + \beta^y_{Bx^2} \cdot x^2_{Bi} + \beta^y_{Bz} \cdot z_{Bi} + \zeta^y_{Bi} \tag{18}$$

In Equation (18), $\beta^y_{Bx^2}$ represents a quadratic effect of the stable B part of x. As an example, from Lin et al. (2017), if they had longitudinal panel data this could be used to test a nonlinear effect of the stable B part of perceived underemployment (x_{Bi}) on the stable part of task crafting (y_{Bi}). This relationship could also be moderated by the stable B part of organizational identification (z_{Bi}).

Now consider three unique cross-level moderation cases that combine the logic of the two previous W^2 and B^2 cases. We will refer to these as B^2 xW, Bx W^2, and B^2 xW^2 cases. In the first B^2 xW case, a quadratic B term moderates a linear W relationship. We start with the purely B^2 (quadratic) relationship in Eq. 18, adding a purely W linear relationship between y and x from Eq. 12 to form the final B^2 xW relationship as follows:

$$\begin{aligned} y_{Wit} = \beta^y_{Wy} \cdot y_{Wit-1} + \beta^y_{Wx} \cdot x_{Wit-1} + u^y_{Wit} \\ + y_{Bi}(\beta^y_{ByWy} \cdot y_{Wit-1} + \beta^y_{ByWx} \cdot x_{Wit-1}) \\ + x_{Bi}(\beta^y_{BxWy} \cdot y_{Wit-1} + \beta^y_{BxWx} \cdot x_{Wit-1}) \\ + x^2_{Bi}(\beta^y_{Bx^2Wy} \cdot y_{Wit-1} + \beta^y_{Bx^2Wx} \cdot x_{Wit-1}) \\ + z_{Bi}(\beta^y_{BzWy} \cdot y_{Wit-1} + \beta^y_{BzWx} \cdot x_{Wit-1}) \end{aligned} \tag{19}$$

In Equation (19), the nonlinear cross-level B^2 xW moderation is represented by $\beta^y_{Bx^2Wy}$ and $\beta^y_{Bx^2Wx}$. For example, if the stable part of work-family conflict (x_{Bi}) were posited to have a curvilinear moderating effect on the W effect of work-family conflict (x_{Wit-1}) on job satisfaction (y_{Wit}), this can be modeled in our approach using the term $\beta^y_{Bx^2Wx}$ in Equation (19).

In the second case of nonlinear Bx W^2 moderation, consider a nonlinear W^2 relationship that is moderated by the stable B parts of modeled variables. We start by writing purely W^2 relationship from Equation (17) and add cross-level interaction effects as follows:

$$\begin{aligned} y_{Wit} = &\beta^{y}_{Wy} \cdot y_{Wit-1} + \beta^{y}_{Wy^2} \cdot y^2_{Wit-1} + \beta^{y}_{Wx} \cdot x_{Wit-1} + \beta^{y}_{Wx^2} \cdot x^2_{Wit-1} + u^{y}_{Wit} \\ &+ y_{Bi}(\beta^{y}_{ByWy} \cdot y_{Wit-1} + \beta^{y}_{ByWy^2} \cdot y^2_{Wit-1} + \beta^{y}_{ByWx} \cdot x_{Wit-1} + \beta^{y}_{ByWx^2} \cdot x^2_{Wit-1}) \\ &+ x_{Bi}(\beta^{y}_{BxWy} \cdot y_{Wit-1} + \beta^{y}_{BxWy^2} \cdot y^2_{Wit-1} + \beta^{y}_{BxWx} \cdot x_{Wit-1} + \beta^{y}_{BxWx^2} \cdot x^2_{Wit-1}) \\ &+ z_{Bi}(\beta^{y}_{BzWy} \cdot y_{Wit-1} + \beta^{y}_{BzWy^2} \cdot y^2_{Wit-1} + \beta^{y}_{BzWx} \cdot x_{Wit-1} + \beta^{y}_{BzWx^2} \cdot x^2_{Wit-1}) \end{aligned} \tag{20}$$

In Equation (20), the cross-level B moderation effects of the nonlinear W^2 terms are $\beta^{y}_{ByWy^2}$, $\beta^{y}_{ByWx^2}$, $\beta^{y}_{BxWy^2}$, $\beta^{y}_{BxWx^2}$, $\beta^{y}_{BzWy^2}$, and $\beta^{y}_{BzWx^2}$. These terms imply that nonlinear W effects may vary across the levels of the B parts of the variables x, y, and z. For example, if a curvilinear effect of work-family conflict (x^2_{Wit-1}) on job satisfaction (y_{Wit}) were theorized to be moderated by gender (z_{Bi}), this could be tested as $\beta^{y}_{BzWx^2}$, which might show that family-work conflict has a stronger effect on job satisfaction at higher levels of family-work conflict, but only for women.

In the third case of nonlinear cross-level moderation, B^2 xW^2, we extend Equation (20) to the case of a purely W quadratic relationship between x and y moderated with a quadratic effect of the B part of a predictor variable z:

$$\begin{aligned} y_{Wit} = &\beta^{y}_{Wy} \cdot y_{Wit-1} + \beta^{y}_{Wy^2} \cdot y^2_{Wit-1} + \beta^{y}_{Wx} \cdot x_{Wit-1} + \beta^{y}_{Wx^2} \cdot x^2_{Wit-1} + u^{y}_{Wit} \\ &+ y_{Bi}(\beta^{y}_{ByWy} \cdot y_{Wit-1} + \beta^{y}_{ByWy^2} \cdot y^2_{Wit-1} + \beta^{y}_{ByWx} \cdot x_{Wit-1} + \beta^{y}_{ByWx^2} \cdot x^2_{Wit-1}) \\ &+ x_{Bi}(\beta^{y}_{BxWy} \cdot y_{Wit-1} + \beta^{y}_{BxWy^2} \cdot y^2_{Wit-1} + \beta^{y}_{BxWx} \cdot x_{Wit-1} + \beta^{y}_{BxWx^2} \cdot x^2_{Wit-1}) \\ &+ z_{Bi}(\beta^{y}_{BzWy} \cdot y_{Wit-1} + \beta^{y}_{BzWy^2} \cdot y^2_{Wit-1} + \beta^{y}_{BzWx} \cdot x_{Wit-1} + \beta^{y}_{BzWx^2} \cdot x^2_{Wit-1}) \\ &+ z^2_{Bi}(\beta^{y}_{Bz^2Wy} \cdot y_{Wit-1} + \beta^{y}_{Bz^2Wy^2} \cdot y^2_{Wit-1} + \beta^{y}_{Bz^2Wx} \cdot x_{Wit-1} + \beta^{y}_{Bz^2Wx^2} \cdot x^2_{Wit-1}) \end{aligned} \tag{21}$$

In Equation (21), the nonlinear cross-level moderation effects of interest are $\beta^{y}_{Bz^2Wy^2}$ and $\beta^{y}_{Bz^2Wx^2}$. This model allows for a quadratic cross-level moderating effect of z^2_{Bi} on the quadratic W relationship between work-family conflict x^2_{Wit-1} and job satisfaction y_{Wit}.

In all cases above, our approach decomposes the dynamic W and stable B parts of observed variables *before* forming product terms. This yields unbiased estimates of same-level and cross-level effects in CLPMs, while helping researchers focus on theory and hypothesis tests that are sensitive to interactions at different levels of analysis. As a beneficial byproduct, any heteroskedasticity that would otherwise exist in residuals can be accounted for by the interaction effects in the model (Raudenbush & Bryk, 2002). Before continuing, however, we would emphasize that simply because so many interactions and nonlinearities can be modeled does not imply that they should be. Indeed, a benefit of 'unconflating' B and W effects is that researchers need only model effects that are theoretically relevant.

Confounding in Tests of Interaction Effects

Given our presentation thus far, it is now straightforward to show how confounding occurs when variables are not disaggregated into B and W parts. For example, when the product of x and y is included in a typical CLPM, the net result can be shown as follows for the y variable (again, temporarily ignoring intercepts or an 'occasion effect' α_t):

$$y_{it} = \beta^{y}_{y} y_{it-1} + \beta^{y}_{x} \cdot x_{it-1} + \beta^{y}_{xy} \cdot x_{it-1} \cdot y_{it-1} + e^{y}_{it} \tag{22}$$

where the product of x and y is lagged and the coefficient β_{xy}^{y} is meant to indicate the interaction effect or degree of moderation for either the AR or CL term (as is well known, either x or y can be understood as the moderator of the other's effect in such a case).

However, the multilevel decomposition in Equations (1) and (2) suggests that this formulation has various potential shortcomings, which we help illustrate by using a compound residual term e_{it}^{y} rather than the separate B and W parts ζ_{Bi}^{y} and u_{Wit}^{y}, respectively. To further show the problems here, we can substitute the multilevel B and W counterparts in Eq. 22 as follows:

$$\begin{aligned} y_{it} &= \beta_{y}^{y}(y_{Bi} + y_{Wit-1}) + \beta_{x}^{y}(x_{Bi} + x_{Wit-1}) \\ &\quad + \beta_{xy}^{y}(x_{Bi} + x_{Wit-1})(y_{Bi} + y_{Wit-1}) + e_{it}^{y} \end{aligned} \tag{23}$$

In Equation (23), the first two terms $\beta_{y}^{y}(y_{Bi} + y_{Wit-1})$ and $\beta_{x}^{y}(x_{Bi} + x_{Wit-1})$ exemplify the well-known multilevel problem of conflating potential B and W effects (for general insight see Preacher et al., 2010; Zhang et al., 2009; for the CLPM case specifically see Hamaker et al., 2015; Zyphur, Allison, et al., 2020; Zyphur, Voelkle, et al., 2020).

By not decomposing y and x into their level-specific components, interaction terms that conflate B and W parts can be shown by expanding the β_{xy}^{y} product term as follows:

$$\begin{aligned} y_{it} = {} & \beta_{y}^{y}(y_{Bi} + y_{Wit-1}) + \beta_{x}^{y}(x_{Bi} + x_{Wit-1}) \\ & + \beta_{xy}^{y} \cdot x_{Bi} \cdot y_{Bi} + \beta_{xy}^{y} \cdot x_{Bi} \cdot y_{Wit-1} + \beta_{xy}^{y} \cdot y_{Bi} \cdot x_{Wit-1} \\ & + \beta_{xy}^{y} \cdot x_{Wit-1} \cdot y_{Wit-1} + e_{it}^{y} \end{aligned} \tag{24}$$

As we already showed above, Eq. 24 makes it apparent that the single coefficient β_{xy}^{y} actually represents four separate interactions: a purely between-level or BxB term $(x_{Bi} \cdot y_{Bi})$, two cross-level or BxW terms $(x_{Bi} \cdot y_{Wit-1})$ and $(y_{Bi} \cdot x_{Wit-1})$, and a purely within-level or WxW term $(x_{Wit-1} \cdot y_{Wit-1})$. By extension, as implied by Eq. 24, adding nonlinearity and/or three-way interactions would exacerbate the interpretational difficulties and degrees of confounding. In what follows, we illustrate our proposed latent interaction CLPM approach by applying it to a longitudinal panel dataset. Following this, we report the results of simulations that lend support for our proposed approach with latent interactions by comparison with observed-means centering and an uncentered approach (i.e., raw variables).

Method

Participants

Data were taken from the Swiss Household Panel (SHP) survey, which is a nationally representative random sample measured annually, beginning in 1999 with around 5,000 Swiss households and adding two additional samples in 2004 and 2013 (see Voorpostel et al., 2018). Individual household members are surveyed via telephone on a broad range of topics including health, employment, and leisure activities. The SHP has been used extensively in fields such as public health (e.g., Knecht et al., 2011), sociology (e.g., Oesch & Lipps, 2013), and economics (e.g., Frei & Sousa-Poza, 2012).

In the present study, we used data collected in the years 2004 through 2007. We chose 2004 as the onset of our study because this year marked the first that all study variables were measured annually, and the first year that included a second SHP sample, which increased our sample size. Although the SHP's sample included around 10,000 working people in 1999, due to attrition this number decreased to 6,000 in 2004. The addition of a second sample in 2004 increased the number of working participants to over 13,000. Furthermore, the waves 2004 to 2007 consistently demonstrated response rates of more than 70%.

We included all working individuals that responded to at least one of the four waves surveyed between 2004 and 2007. This brought our sample size to 7,748 individuals. The participants in our sample (51.6% women, 48.4% men) were on average 39.28 ($SD = 14.50$) years old and had 12.98 ($SD = 3.20$) years of education.

