Autoregressive and cross-lagged panel analysis for longitudinal data

2 authors:

James P. Selig
University of Arkansas for Medical Sciences

Todd D. Little
Texas Tech University

CHAPTER 16

Autoregressive and Cross-Lagged Panel Analysis for Longitudinal Data

JAMES P. SELIG
TODD D. LITTLE

The type of models we discuss in this chapter fall under the larger heading of structural equation models (SEMs) for longitudinal data. Many different names have been used for these models, including causal models (Bentler, 1980; Kenny, 1979), cross-lagged panel models (Mayer, 1986), linear panel models (Greenberg & Kessler, 1982), and autoregressive cross-lagged models (Bollen & Curran, 2006). These models are also related to the autoregressive model and the simplex model. For simplicity, we refer to these models as panel models. What all these hold in common is that they are used to examine the structural relations of repeatedly measured constructs. Recently, several authors have critiqued the frequent use of panel models. We believe that, although many of the criticisms are justified, the panel model remains a useful tool for developmental scientists.

Figure 16.1 shows a path diagram for a two-wave, two-variable panel model. Here two latent variables, X and Y, are measured on two occasions. For convenience, only the structural portion of the model is displayed, and the underlying measurement model with multiple indicators is omitted. See Little (in press) and Little, Preacher, Selig, and Card (2007) for a thorough treatment of measurement models for longitudinal data. This model can be described with the two following equations:

$$X_2 = \beta_1 X_1 + \beta_2 Y_1 + \zeta_X$$
$$Y_2 = \beta_3 Y_1 + \beta_4 X_1 + \zeta_Y$$

Here, X and Y are two different constructs measured at two time points (denoted with a subscript 1 for time 1 and 2 for time 2). The linear regression coefficients β_1 and β_3 describe the autoregressive effects, or the effect of a construct on itself measured at a later time. The autoregressive effects describe the stability of the constructs from one occasion to the next.

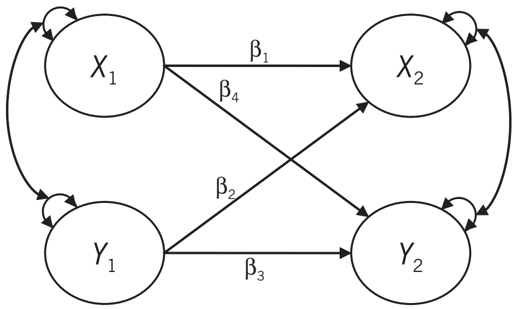

FIGURE 16.1. Path diagram for a two-wave, two-variable panel model.

More precisely, the autoregressive effects describe the stability of individual differences from one occasion to the next. A small or zero autoregressive coefficient means that there has been a substantial reshuffling of the individuals' standings on the construct over time. In contrast, a sizable autoregressive coefficient means that individuals' relative standings on the construct has changed very little over time.

The regression coefficients β_2 and β_4 represent cross-lagged effects, or the effect of a construct on another measured at a later occasion. Here again, these effects depend on individual differences on the constructs. Thus if individuals' standings on X at time 1 are related to their standings on Y at time 2, there will be a significant cross-lagged effect. A feature of the model is that the cross-lagged effects, for example of X_1 predicting Y_2, are estimated controlling for the prior level of the construct being predicted. The inclusion of the autoregressive effects means that the variance in Y_2 that can be predicted by X_1 is residual variance controlling for previous levels of Y_1 (i.e., the stable portion). This statistical feature is the reason this model is sometimes referred to as a residual change model. The fact that prior levels of the outcome construct are controlled for allows one to rule out the possibility that a cross-lagged effect is due simply to the fact that X and Y were correlated at time 1. Gollob and Reichardt (1987) and Cole and Maxwell (2003) make compelling arguments for the need to include autoregressive effects in order to minimize bias in the estimation of cross-lagged effects. The preceding model can be extended to more than two occasions and more than two constructs. The autoregressive and cross-lagged effects retain the same meaning.