Measures

Descriptive statistics are reported in Table 1. We used measures of job satisfaction and work-family conflict. Past research shows that behavior-based types of work-family conflict (i.e., behavioral incompatibilities of different roles) have the strongest relationship to job satisfaction, particularly composite-based satisfaction (averaged across facets) rather than global measures (Bruck et al., 2002).

Job Satisfaction. Participants indicated the extent to which they were satisfied with their job by responding to six items created for the SHP (*0 = not at all satisfied* to *10 = completely satisfied*). The lead-in question was "Can you indicate your degree of satisfaction for each of the following points?" The six items were "your job in general", "the level of interest in your tasks", "the amount of your work", "the income you get from your job", "your work conditions", and "the atmosphere between you and your colleagues." For simplicity, we computed observed-variable mean scores across these six job satisfaction items. However, alternative higher-order latent variable models would also be possible given the general flexibility of SEMs (see Mulder & Hamaker, 2021). The alpha coefficients for the four waves were 0.78, 0.78, 0.79, and 0.79, respectively.

Work-Family Conflict. Participants rated the extent to which their work negatively impacted their family life by a single-item measure: "How strongly does your work interfere with your private activities and family obligations more than you would want this to be?" (0 = *not at all* to 10 = *very strongly*).

Hours Worked. Participants reported the number of hours worked per week, which we included as an observed variable in our model with the same W and B decomposition to control for hours worked per week when interpreting all effects. Before analysis we multiplied this variable by 4 to improve convergence when setting observed variances to small values as noted in the next section.

Gender. In every wave, participants indicated their gender using a dummy variable (*0 = man*, *1 = woman*). Responses across all participants were invariant across all years, allowing this to serve as a purely B observed variable in our CLPM as described next.

Specification and Estimation

We first offer a baseline model decomposing W and B parts of observed variables (as shown in Equations. (1) and (2)). This model includes lagged effects for the W parts (as in Equations (5)and (6)); a B model with stable B job satisfaction predicted by stable B work-family conflict, gender, and hours worked (as in Eq. 7); stable B work-family conflict predicted by gender and hours worked; and gender predicting B hours worked. Details on this SEM specification can be found in Hamaker et al. (2015), including the addition of a time-varying intercept α_t for each observed variable by default (typical in SEM). To elaborate on this, we estimated four other models that included interaction effects: within-level (WxW; as in Equations (9) and (10)); cross-level (BxW; as in Equations (12) to (14)); between-level (BxB; as in Equation (15)); and all simultaneously (WxW, BxB, and BxW). For comparison, we also estimated a traditional CLPM without the B and W

Table 1. Descriptive Statistics and Correlations of Study Variables.

Variable		Mean	σ^2	1	2	3	4	5	6	7	8	9	10	11	12
1	J9	7.825	1.533												
2	J8	7.826	1.503	.634											
3	J7	7.908	1.451	.568	.630										
4	J6	7.966	1.507	.528	.559	.622									
5	C9	3.802	7.620	−.288	−.236	−.201	−.189								
6	C8	3.827	7.476	−.191	−.231	−.183	−.177	.546							
7	C7	3.721	7.766	−.197	−.198	−.245	−.202	.489	.539						
8	C6	3.652	8.218	−.183	−.183	−.176	−.219	.426	.462	.476					
9	H9	13.731	6.918	−.059	−.065	−.078	−.043	.326	.294	.282	.245				
10	H8	13.735	7.025	−.061	−.062	−.066	−.048	.281	.323	.317	.260	.834			
11	H7	13.769	6.955	−.056	−.058	−.068	−.039	.267	.298	.347	.280	.781	.864		
12	H6	13.804	6.723	−.071	−.077	−.088	−.062	.260	.280	.311	.316	.723	.808	.852	
13	G	0.506	0.250	.020	.016	.035	.042	−.111	−.111	−.139	−.101	−.349	−.364	−.364	−.368

Note: $N = 7{,}748$ from the Swiss Household Panel (SHP) with variables named in accordance with the SHP waves; J9-J6 = job satisfaction measured in 2007–2004; C9-C6 = work-family conflict measured in 2007–2004; H9-H6 = hours worked measured in 2007–2004; G = gender coded 0 for men and 1 for women; σ^2 = estimated population variance.

decomposition so that product terms are formed with observed-variable lags, which we term the 'conflated model' below. All Mplus files and a summary of the results can be found in our Online Appendices.

To estimate our models, we considered using latent moderated structural equations or LMS (e.g., Preacher et al., 2016), but this is computationally infeasible in high dimensions (Zyphur et al., 2018). This problem is addressed with the Bayes estimator in Mplus, which provides optimal estimates and faster compute times using parallel processing (Asparouhov & Muthén, 2020). To do this requires first checking compute times for 400 iterations with different numbers of physical (rather than logical or 'virtual') processing cores, which we did for up to 48 cores using Amazon Web Services (AWS) 'c5.24xlarge' EC2 instances (with 48 physical and 96 virtual cores). We found that, in a Windows environment, 4 or 8 cores were optimal for models with fewer interactions, but with many interactions 16 cores was best. However, in the same AWS platform, we found that in a Linux environment estimation scaled much better with higher core counts. For example, estimating the *W*x*W* model below with 60,000 total iterations (as we note below, thinning every other iteration and requesting 30,000 iterations), we observed the following estimation times with our comparatively large model: 2 cores = 25.92 h; 4 cores = 13.33 h; 8 cores = 7 h; 16 cores = 3.58 h; 32 cores = 1.83 h; 48 cores = 1.36 h. As such, we recommend using a Linux environment rather than Windows, particularly given the Linux instance is roughly half price on AWS (~$4.50 hly vs. Windows at ~$9.00 hly).

For all models we checked convergence using posterior scale reduction (PSR) values < 1.05, at least doubling iterations after this to ensure stable convergence (Asparouhov & Muthén, 2010; Zyphur & Oswald, 2015). For this we were conservative and estimated more than will typically be required for other researchers using similar data, for example thinning every other iteration with 30,000 iterations set, for a total of 60,000 in the *W*x*W* case. To aid in convergence, it is notable that decomposing the *W* and *B* parts of observed variables requires accounting for all variance in an observed variable, which implies zero residual variances. This causes slow convergence with Bayes estimators, so we set residual variances to roughly 1–3% of the total observed variation to improve convergence (Asparouhov & Muthén, 2020), while allowing the decomposed *W* and *B* variances to be freely estimated.

Results for Real-Data Example

All model results are shown in Table 2 in a side-by-side format to allow easy comparisons of parameter estimates across models. We present results for each model separately, starting with the baseline model to help draw comparisons with those that follow, including interaction effects and finally the conflated model that does not decompose *B* and *W* parts with observed-variable products for interactions. Standardized effects can be computed using formulas presented in Asparouhov and Muthén (2020), but for brevity we present raw coefficients, as is often done in multilevel literature testing interactions. For brevity, Figure 3 contains results only for the full model with all effects except for the paths involving the control variable (hours worked) and with only two occasions of measurement shown to aid readability. Also, all reported *p*-values are two-tailed and calculated using Bayesian posteriors *SD*s (rather than their frequentist counterpart *SE*s).

Baseline Model. For this model, job satisfaction (J), work-family conflict (C), and hours worked (H) are decomposed into stable *B* and dynamic *W* parts with latent variables, whereas gender (G) is an observed *B* variable (to benefit the reader these variable labels are also in our online Mplus files). We use H as a main-effect control (i.e., no interactions) but decompose it into *W* and *B* parts to properly account for separate *W* and *B* associations with J and C. Notably, typical fit statistics are not available in the presence of latent interactions, but baseline models without interactions have fit equivalent to interaction models (Asparouhov & Muthén, 2020). The baseline model showed adequate fit, but with

Table 2. Summary of Model Estimates.

Variables[(1)]	Baseline Model	*WxW* Model	*BxB* Model	*BxW* Model	Full Model	Conflated Model Variables	Conflated Model
	Estimates	Estimates	Estimates	Estimates	Estimates		Estimates
***W* Effects**						***B* + *W* Effects**	
AR effect C_W	0.086***	0.087***	0.088***	0.206***	0.214***	AR effect for C	0.465***
AR effect J_W	0.203***	0.197***	0.204***	0.395***	0.395***	AR effect for J	0.639***
AR effect H_w	0.493***	0.507***	0.498***	0.557***	0.558***	AR effect for H	0.816***
CL effect $C_W \rightarrow J_W$	0.006	0.005	0.006	0.071***	0.066***	CL effect C→J	0.006
CL effect $J_W \rightarrow C_W$	−0.079	−0.064	−0.075	0.229***	0.207***	CL effect J→C	−0.202***
CL effect $H_W \rightarrow C_W$	0.140***	0.156***	0.146***	0.236***	0.239***	CL effect H→C	0.136***
CL effect $H_W \rightarrow J_W$	−0.005	−0.002	−0.005	0.027**	0.023**	CL effect H→J	−0.007
CL effect $C_W \rightarrow H_W$	0.020	0.025*	0.021*	0.063***	0.065***	CL effect C→H	0.020***
CL effect $J_W \rightarrow H_W$	0.046	0.053	0.046	0.121***	0.115***	CL effect J→H	0.005
***WxW* Interactions**						**Interactions**	
$C_W{*}J_W \rightarrow C_W$		0.044**			0.076***	C*J→C	−0.001
$C_W{*}J_W \rightarrow J_W$		0.030**			0.066***	C*J→J	−0.003
***B* Effects**						**Effects of Gender**	
$C_B \rightarrow J_B$	−0.196***	−0.200***	−0.176***	−0.392***	−0.391***		
$G_B \rightarrow J_B$	0.000	−0.002	−0.011	−0.174***	−0.182***	$G_B \rightarrow J$	−0.010
$H_B \rightarrow J_B$	0.027**	0.026**	0.029**	−0.003	0.002		
$G_B \rightarrow C_B$	0.077	0.072	0.074	−0.080	−0.095	$G_B \rightarrow C$	0.034*
$H_B \rightarrow C_B$	0.357***	0.354***	0.356***	0.269***	0.260***		
$G_B \rightarrow H_B$	−2.000***	−1.999***	−2.106***	−1.987***	−1.987***	$G_B \rightarrow H$	−0.407***
***BxB* Interaction**							
$C_B{*}G_B \rightarrow J_B$			−0.045		−0.047***		
***BxW* Interactions**							
$C_B{*}C_W \rightarrow C_W$				−0.179***	−0.218***		
$J_B{*}C_W \rightarrow C_W$				−0.451***	−0.534***		
$G_B{*}C_W \rightarrow C_W$				−0.132***	−0.130***		
$C_B{*}J_W \rightarrow C_W$				0.121**	−0.021		
$J_B{*}J_W \rightarrow C_W$				0.098	−0.175		
$G_B{*}J_W \rightarrow C_W$				0.072	0.013		
$C_B{*}C_W \rightarrow J_W$				−0.077***	−0.096***		

(continued)

Table 2. (continued)

Variables[1]	Baseline Model	*W*x*W* Model	*B*x*B* Model	*B*x*W* Model	Full Model	Conflated Model Variables	Conflated Model
$J_B{*}C_W \rightarrow J_W$				−0.239***	−0.283***		
$G_B{*}C_W \rightarrow J_W$				−0.041**	−0.045***		
$C_B{*}J_W \rightarrow J_W$				0.365***	0.373***		
$J_B{*}J_W \rightarrow J_W$				0.837***	0.837***		
$G_B{*}J_W \rightarrow J_W$				0.110**	0.096**		

Note: All estimates are unstandardized.
[1] *W* = Within; *B* = Between; J = Job satisfaction; C = Work-family conflict; G = Gender; H = Hours worked; Arrows represent direction.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Source: Swiss Household Panel (SHP).

typical problems of absolute misfit using χ^2 metrics coupled with very good fit for other indices as follows: χ^2 difference for observed versus model-generated data shows a 95% CI ranging from 21.296 to 104.514 (excluding zero, indicating poor fit) with the number of parameters $pD = 53.855$; posterior predictive p-value or $PPP = .001$ (smaller than .05, indicating poor fit); but Bayes-equivalent RMSEA = .006 (smaller than .08, indicating good fit); and both CFI and TLI = .999 (larger than .95, indicating good fit); and DIC = 218442.103.

The following CLPM was implemented in Mplus (see the online files Baseline CLPM.inp and .out) to compute baseline W and B parts and slope coefficients as follows:

$$\begin{aligned} J_{it} = \alpha_t^J + \beta_{B(C)}^J \cdot C_{Bi} + \beta_{B(G)}^J \cdot G_{Bi} + \beta_{B(H)}^J \cdot H_{Bi} + \beta_{W(J)}^J \cdot J_{Wit-1} + \beta_{W(C)}^J \cdot C_{Wit-1} \\ + \beta_{W(H)}^J \cdot H_{Wit-1} + \zeta_{Bi}^J + u_{Wit}^J \end{aligned} \tag{25}$$

$$\begin{aligned} C_{it} = \alpha_t^C + \beta_{B(G)}^C \cdot G_{Bi} + \beta_{B(H)}^C \cdot H_{Bi} + \beta_{W(C)}^C \cdot C_{Wit-1} + \beta_{W(J)}^C \cdot J_{Wit-1} \\ + \beta_{W(H)}^C \cdot H_{Wit-1} + \zeta_{Bi}^C + u_{Wit}^C \end{aligned} \tag{26}$$

$$\begin{aligned} H_{it} = \alpha_t^H + \beta_{B(G)}^H \cdot G_{Bi} + \beta_{W(H)}^H \cdot H_{Wit-1} + \beta_{W(J)}^H \cdot J_{Wit-1} + \beta_{W(C)}^H \cdot C_{Wit-1} \\ + \zeta_{Bi}^H + u_{Wit}^H \end{aligned} \tag{27}$$

The AR effects for work-family conflict, job satisfaction, and hours worked were significant with $\beta_{W(J)}^J = .203$ ($p < .001$), $\beta_{W(C)}^C = .086$ ($p < .001$), $\beta_{W(H)}^H = .493$ ($p < .001$); but most CL effects were not: $\beta_{W(C)}^J = .006$ ($p = .354$), $\beta_{W(J)}^C = -.079$ ($p = .076$), $\beta_{W(H)}^J = -.005$ ($p = .73$), $\beta_{W(C)}^H = .02$ ($p = .092$), $\beta_{W(J)}^H = .046$ ($p = .19$), except for one significant CL effect of hours worked on work-family conflict $\beta_{W(H)}^C = .14$, ($p < .001$). The AR results imply that roughly 20% of a dynamic W perturbation to job satisfaction and 49% for hours worked persists from one occasion to the next, but only about 8% for work-family conflict. For CL effects, typical significance tests support only a direct effect of hours worked on work-family conflict.

Estimates for B terms supported a negative effect of average work-family conflict on average job satisfaction as $\beta_{B(C)}^J = -.196$ ($p < .001$), but not for gender $\beta_{B(G)}^J < 0.001$ ($p = .099$). Gender was negatively associated with hours worked with $\beta_{B(G)}^H = -2.0$ ($p < .001$) but not average work-family conflict $\beta_{B(G)}^C = .077$ ($p = .230$). This indicates that women (Gender = 1) seem to work fewer hours but do not experience more work-family conflict. Average hours worked predicted average work-family conflict $\beta_{B(H)}^C = .357$ ($p < .001$) and average job satisfaction $\beta_{B(H)}^J = .027$ ($p = .01$). Taken together these B coefficients suggest that as a person works more, they experience more work-family conflict but also more job satisfaction, on average. Gender did not predict average job satisfaction or work-family conflict. If the B coefficients were causal, this would suggest that greater work-family conflict causes reduced job satisfaction, women work fewer hours, and hours worked increases work-family conflict, but working more also ironically increases job satisfaction.

In sum, the baseline model effects are important because researchers who use CLPMs are interested in knowing what dynamic effects are present in the W parts of their models. The B model is also relevant for those who are comfortable with their model specification and any inferences that follow from it. However, given the possibility of estimating different types of latent interactions, this typical CLPM may omit important moderation effects.

Within-Level Interaction (*W*x*W*). Our first interaction model includes W latent interactions of work-family conflict and job satisfaction as in Eqs. 9–10. This model allows studying whether AR and CL terms among the W dynamic variables depend on levels of the lagged variables. As we would expect, AR and CL estimates were very similar to the baseline model—because the

W latent variables are mean-centered by default, the baseline model will reflect the average AR and CL terms in a sample as will any WxW interaction model. Specifically, AR terms were $\beta^J_{W(J)} = .197$ ($p < .001$), $\beta^C_{W(C)} = .087$ ($p < .001$), and $\beta^H_{W(H)} = .507$ ($p < .001$), whereas CL terms were $\beta^J_{W(C)} = .005$ ($p = .464$), $\beta^C_{W(J)} = -.064$ ($p = .12$), $\beta^C_{W(H)} = .156$, ($p < .001$), $\beta^J_{W(H)} = -.002$ ($p = .894$), $\beta^H_{W(C)} = .025$ ($p = .042$), and $\beta^H_{W(J)} = .053$ ($p = .08$). Compared to the baseline model, including the WxW interaction terms reduced some uncertainty by making the CL effect of work-family conflict on hours worked significant, suggesting that a dynamic increase to work-family conflict may cause a slight increase in hours worked. Furthermore, the B estimates were also very similar to those obtained under the baseline model—which is expected because the B and W model parts are orthogonal and therefore adding WxW interactions should not markedly affect B estimates—with $\beta^J_{B(C)} = -.20$ ($p < .001$), $\beta^J_{B(G)} = -.002$ ($p = .954$), $\beta^J_{B(H)} = .026$ ($p = .008$), $\beta^C_{B(G)} = .072$ ($p = .246$), $\beta^C_{B(H)} = .354$ ($p < .001$), and $\beta^H_{B(G)} = -1.999$ ($p < .001$).

Moving on to the WxW interactions, we found statistically significant effects. First, when predicting work-family conflict, we found a positive interaction effect $\beta^C_{W(JC)} = .044$ ($p = .016$) such that greater levels of lagged work-family conflict or job satisfaction increased the effect of the other on future work-family conflict. Figure 4 offers an example of how to visualize such interactions. This means that as work-family conflict increases, the CL effect of job satisfaction on conflict ($\beta^C_{W(J)}$) increases, and as job satisfaction increases, the AR effect of work-family conflict ($\beta^C_{W(C)}$) increases. Similarly, when predicting job satisfaction, we found a positive interaction effect $\beta^J_{W(JC)} = .03$ ($p = .02$). This is plotted in Figure 5, which shows that greater levels of lagged job satisfaction increase the CL effect of conflict ($\beta^J_{W(C)}$), and greater levels of lagged conflict increase the AR effect of job satisfaction ($\beta^J_{W(J)}$). Although either variable could be theoretically seen as a moderator, our point is that these interaction effects are of potential interest to researchers who can always estimate them in CLPMs. Our results here show that the dynamic W parts of work-family conflict and job satisfaction appear to have implications for the other's dynamic effects.

Between-Level Interactions (*B*x*B*). Here we include B interactions as in Eq. 15. As Table 2 shows, AR and CL effects were again essentially unchanged compared to the baseline because the B and W models are orthogonal to each other, but the CL term for work-family conflict on hours worked became significant, as in the WxW model. As Table 2 shows, for the AR terms $\beta^J_{W(J)} = .204$ ($p < .001$), $\beta^C_{W(C)} = .088$ ($p < .001$), and $\beta^H_{W(H)} = .498$ ($p < .001$), whereas for the CL terms $\beta^J_{W(C)} = .006$ ($p = .334$), $\beta^C_{W(J)} = -.075$ ($p = .082$), $\beta^C_{W(H)} = .146$ ($p < .001$), $\beta^J_{W(H)} = -.005$ ($p = .698$), $\beta^H_{W(C)} = .021$ ($p = .044$), and $\beta^H_{W(J)} = .046$ ($p = .166$).

Again, as in the WxW model, here the B effects were very similar to the baseline, with average hours worked associated with increased job satisfaction, $\beta^J_{B(H)} = .029$ ($p = .002$), and average work-family conflict, $\beta^C_{B(H)} = .356$ ($p < .001$). Being a woman was associated with fewer hours worked, $\beta^H_{B(G)} = -2.106$ ($p < .001$). The key difference in this model is the B interaction among average work-family conflict and gender predicting job satisfaction, which was negative and significant, $\beta^J_{B(CG)} = -.045$ ($p = .008$). By including this interaction with the gender variable, the significant effect of average work-family conflict on average job satisfaction for men ($G = 0$) was $\beta^J_{B(C)} = -.176$ ($p < .001$). In turn, we can compute this coefficient for women by adding the interaction effect to net $-.045 + (-.176) = -.221$. This is sensible given that the estimated coefficient in the baseline model, which averages across men and women, is approximately the average of these two coefficients at $-.196$. The implication is that, on average over time, women appear to have a stronger association between work-family conflict and job satisfaction (see e.g., Grandey et al., 2005).

Finally, the estimated effect of gender on average job satisfaction was similar yet slightly smaller, $\beta^J_{B(G)} = -.011$ ($p = .694$). However, in this model, this represents the effect of gender on average job satisfaction when average work-family conflict $= 0$, which is the grand-mean of the latent B work-family variable. However, in our model all latent variables have means of zero. Thus, this effect

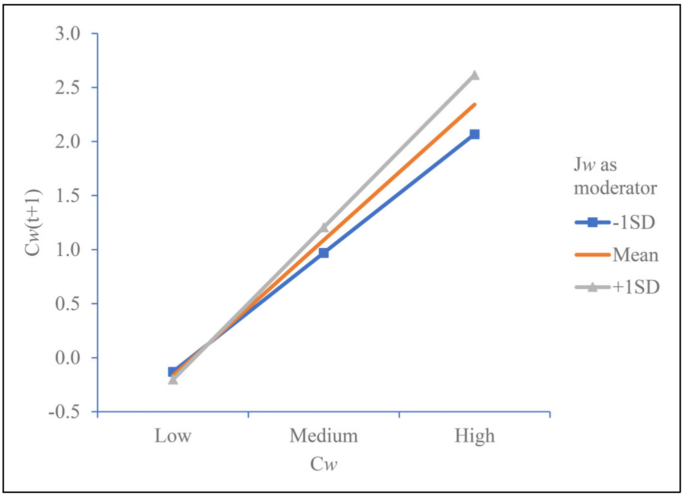

Figure 4. Plot of The *WxW* model's interaction effect of work-family conflict (C) and Job satisfaction (J).
Note: Plotted using only AR and CL effects for simplicity, showing how the *W* part of C and the *W* part J can interact to influence the future state of the *W* part of C (from the model estimates: $C_{Wt+1} = .087 * C_{Wt} - .064 * J_{Wt} + .044\ C_{Wt} * J_{Wt}$).

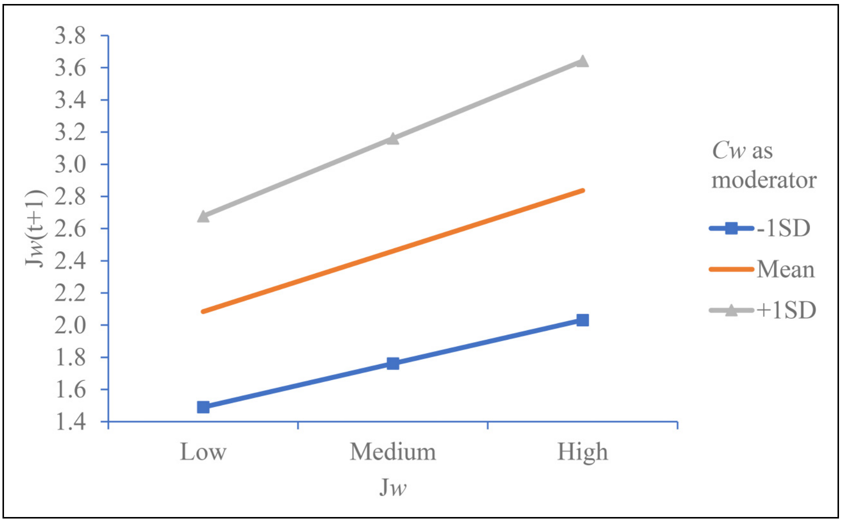

Figure 5. Plot of The *WxW* model's interaction effect of work-family conflict (C) and Job satisfaction (J).
Note: Plotted using only AR and CL effects for simplicity, showing how the *W* part of C and the *W* part J can interact to influence the future state of *W* part of J (from the model estimates: $J_{Wt+1} = .197 * J_{Wt} + .005 * C_{Wt} + .03\ C_{Wt} * J_{Wt}$).

has a similar interpretation as in the baseline model. Of course, although not statistically significant, the interaction effect could imply that average work-family conflict changes the effect of gender on job satisfaction, making job satisfaction even lower for women by −.045 as when work-family conflict increases by 1 unit. Furthermore, as we noted previously, we are reticent to give causal meanings to the *B* coefficients, but we offer this as an example.

Cross-level Interactions (*B*x*W*). In this model, the latent *B* and *W* interaction effects among work-family conflict, gender, hours worked, and job satisfaction are tested based on Eqs. 12–14.