Conceptual Principles and Statistical Assumptions

In the next sections, we review some of the fundamental issues in the use of panel models in developmental research. We consider critical views of the use of the panel model, in addition to the strengths of the model. As with any statistical model, panel models have numerous assumptions that must be met and various threats to validity that must be addressed.

Theory of Change

Collins (2006), in her comprehensive review of longitudinal modeling, argues that any analysis of longitudinal data must be preceded by considering "the nature of the change phenomenon that is to be observed" (p. 507). Aspects of the theoretical model of change can entail, among many other factors, the expected pattern or trajectory of change for a variable; whether change in the variable is monotonic (i.e., changing only in one direction); and what other variables will predict change. Collins's description of the theoretical model of change focuses primarily on how one variable will change over time, but it certainly encompasses how one variable is expected to induce change in another.

A potential limitation of the panel model is that it lacks an explicit theory of change, or at least that the model's theory of change may not align with the researcher's implicit theory of change. Though it is possible, using SEM, to include the mean structure in panel models, most applications of the panel model ignore information about variable means and change in means over time. Thus the way that X and Y are changing over time is not emphasized in the typical panel application (but means can be included if one has a theory of mean-level changes).

Hertzog and Nesselroade (1987) address just this issue in critiquing autoregressive models. The problem with autoregressive effects is that they describe stability in individual differences but do not describe within-person stability for the measured variable. For example, a large autoregressive coefficient for Y can mean any of the following: (1) individuals do not change on Y over time, (2) individuals uniformly increase or decrease in their levels of Y over time, or (3) individuals systematically increase and decrease over time such that one's initial level determines the direction and degree of change over time (i.e., displaying a fan-spread pattern of change over time).

A similar state of affairs exists for cross-lagged effects in the panel model. A relatively weak or nonspecific model of change for a cross-lagged effect could specify that increases in X, or the presence of X, should lead to decreasing levels of Y. For example, a researcher may believe that cognitive behavioral therapy (CBT) at time 1 will lead to a decrease in depressive symptoms at time 2. A panel model based on measures of the number of CBT sessions attended and levels of depressive symptoms for two occasions could examine whether individuals who attended more sessions also had lower scores on depressive symptoms.

Interindividual versus Intraindividual Change

Many of the potential limitations of the panel model stem from the fact that panel models focus on individual differences (interindividual variability) whereas some theories of change pertain to within-individual change (intraindividual variability). Given the previous discussion of autoregressive effects, it should be clear that although the parameters of the panel model are affected by intraindividual change (e.g., the magnitude of the autoregressive coefficient is affected by individual-level change), the parameters of the panel model are not sensitive to the type of individual-level change. The panel model could then be viewed as a potentially useful model with the limitations that it cannot easily incorporate a theory of intraindividual change and that the autoregressive and cross-lagged effects are not specific to the type of individual level-change observed over time.

Panel Models in Developmental Research

To examine the place of panel models in developmental research, we briefly review critiques of the model and then highlight its potential for use in developmental research. Critics of panel models, such as Rogosa (1987), argue that path models, such as the panel model, should be avoided because they do not begin with an explicit statement of the expected change process. Rogosa offers latent growth curve models (LGCMs) as an alternative that begins by specifying how individuals are expected to change and proceeds to identify correlates of within-person change. On the other hand, LGCMs can be viewed as models of mean structure change that are not sensitive to covariation patterns across time among the variables. For those situations in which there is neither strong theory for the change process nor a strong expectation for the trajectory of change over time or in which the research question is about the pattern of influence (and not direction of change), the panel model can be useful to explore the relations between variables over time. In this light, the results of the panel model can provide a piece of the puzzle regarding the examined association.