As noted, in these models the interpretation of the AR and CL coefficients is equivalent to a 'slopes-as-outcomes' model in the multilevel context (e.g., Raudenbush & Bryk, 2002), with *W* estimates such as an AR term $\beta^{J}_{W(J)}$ being equal to the estimated average *W* effect across all individuals when all *B* interacting variables are mean-centered. The stable *B* parts of work-family conflict (C_{Bi}) and job satisfaction (J_{Bi}) are mean-centered by default because means are accounted for by observed-variable intercepts α_t. However, gender is coded 0 = *man* and 1 = *woman*, so all *W* estimates of AR and CL terms are for men, and cross-level interactions with gender indicate the difference in AR and CL terms for women (compared to estimates for men). By accounting for such cross-level interactions, we expect more efficient estimates and therefore smaller *p*-values for *W* terms.

Starting with AR terms, we estimated $\beta^{J}_{W(J)} = .395$ ($p < .001$), $\beta^{C}_{W(C)} = .206$ ($p < .001$), and $\beta^{H}_{W(H)} = .557$ ($p < .001$), which is similar to our baseline estimates for hours worked but roughly double for job satisfaction and work-family conflict, indicating greater dynamic persistence (i.e., slower regression to the mean). For women, the cross-level interactions with gender for job satisfaction was $\beta^{J}_{B(G)W(J)} = .11$ ($p = .002$), implying slower regression to the mean in job satisfaction for women as their AR term is $.11 + .395 = .505$. But, for work-family conflict the estimate was $\beta^{C}_{B(G)W(C)} = -.132$ ($p < .001$), which implies an AR term for women of $-.132 + .206 = .075$, meaning only 7.5% of a temporary increase to work-family conflict persists from one year to the next. In sum, women tend to regress to their long-run average job satisfaction more slowly than men, but women tend to regress to their long-run average for work-family conflict faster than men after experiencing a temporary change (i.e., temporary changes to work-family conflict last longer for men).

For average (stable *B*) job satisfaction and work-family conflict the story is simpler, as greater levels of both appeared to increase persistence of dynamic *W* perturbations for job satisfaction, respectively $\beta^{J}_{B(J)W(J)} = .837$ ($p =< .001$) and $\beta^{J}_{B(C)W(J)} = .365$ ($p < .001$). This indicates that higher levels of average *B* job satisfaction and work-family conflict appear to be associated with slower regression to the mean for *W* job satisfaction. However, the story is reversed for *W* work-family conflict, with average (stable *B*) job satisfaction and work-family conflict associated with reduced persistence of *W* dynamic work-family conflict, $\beta^{C}_{B(J)W(C)} = -.451$ ($p < .001$) and $\beta^{C}_{B(C)W(C)} = -.179$ ($p < .001$), implying faster regression to the mean.

In terms of CL effects, accounting for cross-level interactions leads *all* of these to become significant compared to the baseline model, with all terms interpreted as the average CL effect across the sample for men (when gender = 0) as follows: $\beta^{J}_{W(C)} = .071$ ($p < .001$), $\beta^{C}_{W(J)} = .229$ ($p < .001$), $\beta^{C}_{W(H)} = .236$ ($p < .001$), $\beta^{J}_{W(H)} = .027$ ($p = .004$), $\beta^{H}_{W(C)} = .063$ ($p < .001$), $\beta^{H}_{W(J)} = .121$ ($p < .001$). Similar to the AR effects, most estimates were larger compared to the baseline model, which can be put into some context by considering cross-level interactions.

In terms of gender: women might have a larger *W* effect of job satisfaction on work-family conflict but this interaction is not significant, $\beta^{C}_{B(G)W(J)} = .072$ ($p = .25$); and women show a significantly smaller *W* effect of work-family conflict on job satisfaction, $\beta^{J}_{B(G)W(C)} = -.041$ ($p < .001$), meaning that (relative to men) women's job satisfaction appears to be less affected by a temporary *W* increase in work-family conflict. In terms of average (stable *B*) work-family conflict: a higher average level of work-family conflict seems to amplify the effect of a temporary increase of job satisfaction on work-family conflict, $\beta^{C}_{B(C)W(J)} = .121$ ($p = .01$); and an increase in average work-family conflict seems to diminish the impact of work-family conflict's effect on job satisfaction levels over time, $\beta^{J}_{B(C)W(C)} = -.077$ ($p < .001$), suggesting greater average work-family conflict may make people less sensitive to dynamic changes in work-family conflict in terms of its effect on future job satisfaction. In terms of average (stable *B*) job satisfaction: higher average job satisfaction was related with a weaker *W* effect of work-family conflict on job satisfaction, $\beta^{J}_{B(J)W(C)} = -.239$ ($p < .001$), which implies that higher average job satisfaction may buffer people against the negative effects of dynamic changes to work-family conflict. In contrast, average job satisfaction was not associated with the *W* effect of job satisfaction on work-family conflict, $\beta^{C}_{B(J)W(J)} = .098$ ($p = .348$).

In terms of the B model, cross-level interactions appear to have reduced uncertainty in these estimates, leading to smaller p-values. However, cross-level interactions also markedly change the point estimates, which was initially vexing given that cross-level interactions should impact only W rather than B parameters (Preacher et al., 2016). Specifically, the B estimates were: $\beta^{J}_{B(C)} = -.392$ ($p < .001$), which is twice as large as in the baseline model; $\beta^{J}_{B(G)} = -.174$ ($p < .001$), which is negative and significant rather than approximately zero in the baseline model; $\beta^{J}_{B(H)} = -.003$ ($p = .726$), which is no longer positive and significant as in the baseline model; $\beta^{C}_{B(G)} = -.08$ ($p = .224$), which is reversed in sign but non-significant as in the baseline model; $\beta^{C}_{B(H)} = .269$ ($p < .001$), which is similar to the baseline model estimate; and $\beta^{H}_{B(G)} = -1.987$ ($p < .001$), which is similar to the baseline estimate.

Importantly, the reason for these changes appears to be the use of the B parts of the job satisfaction and work-family conflict variables for BxW interactions. In brief, the B parts of these or any other variables may correlate with the B variation in the W effects implied by cross-level interactions—in the multilevel context this would be typical covariation among a random intercept (stable B part) and a random slope for a W coefficient (implying a potential cross-level interaction). By estimating BxW interactions using the B parts of the observed variables, this controls for the B part of job satisfaction and work-family conflict associated with BxW interactions. Thus, when introducing BxW interactions we eliminate any B variance in the observed variables that is correlated with the 'random slope' implied by the BxW interactions, which in turn adjusts the B parameters. Therefore, researchers may want to interpret B parameters *separately* from models that include such BxW interactions.

Full Model (*W*x*W*, *B*x*B*, and *B*x*W*). To provide the reader with a sense for how estimating all CLPM interaction effects simultaneously differs from models that estimate them separately, we also estimated a 'full' model that includes all types of interactions. To save space we refer the reader to Table 2 to compare each previously discussed model's results with the full model, with a few notable similarities and differences. First, results in the full model were very similar to the BxW model, with substantial differences in B coefficients compared to baseline, WxW, and BxB models for the reasons we have just described. Again, researchers interested in B parameters may want to avoid interpreting these B terms in models that also include latent cross-level interactions (i.e., BxW terms).

Second, the full model contains two WxW interactions whereas the BxW does not, and therefore it is relevant to compare the WxW terms against those in the full model. Consistent with accounting for significant variance with the addition of many interaction terms in the full model, the WxW interactions in the full model have smaller p-values (the estimates are also somewhat larger, but this may be merely due to variance caused by the estimator).

Conflated Model. Finally, we present a traditional CLPM with work-family conflict, gender, hours worked, and job satisfaction with interactions specified *without* decomposing the variables into B and W parts. We could have used a maximum likelihood estimator, but we used a Bayes estimator with comparable fit statistics (parameter estimates were almost identical across the two types of estimators). Model fit statistics worsened compared to the baseline model: χ^2 difference for observed versus model-generated data shows a 95% CI from 2069.879 to 2180.863 (not encompassing zero, indicating poor fit) with the number of parameters $pD = 76.362$; posterior predictive p-value or $PPP < .001$ (smaller than .05, indicating poor fit); but Bayes-equivalent RMSEA = .06 (smaller than .08, indicating good fit); CFI = .973 and TLI = .957 (larger than .95, indicating good fit); and DIC = 298804.233 (larger than the baseline model's DIC of 218442.103, indicating worse fit for this model).

AR effects were all significant and substantially larger than in the baseline model: $\beta^{J}_{J} = .639$ ($p <$ .001), much larger than the baseline model's .086; $\beta^{C}_{C} = .465$ ($p < .001$), which can be compared to

.203; and $\beta_H^H = .816$ ($p < .001$), which again is much larger than the estimate obtained in the baseline model (.493). The issue here is that all stable B (co)variance in this typical CLPM must be accounted for by AR and CL terms in the model, which will often substantially overestimate AR terms in a positive direction (see Hamaker et al., 2015; Zyphur, Voelkle, et al., 2020).

The CL terms were also very different, including two that became significant in the conflated model. First, job satisfaction predicting work-family conflict $\beta_J^C = -.202$ ($p < .001$) was over twice the baseline model's estimate (of $-.079$). Second, work-family conflict predicting hours worked $\beta_C^H = .02$ ($p < .001$) was similar in magnitude but with a much smaller p-value. The other CL terms, which were similar both in magnitude and statistical significance, were: $\beta_C^J = .006$ ($p = .792$); $\beta_H^C = .136$ ($p < .001$); $\beta_H^J = -.007$ ($p = .054$), $\beta_J^H = .005$ ($p = .67$).

In terms of gender, a stable B variable, we used this to predict all occasions past the first (i.e., $t > 1$). Although this specification is uncommon (e.g., Shin & Konrad, 2017), it is the correct way to control for a B variable by using it to predict all dependent variables in the model. Results show $\beta_G^J = -.01$ ($p = .608$), which is almost identical to the baseline, but $\beta_G^C = .034$ ($p = .012$) is about half of the baseline model estimate and statistically significant. Last, $\beta_G^H = -.407$ ($p < .001$) is less than one quarter of the baseline estimate and significant.

Finally, the product term estimates were not significant, with $\beta_{CJ}^J = -.003$ ($p = .34$) and $\beta_{CJ}^C = -.001$ ($p = .92$). These estimates are not very illuminating given the fact that the product terms conflate WxW, BxB, and cross-level BxW interaction effects, as we have noted. Estimating interactions with gender would result in the same problem except we would then be conflating only BxW and any BxB interaction effects (because gender was a purely B variable). Indeed, the fact that we are unable to estimate any such interactions with the latent B parts of job satisfaction and work-family conflict further clarifies the limitations of traditional CLPMs—apart from the difficulty controlling for the stable B parts as we and others have pointed out (see Hamaker et al., 2015; Zyphur, Allison, et al., 2020; Zyphur, Voelkle, et al., 2020). For further details, we point the reader to our Online Appendices which include the Mplus code and output for all models.

Simulations

To provide the reader with some preliminary insights into how CLPMs with latent interactions perform, we also ran simulations for the two cases that we believe will be most common in future research using our approach: purely within (WxW) models and (BxW) models. In the WxW case, this allowed us to compare our approach to observed-means centering and an uncentered approach with raw variables (this uncentered approach is equivalent to the 'conflated model' described previously and in Table 2 as a traditional CLPM; for all simulation results see Table 3). Alternatively, in the BxW case, we compared our approach to only observed-means centering because it would seem incoherent to propose a cross-level interaction but use uncentered variables to test this (see results in Tables 4a and 4b). As we will now describe, our proposed CLPM approach outperformed the alternatives, and for some parameters drastically so.

Each simulation condition was replicated 500 times. The true model always had a B variance of 1.0 and a model-implied W variance of 1.0 (adjusting W residuals appropriately to account for W effects), giving an ICC(1) = .50 as is often found in longitudinal data. These unit variances imply estimated AR, CL, and interaction effects can be interpreted as being standardized. In all simulation conditions, AR terms were fixed to .5 to reflect a reasonable degree of persistence and CL terms were fixed at .3 to reflect modest cross-lagged effects. We also specified a covariance of .3 among B components to reflect a moderate correlation among B factors. To improve convergence with the Bayes estimator, we allowed for a small observed-variable residual of .03, so the total observed variance for any variable was 2.03. To account for this, we fixed observed-variable residuals to .03 when

Table 3. Results of Simulation Study 1 Examining *WxW* Interaction Effects.

	Proposed CLPM Approach					Observed-Means Approach					Uncentered Predictors Approach				
	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients
***WxW* Interaction Effect = −0.3**															
W effects															
AR effect X	0.502	0.061	−0.4	.957	.998	−0.373	0.032	174.6	.000	1.000	0.747	0.024	−49.4	.000	1.000
AR effect Y	0.500	0.061	0.0	.951	.998	−0.372	0.033	174.4	.000	1.000	0.748	0.024	−49.6	.000	1.000
CL effect $Y_W \rightarrow X_W$	0.299	0.052	0.3	.949	.992	0.044	0.033	85.3	.000	.314	0.175	0.024	41.7	.000	1.000
CL effect $X_W \rightarrow Y_W$	0.298	0.051	0.1	.947	.992	0.048	.033	84.0	.000	.330	0.175	0.024	41.7	.000	1.000
WxW interaction															
$X_W*Y_W \rightarrow X_W$	−0.300	0.046	0.0	.943	.998	−0.036	0.050	88.0	.022	.294	−0.079	0.015	73.7	.000	0.990
$X_W*Y_W \rightarrow Y_W$	−0.297	0.046	1.0	.943	.998	−0.038	0.051	87.3	.018	.276	−0.079	0.016	73.7	.000	0.994
B covariance															
X_B - Y_B	0.315	0.101	−5.0	.951	.927	0.707	0.111	−135.7	.024	1.000	—	—	—	—	—
Number of converged replications				494					500					500	
***WxW* Interaction Effect = 0**															
W effects															
AR effect X	0.489	0.090	2.2	.941	.988	−0.373	0.034	174.6	.000	1.000	0.725	0.026	−45.0	.000	1.000
AR effect Y	0.488	0.090	2.4	.946	.995	−0.373	0.036	174.6	.000	1.000	0.725	0.027	−45.0	.000	1.000
CL effect $Y_W \rightarrow X_W$	0.276	0.071	8.0	.939	.955	0.039	0.035	87.0	.000	.216	0.172	0.026	42.7	.000	1.000
CL effect $X_W \rightarrow Y_W$	0.277	0.071	7.7	.932	.962	0.043	0.036	85.7	.000	.252	0.173	0.026	42.3	.000	1.000
WxW interaction															
$X_W*Y_W \rightarrow X_W$	0.001	0.061	—	.927	.073	0.000	0.053	—	.932	.068	0.000	0.018	—	.928	.072
$X_W*Y_W \rightarrow Y_W$	−0.001	0.061	—	.951	.049	−0.002	0.055	—	.920	.080	0.000	0.018	—	.950	.050
B covariance															
X_B - Y_B	0.384	0.167	−28.0	.915	.594	0.620	0.103	−106.7	.100	1.000	—	—	—	—	—
Number of converged replications				426					500					500	
***WxW* Interaction Effect = 0.3**															
W effects															
AR effect X	0.505	0.060	−1.0	.946	.998	−0.372	0.032	174.4	.000	1.000	0.749	0.024	−49.8	.000	1.000

(continued)

Table 3. (continued)

	Proposed CLPM Approach					Observed-Means Approach					Uncentered Predictors Approach				
	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients
AR effect Y	0.505	0.060	−1.0	.926	.998	−0.372	0.033	174.4	.000	1.000	0.748	0.024	−49.6	.000	1.000
CL effect $Y_W \rightarrow X_W$	0.300	0.051	0.0	.948	.992	0.044	0.033	85.3	.000	.278	0.176	0.024	41.3	.000	1.000
CL effect $X_W \rightarrow Y_W$	0.299	0.051	0.3	.956	.992	0.047	0.033	84.3	.000	.316	0.176	0.024	41.3	.000	1.000
WxW interaction															
$X_W{*}Y_W \rightarrow X_W$	0.292	0.046	2.7	.946	.998	0.032	0.049	89.3	.014	.290	0.079	0.015	73.7	.000	.994
$X_W{*}Y_W \rightarrow Y_W$	0.294	0.046	2.0	.960	1.000	0.032	0.050	89.3	.016	.258	0.079	0.016	73.7	.000	.992
B covariance															
X_B - Y_B	0.318	0.101	−6.0	.950	.914	0.710	0.111	−136.7	.024	1.000	—	—	—	—	—
Number of converged replications				499					500					500	

Note. Parameter estimates are based on 500 replications. The AR effects were fixed at 0.5. The CL effects were fixed at 0.3. The *B* covariance was fixed at 0.3. Each replication has $N = 300$. Bias was calculated using the formula (Population parameter – Parameter estimate) / (Population parameter). 95% Coverage indicates the proportion of replications whose 95% confidence intervals include the population parameter. Proportion of Significant Coefficients indicates the proportion of replications for which the null hypothesis is rejected at $p = .05$. This value represents the power when the population parameter is nonzero and the Type-I error rate when the population parameter is zero.

Table 4a. Results of Simulation Study 2 Examining *BxW* Interaction Effects for $N = 300$ (with $T = 3/T = 4$ Results Shown).

	Proposed CLPM Approach					Observed-Means Approach				
	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients
***BxW* Interaction Effect = –0.3**										
W effects										
AR effect X	0.499/0.498	0.078/0.054	0.2/0.4	.943/.946	1.000/1.000	−0.371/−0.136	0.035/0.033	174.2/127.2	.000/.000	1.000/.998
AR effect Y	0.505/0.502	0.073/0.049	−1.0/−0.4	.959/.958	1.000/1.000	−0.349/−0.097	0.033/0.031	169.8/119.4	.000/.000	1.000.908
CL effect $Y_W \rightarrow X_W$	0.291/0.293	0.059/0.041	3.0/2.3	.936/.960	.992/1.000	0.039/0.069	0.036/0.033	87.0/77.0	.000/.000	.206/.508
CL effect $X_W \rightarrow Y_W$	0.295/0.296	0.058/0.042	1.7/1.3	.943/.966	.992/1.000	0.054/0.084	0.033/0.029	82.0/72.0	.000/.000	.366/.810
BxW interaction										
$X_W{*}Y_B \rightarrow Y_W$	−0.302/−0.303	0.073/0.047	−0.7/−1.0	.941/.958	1.000/1.000	−0.082/−0.124	0.025/0.023	−72.7/−58.7	.000/.000	0.910/.996
B covariance										
X_B - Y_B	0.306/0.304	0.127/0.107	−2.0/−1.3	.941/.944	.697/.859	0.621/0.660	0.107/0.102	−107.0/−120.0	.100/.032	1.000/1.000
Number of converged replications				488/498					500/500	
***BxW* Interaction Effect = 0**										
W effects										
AR effect X	0.492/0.496	0.090/0.060	1.6/0.8	.924/.945	.997/1.000	−0.371/−0.136	0.035/0.033	174.2/127.2	.000/.000	1.000/1.000
AR effect Y	0.510/0.506	0.089/0.060	−2.0/−1.2	.918/.916	.997/1.000	−0.373/−0.134	0.035/0.032	174.6/126.8	.000/.000	1.000/.998
CL effect $Y_W \rightarrow X_W$	0.277/0.291	0.071/0.046	7.7/3.0	.921/.933	.952/1.000	0.041/0.072	0.036/0.033	86.3/76.0	.000/.000	.226/.574
CL effect $X_W \rightarrow Y_W$	0.285/0.296	0.701/0.046	5.0/1.3	.946/.954	.983/1.000	0.043/0.072	0.035/0.031	85.7/76.0	.000/.000	.212/.628
BxW interaction										
$X_W{*}Y_B \rightarrow Y_W$	−0.003/−0.005	0.096/0.054	—/—	.952/.933	.048/.067	0.000/−0.002	0.026/0.025	—/—	.930/.910	.070/.090
B covariance										
X_B - Y_B	0.319/0.282	0.181/0.151	−6.3/6.0	.898/.887	.414/.484	0.620/0.653	0.106/0.102	−106.7/−117.7	.096/.028	1.000/1.000
Number of converged replications				353/417					500/500	

(continued)

Table 4a. (continued)

	Proposed CLPM Approach					Observed-Means Approach				
	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficients
***BxW* Interaction Effect = 0.3**										
W effects										
AR effect X	0.498/0.501	0.079/0.054	0.4/−0.2	.944/.932	.998/1.000	−0.372/−0.136	0.035/0.033	174.4/127.2	.000/.000	1.000/.998
AR effect Y	0.505/0.501	0.073/0.049	−1.0/−0.2	.944/.918	.998/1.000	−0.348/−0.099	0.033/0.031	169.9/119.8	.000/.000	1.000/.916
CL effect $Y_W \rightarrow X_W$	0.292/0.294	0.059/0.041	2.7/2.0	.963/.930	.994/1.000	0.040/0.069	0.036/0.033	86.7/77.0	.000/.000	.234/.528
CL effect $X_W \rightarrow Y_W$	0.295/0.296	0.059/0.042	1.7/1.3	.944/.954	.996/1.000	0.053/0.084	0.033/0.029	82.3/72.0	.000/.000	.358/.768
BxW interaction										
$X_W{*}Y_B \rightarrow Y_W$	0.303/0.300	0.072/0.047	−1.0/0.0	.963/.930	1.000/1.000	0.081/0.120	0.025/0.023	73.0/60.0	.000/.000	.862/.994
B covariance										
X_B - Y_B	0.305/0.302	0.128/0.108	−1.7/−0.7	.944/.964	.649/.836	0.622/0.662	0.107/.103	−107.3/−120.7	.088/.024	1.000/1.000
Number of converged replications				481/499					500/500	

Note. Parameter estimates are based on 500 replications. The first value in a cell indicates the findings for $T = 3$. The second value in a cell indicates the findings for $T = 4$. The AR effects were fixed at 0.5. The CL effects were fixed at 0.3. The *B* covariance was fixed at 0.3. Each replication has $N = 300$. Bias was calculated using the formula (Population parameter – Parameter estimate) / (Population parameter). 95% Coverage indicates the proportion of replications whose 95% confidence intervals include the population parameter. Proportion of Significant Coefficient indicates the proportion of replications for which the null hypothesis is rejected at $p = .05$. This value represents the power when the population parameter is nonzero and the Type-I error rate when the population parameter is zero.

Table 4b. Results of Simulation Study 2 Examining *BxW* Interaction Effects for $N = 2{,}000$ (with $T = 3/T = 4$ Results Shown).

	Proposed CLPM Approach					Observed-Means Approach				
	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficient	Estimates	SE	Bias in %	95% Coverage	Proportion of Significant Coefficient
***BxW* Interaction Effect = –0.3**										
W effects										
AR effect X	0.501/0.499	0.030/0.021	−0.2/0.2	.927/.950	1.000/1.000	−0.371/−0.134	0.013/0.013	174.2/126.8	.000/.000	1.000/1.000
AR effect Y	0.500/0.499	0.028/0.019	0.0/0.2	.941/.940	1.000/1.000	−0.348/−0.098	0.013/0.012	169.6/119.6	.000/.000	1.000/1.000
CL effect $Y_W \rightarrow X_W$	0.300/0.300	0.023/0.016	0.0/0.0	.949/.954	1.000/1.000	0.040/0.069	0.013/0.012	86.7/77.0	.000/.000	.832/.998
CL effect $X_W \rightarrow Y_W$	0.299/0.299	0.022/0.016	0.3/0.3	.943/.940	1.000/1.000	0.052/0.081	0.013/0.012	82.7/73.0	.000/.000	.982/1.000
BxW interaction										
$X_W{}^*Y_B \rightarrow Y_W$	−0.301/−0.300	0.024/0.017	−0.3/0.0	.947/.958	1.000/1.000	−0.083/−0.122	0.009/0.009	72.3/59.3	.000/.000	1.000/1.000
B covariance										
X_B - Y_B	0.302/0.300	0.048/0.040	−0.7/0.0	.927/.928	1.000/1.000	0.617/0.657	0.041/0.038	−105.7/−119.0	.000/.000	1.000/1.000
Number of converged replications				491/499					500/500	
***BxW* Interaction Effect = 0**										
W effects										
AR effect X	0.501/0.499	0.035/0.024	−0.2/0.2	.933/.952	1.000/1.000	−0.371/−0.134	0.013/0.013	174.2/126.8	.000/.000	1.000/1.000
AR effect Y	0.501/0.499	0.035/0.024	−0.2/0.2	.914/.952	1.000/1.000	−0.373/−0.135	0.013/0.012	174.6/127.0	.000/.000	1.000/1.000
CL effect $Y_W \rightarrow X_W$	0.296/0.300	0.028/0.018	1.3/0.0	.944/.958	1.000/1.000	0.042/0.071	0.013/0.012	86.0/76.3	.000/.000	.842/.998
CL effect $X_W \rightarrow Y_W$	0.296/0.299	0.028/0.018	1.3/0.3	.925/.932	1.000/1.000	0.040/0.069	0.013/0.012	86.7/77.0	.000/.000	.826/1.000
BxW interaction										
$X_W{}^*Y_B \rightarrow Y_W$	−0.002/0.000	0.022/0.016	—/—	.944/.952	.056/.048	−0.001/0.000	0.010/0.009	—/—	.914/.926	.086/.074
B covariance										
X_B - Y_B	0.303/0.297	0.077/0.060	−1.0/1.0	.880/.888	.889/.974	0.614/0.684	0.041/0.038	−104.7/−128.0	.000/.000	1.000/1.000
Number of converged replications				359/455					500/500	