Reciprocal Effects

In spite of the limitations of the panel model, much of its appeal derives from the fact that it provides a means of examining relations that stem directly from theory. For example, there is a great deal of interest in the developmental sciences in reciprocal relations. Systems theory (Sameroff, 1983) places a heavy emphasis on such reciprocal relations such as those that exist between the parent and child or between the individual and his or her environment or context. Panel models make it easy to examine such reciprocal relations in that for a simple bivariate model the existence of cross-lagged effects can describe reciprocal relations. Results from a panel analysis can be used to determine whether cross-lagged effects occur in both directions (i.e., whether X_1 predicts Y_2 and Y_1 predicts X_2) and to assess the relative strength of the cross-lagged effects. For example, data based on the observation of a parent–child dyad could be analyzed to see whether a parent's behavior affects the child's subsequent behavior or the child's behavior affects the parent's subsequent behavior and even to see which of the two cross-lagged effects is stronger.

Mediation

Cole and Maxwell (2003) describe how the panel model can be used to examine mediation. A mediator is a variable that in part explains the association between two other variables. For example, Maxwell and Cole (2007) examine a hypothetical model in which the relation between parental depression and child depression is explained by problematic parenting. Panel models are useful for mediation in that the direct and indirect effects allow a statistical evaluation of mediation and are easy to estimate in the context of a panel model. The longitudinal nature of the data from the panel design provides an advantage over mediation models estimated using cross-sectional data (Cole & Maxwell, 2003; Maxwell & Cole, 2007). See MacKinnon (2008, Chapter 8) and Selig and Preacher (2009) for more information on the use of mediation models with longitudinal data.

Moderation

Panel models also make it relatively easy to examine moderation, or statistical interactions (Aiken & West, 1991). Moderation occurs when the magnitude of the relation between two variables depends on a third variable. When using measured variables, moderation can be tested by using the multiplicative product of two variables as a predictor. The effect of this product term then indicates whether moderation exists. When using latent variables, introducing a moderator is more of a challenge (but see Little, Slegers, and Card, 2006, for an approach to moderation with latent variables). Recent changes to SEM software such as Mplus (Muthén & Muthén, 2009) also make it easier to include a moderator in a latent variable model.

To sum up this section on the use of the panel model in developmental research, we acknowledge that panel models have limitations and should not be thought of as the default model for longitudinal data analysis. However, we do believe there is a place for the use of the panel model in developmental research. Results from a panel model may lead to better understanding of the longitudinal relations between variables. The results could suggest future analyses (e.g., looking for reciprocal effects, mediation, or moderation) that would shed further light on these relations. The results may also provide information on the within-person processes that correspond to the panel analyses and lead to the analysis of intraindividual data to test those ideas. In common with many statistical results, the results from a panel model will rarely be a final destination but instead will be an invitation to further investigation.

Reliability of Measurement

The ability to arrive at unbiased parameter estimates relies on the assumption that the variables are measured without error. Measurement error can be addressed using latent variables with multiple indicators (Brown, 2006). In this situation, there are two important parts of the panel model. The structural model will look much the same whether latent or measured variables are used; however, the latent variable model will have an underlying measurement model in which the specific relations between the observed measured variables and the unobserved latent variables will be specified. Use of latent variables is often desirable because, in addition to correcting for measurement error, the use of the latent variable model requires the researcher to be more explicit in stating the way the constructs in the model are defined, and a confirmatory factor analysis can be used to evaluate the success of the measurement model.

Factorial Invariance

Factorial invariance is another key assumption of longitudinal panel models. That is, is the instrument measuring the same construct at different time points? Factorial invariance is also referred to as measurement equivalence or measurement invariance (all these terms are synonymous; see Meredith, 1993). Factorial invariance does not mean that the constructs are unchanging. It addresses only the equivalence of measurement of the construct to ensure that the differences in the constructs are true differences and not confounded with problems such as changes in the discrimination (loading information) or changes in the sensitivity (mean-level information) of one or more of the indicators of a construct. The basic idea of factorial invariance is that if the construct changes over time, then this change is conveyed as changes in all the indicators in the same direction and the same amount. If one or more indicators changes differently from what was expected, then this item is no longer invariant. Factorial invariance and differential item functioning (DIF) are the same concept, with factorial invariance the method for testing indicators that are roughly continuous and DIF being the method of testing indicators that are dichotomous or polytomous in nature (Reise, Widaman, & Pugh, 1993).