(continued)

Table 4b. (continued)

	Proposed CLPM Approach					Observed-Means Approach				
	Estimates	*SE*	Bias in %	95% Coverage	Proportion of Significant Coefficient	Estimates	SE	Bias in %	95% Coverage	Proportion of Significant Coefficient
***BxW* Interaction Effect = 0.3**										
W effects										
AR effect X	0.500/0.500	0.030/0.021	0.0/0.0	.924/.948	1.000/1.000	−0.371/−0.134	0.013/0.013	174.2/126.8	.000/.000	1.000/1.000
AR effect Y	0.501/0.498	0.029/0.019	−0.2/0.4	.922/.944	1.000/1.000	−0.348/−0.099	0.013/0.011	169.6/119.8	.000/.000	1.000/1.000
CL effect $Y_W \rightarrow X_W$	0.301/0.301	0.022/0.016	−0.3/−0.3	.939/.950	1.000/1.000	0.041/0.069	0.013/0.012	86.3/77.0	.000/.000	.838/.998
CL effect $X_W \rightarrow Y_W$	0.300/0.299	0.022/0.016	0.0/0.3	.947/.946	1.000/1.000	0.052/0.081	0.013/0.012	82.7/73.0	.000/.000	.976/1.000
BxW interaction										
$X_W{*}Y_B \rightarrow Y_W$	0.301/0.300	0.025/0.017	−0.3/0.0	.930/.948	1.000/1.000	0.082/0.122	0.009/0.009	72.7/59.3	.000/.000	1.000/1.000
B covariance										
X_B - Y_B	0.297/0.301	0.048/0.040	1.0/−0.3	.930/.930	1.000/1.000	0.615/0.656	0.041/0.038	−105.0/−118.7	.000/.000	1.000/1.000
Number of converged replications				489/499					500/500	

Note. Parameter estimates are based on 500 replications. The first value in a cell indicates the findings for $T = 3$. The second value in a cell indicates the findings for $T = 4$. The AR effects were fixed at 0.5. The CL effects were fixed at 0.3. The *B* covariance was fixed at 0.3. Each replication has $N = 300$. Bias was calculated using the formula (Population parameter – Parameter estimate) / (Population parameter). 95% Coverage indicates the proportion of replications whose 95% confidence intervals include the population parameter. Proportion of Significant Coefficient indicates the proportion of replications for which the null hypothesis is rejected at $p = .05$. This value represents the power when the population parameter is nonzero and the Type-I error rate when the population parameter is zero.

estimating our latent CLPM. To ensure that this small variance did not impact the performance of the alternative models, for both the observed-means centering and uncentered approaches we specified a latent variable for each observed variable with variance fixed at .03, but uncorrelated with all other variables, thus adding no estimated parameters and not impacting model *df* or fit. For both the *WxW* and the *BxW* simulations, we varied the interaction effects in three conditions: fixed to −0.3, 0, and 0.3 reflecting moderately negative, zero, and moderately positive interaction effects, respectively. We chose these values to be in line with effect sizes found in prior work (e.g., Ernst Kossek & Ozeki, 1998, reported a meta-analyzed relationship of around -.3 between work-family conflict and job satisfaction).

In Simulation Study 1, we generated three waves of data ($T = 3$), as would be fairly common in applied research, and we set the sample size to $N = 300$ with a *W*x*W* interaction among the *W* parts of two observed variables x and y. This Simulation Study 1 thus has a 1 × 1 × 3 design: 1 (Sample size: $N = 300$) by 1 (Occasion: $T = 3$) by 3 (Interaction effect sizes: −.3, 0, .3). Results shown in Table 3 support the latent means approach, showing good statistical coverage and only comparatively small levels of bias for all parameters across all conditions. Alternatively, the observed-means approach underestimates AR terms, the uncentered approach overestimates AR terms, both of these methods underestimate CL terms to different degrees, and only in the no-interaction condition do these approaches offer good estimates of the interaction effect—in the −.3 and .3 interaction cases, the estimates are markedly shrunk towards zero.

Simulation Study 2 examines *B*x*W* interactions with a single interaction among the *B* part of y and the *W* part of x to predict the future *W* part of y (shown as $X_W{*}Y_B \rightarrow Y_W$ in Tables 4a and 4b), contrasting our proposed latent CLPM approach with an observed-means centering approach. In this simulation, we set $T = 3$ and $N = 300$ (shown in Table 4a), but because cross-level interactions are known to be sensitive to the number of lower-level observations per higher-level unit (see Raudenbush & Bryk, 2002), we further varied the number of waves to $T = 4$. Also, for comparison, we included conditions with the sample size of $N = 2{,}000$ (see Table 4b).[4] Simulation Study 2 thus has a 2 × 2 × 3 design: 2 (Sample size: $N = 300$ and $N = 2{,}000$) by 2 (Occasion: $T = 3$ and $T = 4$) by 3 (Interaction effect sizes: −.3, 0, .3). Results show that the latent CLPM approach outperforms the observed-means centering method in all cases, with a similar pattern for the observed-means approach: underestimated AR terms; CL terms shrunk towards zero; and cross-level interactions markedly shrunk towards zero. Notably, the larger sample size of $N = 2{,}000$ in Table 4b does not appear to improve estimates very much, but the larger $T = 4$ sample size does improve estimates for the observed-means case—this is to be expected given the longitudinal panel data modeling literature on the topic (see Nickell, 1981; Zyphur, Allison, et al., 2020; Zyphur, Voelkle, et al., 2020).

Before continuing, a relevant finding in both studies was that fewer latent CLPMs converged when interaction effects were zero, particularly with smaller samples (see Tables 3, 4a, and 4b). This may point to inherent instabilities with very small interaction effects, but in our view, this more likely suggests that a larger number of iterations may be required in the presence of very small interaction effects and small samples—we used a maximum of 20,000 iterations in our simulation runs. Requesting more iterations or thinning iterations should therefore alleviate these issues—something that is more feasible when estimating only one data set instead of 500, as was the case in our simulations. Conveniently, estimation becomes faster as sample sizes decrease, and therefore increasing the number of iterations should not present computational issues in practice. In sum, our results suggest taking a latent mean centering approach over others for estimating interactions in CLPMs. In both simulation studies, we find that our proposed approach outperforms both an observed-means as well as uncentered method with raw variables, providing less biased estimates with greater coverage to improve statistical inference. The full simulation files and results can be found in our Online Appendices.

Table 5. Overview of Software Packages for Estimating Latent Interactions in SEMs.

Software package name	Latent interaction testing procedures in alternative software and their capabilities	Key reference(s)
Mplus	Both latent moderated structural equations (LMS) and Bayesian estimation allowed. The LMS approach runs into difficulties with more than only a few dimensions because it requires numerical integration, which is computationally heavy. Bayesian estimation (as used in our approach here) is significantly faster, less biased, and has much better convergence rates particularly in high dimensions.	Asparouhov & Muthén (2020); Muthén and Muthén (1998–2017)
blavaan and JAGS (R)	The blavaan package may soon be an appealing approach to latent interactions in SEM given their plans for this development, but currently it does not support this feature (Merkle & Rosseel, 2018). In addition, researchers have noted that JAGS seems to have difficulty in handling such complex models (Depaoli et al., 2016).	Depaoli et al. (2016); Merkle & Rosseel (2018)
gllamm and gsem (Stata)	The gllamm package in Stata allows the implementation of our approach and Mplus code logic. However, researchers have noted the long computation times and instability of gllamm when estimating complex models. More precisely, according to Grilli and Rampichini (2006), computation time for interaction effects is "approximately proportional to the product of the number of quadrature points for all latent variables used. For example, if there are two random effects at level 2 (a random intercept and slope) and 8 quadrature points are used for each random effect, the time will be approximately proportional to 64" (p. 4–5). Like Mplus, Stata is a commercial software and thus may not be available to all researchers. The recent gsem package in Stata, which appears to have partially supplanted gllamm, does not appear to allow for same-level latent interactions, which our approach requires.	Grilli and Rampichini (2006); Rabe-Hesketh et al. (2005); Rabe-Hesketh and Skrondal (2012)
lavaan (R)	Cortina et al. (2021) provide R code for implementing a latent moderated structural equations (LMS) method with fully and partially latent approaches. However, they note that latent procedures, specifically moderation, involve "complex or impractical" methods that make researchers abandon latent interaction testing (p. 16). Cheung and colleagues (2021) suggest that lavaan cannot be used for LMS. Similarly, Cortina et al. (2021) recommend the use of Mplus for LMS.	Cheung et al. (2021); Cortina et al. (2021)
nlsem (R)	This package is useful for nonlinear latent interactions. However, the estimation of complex models can be very slow particularly in high dimensions (Umbach et al., 2017).	Umbach et al. (2017)
OpenMx (R)	This package can be useful in simple multilevel SEMs, however, more complex models are currently not supported. Specifically, Neale et al. (2016) state that "the OpenMx development team is working toward a much more general solution that would accommodate cross-classified models as well as large and complex data. We are still working on	Neale et al. (2016)

(continued)

Table 5. (continued)

Software package name	Latent interaction testing procedures in alternative software and their capabilities	Key reference(s)
	syntax for general purpose multilevel models that is both comprehensive and simple to understand" (p. 547).	
xxM (R)	This package is designed for multilevel SEM analysis. However, Pritikin et al. (2017) report challenges associated with estimating complex models. Discussions in the OpenMx forum indicate that there are no recent developments regarding the implementation of latent interactions but interested readers can watch these packages as developers update them with new estimation algorithms and coding methods for latent interactions (AdminNeale, 2020a, 2020b; Pritikin, 2020).	AdminNeale (2020a, 2020b); Pritikin (2020); Pritikin et al. (2017)

Discussion

This paper describes how to estimate interactions among stable (between-level or B) and/or time-varying (within-level or W) parts of latent variables in CLPMs. We propose that prior research examining interaction effects in CLPMs overlooks the potential importance of separating different types of (latent) interaction effects. As a consequence, interpretations of interaction effects in past CLPMs and conclusions with regards to hypothesis testing may have been flawed. At the same time, some researchers may have concluded that interaction effects did not exist when, in fact, they did if B and W components had been properly separated. To address this issue and assist researchers interested in testing interaction effects in CLPMs, we offer a latent interactions approach that can be used to detect previously unconsidered forms of moderation involving B and/or W parts of predictors. As noted above, although not all of the many types of interactions and nonlinear effects that our framework allows may be of interest, the convenient feature of our proposed method is that any one or more of these effects can be precisely specified and estimated to test theory, while excluding others that are not theoretically relevant.

Conveniently, our approach can be easily applied to other methods that have a similar process of forming latent W terms, including the general cross-lagged panel model or GCLM wherein W terms are referred to as 'impulses' u (see Zyphur, Allison, et al., 2020); the 'xtdpdml' panel data model (see Allison et al., 2017; Moral-Benito et al., 2019; Williams et al., 2018); as well as other methods (e.g., Bollen & Brand, 2010). In general, whenever a lagged-effects structure exists that separates stable B and dynamic W components, we recommend formally delineating WxW, BxB, and BxW interactions —although in other models the B terms may need to be rescaled (for discussion see Zyphur, Allison, et al., 2020).

Furthermore, it is notable that WxW and BxW interactions can be applied to indirect effects using AR and CL terms (for general discussion see Preacher et al., 2016; Zyphur et al., 2018). As Zyphur, Allison, et al. (2020) note, the logic of an 'impulse response' can be used to model the long-run effects of a 1-unit change in a predictor variable as effects 'flow' through a system over time along AR and CL paths. The implied impulse responses at low and high levels of a moderating variable involved in an interaction can be plotted to gain insight into potential long-run effects at different values of the W or B components of variables involved in an interaction. The approach we take here can also be extended to additional cases, including latent-indicator models to account for measurement error or multiple-group models (see Mulder & Hamaker, 2021), which can be easily implemented in SEM (see Little, 2013), as well as indirect-effects in the B model.

In terms of our applied example, prior research has come a long way in attempts to tease apart causal associations and potential interactions/moderation effects. However, much of this work has not accounted for potential stable B versus dynamic W components (e.g., Spector et al., 2007). For example, Grandey et al. (2005) found that an annual change in job satisfaction was predicted by work-family conflict for women only. Alternatively, using our data and our B and W decomposition with latent interactions we found an effect of past work-family conflict on future job satisfaction, which was actually stronger for men than for women, after accounting for hours worked. Using the framework we advance here, future studies can better interrogate the causal relationships among these variables including potential higher-order interactions and forms of nonlinearity.

Limitations and Future Research

We note four key limitations of our proposed method. First, although our method is computationally tractable, it can still be burdensome to estimate many interaction effects in large datasets. To speed up estimation, we used an AWS EC2 instance with up to 48 physical cores (96 virtual cores), which could still take a few hours to converge—along with relevant checks for convergence as recommended by Asparouhov and Muthén (2020). Run on a Linux instance rather than Windows allowed using all 48 cores to speed up processing. However, discovering this fact by running Windows and Linux instances on AWS for quite some time, including thousands of simulation runs, cost our research team roughly $3,000 USD—an expense future researchers will not need to incur if using cheaper and much faster Linux instances on AWS. Furthermore, our sample size was comparatively large with almost 8,000 people measured at 4 occasions for 3 separate variables, resulting in 12 latent W variables and 3 latent B variables. Hopefully our efforts will have removed much of the required trial and error, and costs, associated with latent interaction CLPMs.

Furthermore, future research may be able to draw from advances in the field of Bayesian statistics to reduce convergence time (e.g., Hamiltonian Monte Carlo; see Gelman et al., 2013) and utilize upcoming CPUs with even greater core counts and higher performance. Also, in practice, researchers can exclude any latent interactions that are not theoretically relevant, thus reducing computation times. In fact, for Simulation Study 1, it took about 50 h to estimate a total of 1,500 models involving two WxW interaction effects ($T = 3$, $N = 300$) with only 2 time-series variables x and y using our proposed CLPM approach on a standard office computer. Assuming this sample size will be relatively standard, this is a modest 2 min of estimation time per model.

The second limitation is that our approach is currently feasible only in Mplus using a wide (single-level) SEM framework and is thus limited to approximately 10 occasions of measurement ($T = 10$). Mplus is currently the only available software package, to our knowledge, that can deal effectively with many latent interactions in SEMs. For the interested reader, we present commonly-used software packages and their limitations with regards to our proposed approach in Table 5. We hope that future work will provide researchers with alternatives to proprietary software. Regarding the number of measurement occasions, we suppose that for $T > 10$, there may be too many observed variables to ensure reasonable computation times. In such cases, researchers may opt for a multilevel framework such as DSEM (see Asparouhov et al., 2018), which currently allows cross-level BxW interactions but not yet latent WxW or BxB interactions (see Asparouhov et al., 2018; Asparouhov & Muthén, 2019; Hamaker & Muthén, 2020).

Relatedly, a third limitation is that our proposed approach makes it burdensome to estimate 'random slopes' for AR and CL terms, as in multilevel models (including DSEM). Although BxW interactions typically involve a 'slopes-as-outcomes' model, where a W effect is allowed to have a B variance (see Raudenbush & Bryk, 2002), the model input and logic in the single-level case (such as ours) is rather complicated. Although it is technically possible to specify a latent variable to capture B variation in AR and CL terms, results of our second simulation study suggest that estimates of this B variation will probably not be trustworthy given the small T involved (contrast $T = 3$ and $T = 4$ results in Tables 4a and 4b). Therefore, we have not pursued this approach here and we do not recommend it. Instead, we recommend our approach wherein BxW terms are treated as fixed, which the reader can selectively interpret as a slopes-as-outcomes model that excludes a B residual for the slope.

The fourth limitation is the fact that models that include BxW interactions with the stable B part of a longitudinal variable may bias estimates of purely B regressions, and likewise B regressions may bias the estimation of BxW interactions. This may occur with few occasions of measurement because of the inherent uncertainty in B estimates that are involved in these model components with small T. Therefore, if researchers are interested in testing hypotheses associated with purely B effects, then these should likely be evaluated in models that exclude BxW interactions among the B variables involved. However, we would point out that this does not seem like a very serious problem given that the entire point of lagged-effects models, such as CLPMs, is to avoid having to make causal inferences based on purely B 'effects'. Instead, we recommend focusing on AR and CL terms in the W part of a CLPM and estimating WxW and BxW interactions as needed to test theory (while allowing for covariance among B parts to hold them constant, i.e., a 'fixed-effects' method).

In terms of future research, we propose that a key direction for innovation is the development of standardization methods for latent interaction effects in CLPMs. Typically, obtaining standardized effect estimates when testing interactions is done by standardizing variables before forming product terms. However, as we have decomposed observed variables into latent B and W parts for analysis, standardizing observed variables before analysis is not possible. Instead, it is theoretically possible to standardize the latent B and W variables in our model by imposing a unit variance and zero mean on them. To this end, Asparouhov and Muthén (2020) suggest a standardization approach that could overcome the need to impose such a model specification. When doing so, previous work has pointed out that it may be important to account for B variation in W variances used for standardization (see Schuurman et al., 2016). Future research should examine how this process can be facilitated in the context of latent interaction effects like those we have presented here.

Acknowledgment

The first three authors Ozlem Ozkok, Manuel J. Vaulont, and Michael J. Zyphur contributed equally to this manuscript. All authors are grateful for input provided by Bengt Muthén, Linda Muthén, and Tihomir Asparouhov. This study has been realized using data collected by the Swiss Household Panel (SHP), which is based at the Swiss Centre of Expertise in the Social Sciences FORS. The SHP project is supported by the Swiss National Science Foundation.

ORCID iDs

Ozlem Ozkok https://orcid.org/0000-0003-0764-5568
Michael J. Zyphur https://orcid.org/0000-0003-3237-7892

Supplemental Material

Supplemental material for this article is available online.

Notes

1. In the present paper, we use the terms *interaction* and *moderation* interchangeably because these methods are statistically equivalent.
2. We note that the inclusion of a direct-rating measure (e.g., trait job satisfaction) along with typical repeated measures capturing state job satisfaction would require treating the direct-rating measure as being subjected to sampling error due to its one-time measurement, as well as the removal of B components from the repeated

measures so they can reflect only the intended W variation. Because a person's average state job satisfaction is indicative of their level of trait job satisfaction (i.e., the B component), failing to remove the B variation in the repeated measure would result in conflated estimates. We leave a longer treatment of this topic to future work.

3. Furthermore, because longitudinal data are not strictly nested but instead are crossed with time/measurement occasions, it is important to control or otherwise account for shared occasions-specific effects with an observed-variable intercept (automatically done in SEM; see 'occasion effects' in Zyphur, Allison, et al., 2020). For now, we focus on B and W terms, and later add when describing our formal SEM specification.
4. A number of management studies with cross-lagged panel models reported sample sizes between 200 and 3,000 (Alisic & Wiese, 2020, $N = 3{,}118$; Cieslak et al., 2007, $N = 247$; Guan et al., 2017, $N = 228$; Hakanen et al., 2008, $N = 2{,}555$; Praskova et al., 2014, $N = 216$; Rantanen et al., 2008, $N = 365$). Thus, for our simulation studies we specified $N = 300$ and $N = 2000$ to reflect this range of sample sizes.

References

Abrigo, M. R., & Love, I. (2016). Estimation of panel vector autoregression in Stata. *The Stata Journal*, *16*(3), 778-804. doi:10.1177/1536867X1601600314

AdminNeale (2020a, 25 September). *Re: Recent developments? [discussion post]*. OpenMX Latent Interactions Discussion Board. Retrieved in February 2021 from https://openmx.ssri.psu.edu/thread/466.

AdminNeale (2020b, 25 May). *Re: G x E with latent moderator? Several methods [discussion post]*. OpenMX Discussion Board. Retrieved in February 2021 from https://openmx.ssri.psu.edu/node/4619.

Alisic, A., & Wiese, B. S. (2020). Keeping an insecure career under control: The longitudinal interplay of career insecurity, self-management, and self-efficacy. *Journal of Vocational Behavior*, *120*, 103431. https://doi.org/10.1016/j.jvb.2020.103431

Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum likelihood for cross-lagged panel models with fixed effects. *Socius*, *3*, 1-17. https://doi.org/10.1177/2378023117710578

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359-388. https://doi.org/10.1080/10705511.2017.1406803

Asparouhov, T., & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling*, *26*(1), 119-142. https://doi.org/10.1080/10705511.2018.1511375

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis using Mplus: Technical implementation*. Muthén & Muthén. https://www.statmodel.com/download/Bayes3.pdf

Asparouhov, T., & Muthén, B. (2020). Bayesian estimation of single and multilevel models with latent variable interactions. *Structural Equation Modeling*, *28*(2), 1-15. https://doi.org/10.1080/10705511.2020.1761808

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1-26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Beal, D. J. (2015). ESM 2.0: State of the art and future potential of experience sampling methods in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*, 383-407. https://doi.org/10.1146/annurev-orgpsych-032414-111335

Bliese, P. D., Schepker, D. J., Essman, S. M., & Ployhart, R. E. (2020). Bridging methodological divides between macro-and microresearch: Endogeneity and methods for panel data. *Journal of Management*, *46*(1), 70-99. https://doi.org/10.1177/0149206319868016

Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A structural equations approach. *Social Forces*, *89*(1), 1-34. https://doi.org/10.1353/sof.2010.0072

Bruck, C. S., Allen, T. D., & Spector, P. E. (2002). The relation between work–family conflict and job satisfaction: A finer-grained analysis. *Journal of Vocational Behavior*, *60*, 336-353. https://doi.org/10.1006/jvbe.2001.1836

Carpenter, M. A., & Fredrickson, J. W. (2001). Top management teams, global strategic posture, and the moderating role of uncertainty. *Academy of Management Journal*, *44*(3), 533-545. https://doi.org/10.5465/3069368

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods*, *1*(4), 421-483. https://doi.org/10.1177%2F109442819814004

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2021). Testing moderation in business and psychological studies with latent moderated structural equations. *Journal of Business and Psychology*. https://doi.org/10.1007/s10869-020-09717-0

Cieslak, R., Knoll, N., & Luszczynska, A. (2007). Reciprocal relations among job demands, job control, and social support are moderated by neuroticism: A cross-lagged analysis. *Journal of Vocational Behavior*, *71*(1), 84-96. https://doi.org/10.1016/j.jvb.2007.03.001

Cortina, J. M., Markell-Goldstein, H. M., Green, J. P., & Chang, Y. (2021). How are we testing interactions in latent variable models? Surging forward or fighting shy? *Organizational Research Methods*, *24*(1), 26-54. https://doi.org/10.1177/1094428119872531

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, *38*(4), 529-569. https://doi.org/10.1207/s15327906mbr3804_5

Depaoli, S., Clifton, J. P., & Cobb, P. R. (2016). Just another gibbs sampler (JAGS) flexible software for MCMC implementation. *Journal of Educational and Behavioral Statistics*, *41*(6), 628-649. https://doi.org/10.3102%2F1076998616664876

Dormann, C., & Zapf, D. (1999). Social support, social stressors at work, and depressive symptoms: Testing for main and moderating effects with structural equations in a three-wave longitudinal study. *Journal of Applied Psychology*, *84*(6), 874-884. https://doi.org/10.1037/0021-9010.84.6.874

Eby, L. T., Butts, M. M., Hoffman, B. J., & Sauer, J. B. (2015). Cross-lagged relations between mentoring received from supervisors and employee OCBs: Disentangling causal direction and identifying boundary conditions. *Journal of Applied Psychology*, *100*(4), 1275-1285. https://psycnet.apa.org/doi/10.1037/a0038628

Ernst Kossek, E., & Ozeki, C. (1998). Work–family conflict, policies, and the job-life satisfaction relationship: A review and directions for organizational behavior-human resources research. *Journal of Applied Psychology*, *83*(2), 139-149. https://doi.org/10.1037/0021-9010.83.2.139

Fleeson, W., & Gallagher, P. (2009). The implications of Big five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, *97*(6), 1097. https://psycnet.apa.org/doi/10.1037/a0016786

Frei, C., & Sousa-Poza, A. (2012). Overqualification: Permanent or transitory? *Applied Economics*, *44*(14), 1837-1847. https://doi.org/10.1080/00036846.2011.554380

Gabriel, A. S., Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts, M. M. (2019). Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement. *Organizational Research Methods*, *22*(4), 969-1006. https://doi.org/10.1177%2F1094428118802626

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). *CRC Press*.