Factorial invariance is assumed in any model with multiple time points or multiple groups. When only observed, or manifest, variables are used in a panel model, factorial invariance is an untestable assumption. By using latent variables with multiple indicators, however, this assumption can be tested. The procedures for testing factorial invariance are spelled out in detail in sources such as Little, Card, Slegers, and Ledford (2007).

Sample Selectivity and Selective Attrition

Two related threats to validity cover the researcher's ability to generalize results to the target population. *Sample selectivity* occurs when the sample is a nonrepresentative sample from the population. This threat to validity for panel studies (and any developmental study, for that matter) is often ignored in most research reports. Sample selectivity can be related to a poor sampling framework on the part of the experimenter or to selective participation. To the degree that population characteristics are known, one can compare the sample characteristics with the population characteristics to determine the nature of the selectivity. Reweighting the sample to adjust the sample characteristics to follow more closely the population can help reduce any bias due to sample selectivity. *Selective attrition* occurs when participants drop out of a study in a predictable way (e.g., low-socioeconomic-status [SES] participants may be more likely to drop out than higher SES participants). To the degree that characteristics of the dropout process are measured and represented on the data set, the missing data mechanism would be classified as missing at random (MAR). When missing data are due to the MAR mechanism, the information lost can be recovered with modern techniques for handling missing data.

Retest Effects

Another threat to the validity of conclusions from panel models involves retest effects. Retest effects arise from the fact that construct-irrelevant variance is often introduced when participants are measured repeatedly with the same instrument. These effects can occur when respondents remember presented material, react to repeated questioning, or lose interest in the study, to list a few examples. Such effects are mostly ignored and rarely estimated, even though powerful designs exist that allow for the control of retest effects. With the advancement of modern methods for handling missing data, planned missing-data-collection designs can also be used both to control for retest effects and to estimate the size and direction of retest effects (see discussion of planned missing designs later in the chapter and in Enders, 2010).

Model Specification

Another assumption is that the panel model is properly specified. The broad issue of specification incorporates separate assumptions. For example, all the relations examined using a panel model are assumed to be linear, or at least linear in the parameters. Model specification also entails properly structuring the associations among the variables. For example, omitting an important cross-lagged effect in an estimated model can bias the other parameter estimates. This issue is often less relevant for panel models with a smaller number of variables, but when there are many variables and many occasions, analysts will often omit paths in the pursuit of model parsimony. Therefore, care in parameter pruning is warranted.

Correct model specification also entails the assumption that all the important predictors are included in a model. For example, given the two-variable, two-wave model previously described, the cross-lagged effect of X on Y may be biased if in fact the relation between X and Y can be completely explained by their mutual dependence on some third variable Z. This assumption is arguably one of the most difficult to satisfy because there are often dozens, if not hundreds, of possible determinants of human behavior. It is difficult to argue that a bivariate model includes all relevant variables, although logic, prior research, and plausibility arguments can help narrow the field in terms of potential confounds to a study's findings.

Panel Models and Causal Claims

In the past, it was common to refer to panel models as causal models (Bentler, 1980; Kenny, 1979). Their introduction into many fields created great optimism that panel models could determine causal relations between variables in a manner similar to the way an experiment with random assignment can lend support to a causal effect. This optimism may have stemmed from the fact that panel models are in alignment with two fundamental aspects of causal inference. First, by measuring putative causes prior to the effects, temporal precedence of the cause is supported, and second, by simultaneously modeling the unique effect of several causes, it may be possible to support a causal explanation of one variable over another.