Grandey, A. A., Cordeiro, B. L., & Crouter, A. C. (2005). A longitudinal and multi-source test of the work–family conflict and job satisfaction relationship. *Journal of Occupational and Organizational Psychology*, *78*(3), 305-323. https://doi.org/10.1348/096317905X26769

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424. https://doi.org/10.2307/1912791

Grilli, L., & Rampichini, C. (2006). *A review of random effects modelling using gllamm in Stata*. Department of Statistics, University of Florence.

Guan, Y., Zhuang, M., Cai, Z., Ding, Y., Wang, Y., Huang, Z., & Lai, X. (2017). Modeling dynamics in career construction: Reciprocal relationship between future work self and career exploration. *Journal of Vocational Behavior*, *101*, 21-31. https://doi.org/10.1016/j.jvb.2017.04.003

Hakanen, J. J., Schaufeli, W. B., & Ahola, K. (2008). The Job demands-resources model: A three-year cross-lagged study of burnout, depression, commitment, and work engagement. *Work & Stress*, *22*(3), 224-241. https://doi.org/10.1080/02678370802379432

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, *53*(6), 820-841. https://doi.org/10.1080/00273171.2018.1446819

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102-116. https://psycnet.apa.org/doi/10.1037/a0038889

Hamaker, E. L., & Muthén, B. (2020). The fixed versus random effects debate and how it relates to centering in multilevel modeling. *Psychological Methods*, *25*(3), 365-379. https://psycnet.apa.org/doi/10.1037/met0000239

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, *50*(1), 1-22. https://doi.org/10.1080/00273171.2014.962683

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches*. Routledge. https://doi.org/10.4324/9780429060274

Houkes, I., Janssen, P. P., de Jonge, J., & Bakker, A. B. (2003). Personality, work characteristics and employee well-being: A longitudinal analysis of additive and moderating effects. *Journal of Occupational Health Psychology*, *8*(1), 20-38. https://psycnet.apa.org/doi/10.1037/1076-8998.8.1.20

Knecht, M. K., Bauer, G. F., Gutzwiller, F., & Hämmig, O. (2011). Persistent work-life conflict and health satisfaction–A representative longitudinal study in Switzerland. *BMC Public Health*, *11*(1), 271. https://doi.org/10.1186/1471-2458-11-271

Koval, P., & Kuppens, P. (2012). Changing emotion dynamics: Individual differences in the effect of anticipatory social stress on emotional inertia. *Emotion*, *12*(2), 256–267. https://doi.org/10.1037/a0024756

Koval, P., Sütterlin, S., & Kuppens, P. (2016). Emotional inertia is associated with lower well-being when controlling for differences in emotional context. *Frontiers in Psychology*, *6*, 1997. https://doi.org/10.3389/fpsyg.2015.01997

Lian, H., Ferris, D. L., Morrison, R., & Brown, D. J. (2014). Blame it on the supervisor or the subordinate? Reciprocal relations between abusive supervision and organizational deviance. *Journal of Applied Psychology*, *99*(4), 651-664. https://psycnet.apa.org/doi/10.1037/a0035498

Lin, B., Law, K. S., & Zhou, J. (2017). Why is underemployment related to creativity and OCB? A task-crafting explanation of the curvilinear moderated relations. *Academy of Management Journal*, *60*(1), 156-177. https://doi.org/10.5465/amj.2014.0470

Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford.

Love, I., & Zicchino, L. (2006). Financial development and dynamic investment behavior: Evidence from panel VAR. *The Quarterly Review of Economics and Finance*, *46*(2), 190-210. https://doi.org/10.1016/j.qref.2005.11.007

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203-229. https://psycnet.apa.org/doi/10.1037/a0012869

Merkle, E. C., & Rosseel, Y. (2018). Blavaan: Bayesian structural equation models via parameter expansion. *Journal of Statistical Software*, *85*(4), 1–30. https://doi.org/10.18637/jss.v085.i04

Moral-Benito, E., Allison, P., & Williams, R. (2019). Dynamic panel data modelling using maximum likelihood: An alternative to arellano-bond. *Applied Economics*, *51*(20), 2221-2232. https://doi.org/10.1080/00036846.2018.1540854

Mulder, J. D., & Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling*, *28*(4), 638-648. https://doi.org/10.1080/10705511.2020.1784738

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide (8th Ed.)*. Muthén & Muthén.

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., & Boker, S. M. (2016). Openmx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535-549. https://doi.org/10.1007/s11336-014-9435-8

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, *49*(6), 1417-1426. https://doi.org/10.2307/1911408

Oesch, D., & Lipps, O. (2013). Does unemployment hurt less if there is more of it around? A panel analysis of life satisfaction in Germany and Switzerland. *European Sociological Review*, *29*(5), 955-967. https://doi.org/10.1093/esr/jcs071

Ozkok, O., Zyphur, M. J., Barsky, A. P., Theilacker, M., Donnellan, M. B., & Oswald, F. L. (2019). Modeling measurement as a sequential process: Autoregressive confirmatory factor analysis (AR-CFA). *Frontiers in Psychology*, *10*, 2108. https://doi.org/10.3389/fpsyg.2019.02108

Praskova, A., Hood, M., & Creed, P. A. (2014). Testing a calling model of psychological career success in Australian young adults: A longitudinal study. *Journal of Vocational Behavior*, *85*(1), 125-135. https://doi.org/10.1016/j.jvb.2014.04.004

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*(4), 437-448. https://doi.org/10.3102%2F10769986031004437

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods*, *21*(2), 189-205. https://psycnet.apa.org/doi/10.1037/met0000052

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*(3), 209-233. https://psycnet.apa.org/doi/10.1037/a0020141

Pritikin, J. (2020, 25 September). *Re: Recent developments? [discussion post]*. OpenMX Latent Interactions Discussion Board. Retrieved in February 2021 from https://openmx.ssri.psu.edu/thread/466. https://doi.org/10.1080/10705511.2017.1293542

Pritikin, J. N., Hunter, M. D., von Oertzen, T., Brick, T. R., & Boker, S. M. (2017). Many-level multilevel structural equation modeling: An efficient evaluation strategy. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(5), 684-698.

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed.). Stata Press.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2005). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics*, *128*(2), 301-323. https://doi.org/10.1016/j.jeconom.2004.08.017

Rantanen, J., Kinnunen, U., Feldt, T., & Pulkkinen, L. (2008). Work–family conflict and psychological well-being: Stability and cross-lagged relations within one-and six-year follow-ups. *Journal of Vocational Behavior*, *73*(1), 37-51. https://doi.org/10.1016/j.jvb.2008.01.001

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). *Sage*.

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*(2), 206-221. https://doi.org/10.1037/met0000062

Selig, J. P., Preacher, K. J., & Little, T. D. (2012). Modeling time-dependent association in longitudinal data: A lag as moderator approach. *Multivariate Behavioral Research*, *47*(5), 697-716. https://doi.org/10.1080/00273171.2012.715557

Shin, D., & Konrad, A. M. (2017). Causality between high-performance work systems and organizational performance. *Journal of Management*, *43*(4), 973-997. https://doi.org/10.1177%2F0149206314544746

Spector, P. E., Allen, T. D., Poelmans, S. A., Lapierre, L. M., Cooper, C. L., & Brough, P. (2007). Cross-national differences in relationships of work demands, job satisfaction, and turnover intentions with work-family conflict. *Personnel Psychology*, *60*(4), 805-835. https://doi.org/10.1111/j.1744-6570.2007.00092.x

Suls, J., Green, P., & Hillis, S. (1998). Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin*, *24*(2), 127-136. https://doi.org/10.1177%2F0146167298242002

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*(3), 500-517. https://psycnet.apa.org/doi/10.1037/0021-9010.88.3.500

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*(4), 703-742. https://doi.org/10.1111/j.1744-6570.1991.tb00696.x

Umbach, N., Naumann, K., Brandt, H., & Kelava, A. (2017). Fitting nonlinear structural equation models in R with package nlsem. *Journal of Statistical Software*, *77*(7), 1-20. http://dx.doi.org/10.18637/jss.v077.i07

Voorpostel, M., Tillmann, R., Lebert, F., Kuhn, U., Lipps, O., Ryser, V.-A., & Wernli, B. (2018). *Swiss Household panel userguide (1999–2017), wave 19, December 2018*. FORS.

Williams, R., Allison, P. D., & Moral-Benito, E. (2018). Linear dynamic panel-data estimation using maximum likelihood and structural equation modeling. *The Stata Journal*, *18*(2), 293-326. https://doi.org/10.1177%2F1536867X1801800201

Yuan, K. H., Cheng, Y., & Maxwell, S. (2014). Moderation analysis using a two-level regression model. *Psychometrika*, *79*(4), 701-732. https://doi.org/10.1007/s11336-013-9357-x

Zablah, A. R., Carlson, B. D., Donavan, D. T., Maxham, J. G.III, & Brown, T. J. (2016). A cross-lagged test of the association between customer satisfaction and employee job satisfaction in a relational context. *Journal of Applied Psychology*, *101*(5), 743-755. https://doi.org/10.1037/apl0000079

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods*, *12*(4), 695-719. https://doi.org/10.1177%2F1094428108327450

Zhou, L., Wang, M., & Zhang, Z. (2021). Intensive longitudinal data analyses with dynamic structural equation modeling. *Organizational Research Methods*, *24*(2), 219-250. https://doi.org/10.1177%2F1094428119833164

Zyphur, M. J., Allison, P. D., Tay, L., Voelkle, M. C., Preacher, K. J., Zhang, Z., & Diener, E. (2020). From data to causes I: Building a general cross-lagged panel model (GCLM). *Organizational Research Methods*, *23*(4), 651-687. https://doi.org/10.1177%2F1094428119847278

Zyphur, M. J., & Oswald, F. L. (2015). Bayesian estimation and inference: A user's guide. *Journal of Management*, *41*(2), 390-420. https://doi.org/10.1177/0149206313501200

Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., & Diener, E. (2020). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods*, *23*(4), 688-716. https://doi.org/10.1177%2F1094428119847280

Zyphur, M. J., Zhang, Z., Preacher, K. J., & Bird, L. J. (2018). Moderated mediation in multilevel structural equation models: Decomposing effects of race on math achievement within versus between high schools in the United States. In S. E. Humphrey, & J. M. LeBreton (Eds.), *The handbook of multilevel theory, measurement, and analysis* (pp. 473-494). American Psychological Association. https://psycnet.apa.org/doi/10.1037/0000115-021

Author Biographies

Ozlem Ozkok is an assistant professor of marketing at Rennes School of Business, France. Her areas of research are dynamic panel data models, marketing frontlines (sales/service employees and technologies), social network analysis, and (open) service innovation.

Manuel J. Vaulont is an assistant professor of management at HKUST Business School at the Hong Kong University of Science and Technology. Manuel received his Ph.D. in management from Arizona State University. His research interests focus on leadership, teams, and alternative work arrangements. Manuel won the 2020 Best Student Conference Paper Award of AOM’s Research Methods Division.

Michael J. Zyphur graduated in 2006 with a PhD in Industrial and Organizational Psychology from Tulane University, New Orleans. His research focuses on the application of quantitative methods to answer questions in the social and health sciences, as well as the historical and ethical nature of statistics, probability, and social

indices. The new EDM Stata package resulting from his collaborative efforts with natural scientists in ecology and physics is now available for free download at https://jinjingli.github.io/edm/

Zhen Zhang is a professor of Management in the Edwin L. Cox School of Business at Southern Methodist University. He obtained his PhD in Human Resources and Industrial Relations in 2008 from the University of Minnesota. His research focuses on leadership process and development, biological basis of work behavior, the intersection between organizational behavior and entrepreneurship, and advanced research methods.

Kristopher J. Preacher is a professor in the Quantitative Methods (QM) program at Vanderbilt University. He is interested in bridging the gap between theory construction and model specification and assessment. His research interests include factor analysis, structural equation modeling (SEM), multilevel modeling (MLM), latent growth curve modeling, model fit, the assessment of mediation and moderation effects, and multilevel structural equation modeling (MSEM), an emerging paradigm that combines the strengths of MLM and SEM.

Peter Koval is a Senior Lecturer at the University of Melbourne. His research interests lie at the intersection of social, personality and clinical psychology with a focus on everyday emotional processes, including how people experience and regulate their emotions in response to everyday events, and how these processes relate to well-being and psychopathology.

Yixia Zheng is a PhD candidate focusing on emotion dynamics and wellbeing at the University of Melbourne.