However, we view it as a positive result that panel models are no longer commonly referred to as causal models. No statistical model can determine causal relations apart from strong theory and solid research design. Furthermore, causal claims are especially difficult to make in developmental research, as a great many potential causes cannot be manipulated due to either practical or ethical limitations (Baltes, Reese, & Nesselroade, 1977).

The possible exception to the rule that results from a panel analysis provide weak support for making causal claims is the case in which the data come from an experimental design. In this case, it is the fact that the putative cause is manipulable through the random assignment of individuals to receive different levels of the causal variable, and not any property of the model itself, that supports causal inference.

There are several reasons to use caution when attempting to draw causal inference based on panel model results. For example, the putative causes often cannot be manipulated or cannot be manipulated independently from other variables in the model. In addition, proper causal inference rests on model assumptions such as including all relevant predictors. As noted earlier, this assumption can be difficult to establish.

Although it can be difficult to support causal claims using results from a panel model, we believe the panel model can be an important tool in building an argument for a causal effect of one variable on another. This support for a causal argument is especially true when the evidence from the panel model is a component in a much broader examination of the effect.

Role of Time in Panel Models

Though it is possible to use panel models with cross-sectional data, there are many advantages to using longitudinal data instead. Using longitudinal data allows the investigator to more precisely specify the hypothesized direction of effects. The temporal precedence of one variable before another can lend support to a causal claim. Finally, it allows the investigator to focus on change over time rather than static relations.

One of the fundamental choices for the use of a panel model is when to measure participants. In other words, at what developmental epoch would we expect the see the hypothesized cross-lagged and autoregressive effects? Here it is assumed that, just as the level of a variable can change as individuals develop, so too can the relation between two or more variables.

The choice of time between observations for a panel model is often not given sufficient attention. Whenever associations between variables change over time, any autoregressive or cross-lagged effect from the model is specific to the chosen lag. Thus the analyst assumes that a reasonable lag was used, meaning that it was not too long or too short and the time was appropriate to see the effect emerge. An interesting implication arises from this assumption. Most panel designs measure all variables on a fixed lag schedule. The fact that all variables are measured at the same time implicitly assumes that the time for the cross-lagged effect of X on Y and Y on X is the same. It is not difficult, however, to imagine a scenario in which the two cross-lagged effects emerge on different schedules.

When choosing time lags for a panel design, the designer must choose a lag that will provide sufficient time for the effect to occur but not so much time that the effect will have disappeared. The analyst using a panel model must assume that an appropriate lag was chosen to detect the hypothesized effect. It cannot be assumed that the appropriate lag for each relation will be the same.

Developmental Applications

Belsky, Fearon, and Bell (2007) used a panel model to examine the longitudinal relations among three constructs: parental sensitivity, child attentional control, and child externalizing behavior. Each construct was assessed on four occasions. The hypothesized relation among these constructs was that the child's attentional control would mediate the relation between quality of parenting and problem behavior. Thus the change suggested by the model was that sensitive parenting would support the child in developing attentional control, and this in turn would serve to reduce problem behaviors. In this analysis, Belsky and colleagues are capitalizing on many of the strengths of panel models. First, they examine longitudinal effects (e.g., the effect of sensitive parenting at time 1 on attentional processes at time 2), which gives them the ability to rule out possible effects (e.g., time 2 variables cannot affect time 1 variables) and the ability to explicitly test for relations in the opposite direction of those specified by theory. For example, the authors directly test for relations, such as earlier attentional processes, that predict later sensitive parenting. The panel model approach can be especially useful in testing for such reciprocal effects. Finally, the analyses of Belsky and colleagues highlight the utility of panel models for testing indirect, or mediated, effects. The authors are also careful to state that "it is risky to embrace causal conclusions without direct experimental manipulation, even with longitudinal data" (Belsky et al., 2007, p. 1239).

Illustration

Children of depressed mothers are often at risk for several negative outcomes (Downey & Coyne, 1990). In particular, children of depressed mothers are themselves at risk for experiencing issues related to depression and anxiety. Several authors, such as Downey and Coyne (1990) and Cummings and Davies (1994) describe models for family functioning in which there are reciprocal parent–child effects such that maternal depression may affect the functioning of the child and the functioning of the child may have an impact on the mother's depressive symptoms. In the following example, we examine the longitudinal association between a mother's level of depressive symptoms and her child's level of internalizing symptoms. The model presented is simplified to illustrate the key elements of the autoregressive cross-lagged model and the information that can and cannot be gained from such a model. This model is incomplete, as many other important constructs are omitted. For example, the quality of the mother–child relationship may act as a mediator between maternal depressive symptoms and the child's internalizing symptoms, or a contextual factor such as family financial strain may moderate the association between these constructs. However, the model does describe the relation between two constructs and highlights the potential reciprocity between these constructs.

The data for this example are from the NICHD Study of Early Childcare (*secc.rti.org*). Mothers and children were assessed once when the focal child was approximately 24 months of age and again when the child was approximately 36 months of age. Maternal depressive symptoms were measured using the Center for Epidemiological Studies—Depression scale (CES-D; Radloff, 1977). The 21 items of the CES-D are often reported as a single score representing depressive symptoms. To create multiple indicators for a maternal depressive symptoms construct, we created three parcels (Little, Cunningham, Shahar, & Widaman, 2002). Here the item parcels are averages of three items. Children's internalizing symptoms were assessed using the Child Behavior Checklist (CBCL; Achenbach, 1992). We used three subscales from the CBCL as indicators of an internalizing symptoms latent variable: Anxious/Depressed, Withdrawn, and Somatic Symptoms.

All analyses were conducted using Mplus 5.21. Full information maximum likelihood estimation was employed to make use of all available data. Of the original mother–child dyads, 130 were missing data for all variables and therefore excluded from the analyses, leaving a total sample size of 1,234 dyads. Latent variables were identified by setting the latent variance to 1.0, and the scale for the mean structure was set by fixing the latent means to 0. When factor loadings were equated (e.g., to test the weak invariance model), latent variances at the second occasion were freely estimated, and when intercepts were equated across occasions (e.g., to test the strong invariance model), latent means at the second occasion were freely estimated.

Prior to estimating the structural model for the longitudinal associations between maternal depressive symptoms and child internalizing symptoms, we conducted tests for longitudinal factorial invariance. Model fit indices for the configural, weak, and strong invariance models, in addition to the formal model comparisons using the nested model chi-square test, are presented in Table 16.1. Based solely on the nested model chi-square comparisons, the weak invariance constraints (i.e., equating factor loadings across occasions) significantly diminished model fit as compared with the configural invariance model ($p = .02$). However, by the change in CFI standards of either Cheung and Rensvold (2002) or Meade, Johnson, and Brady (2008), the weak invariance model is supported. The strong invariance constraints (i.e., equating indicator intercepts across occasions) did not result in a significant decrease in model fit as compared with the weak invariance model ($p = .38$) and are thus supported by any standard.

Table 16.2 shows the results for the two-wave, two-variable model. All autoregressive and cross-lagged effects are statistically significant ($p < .001$). The autoregressive

TABLE 16.1. Model Fit Indices for the Configural, Weak, and Strong Invariance Models

Model	χ^2	*df*	CFI	RMSEA (90% CI)	$\Delta\chi^2$	Δdf	*p*
Configural	81.80	42	0.995	0.028 (0.019–0.037)	—	—	—
Weak	93.19	46	0.994	0.029 (0.020–0.037)	11.39	4	.02
Strong	98.67	50	0.994	0.028 (0.020–0.036)	5.48	4	.24

coefficients for maternal depressive symptoms and child internalizing behavior show that individual differences in both constructs are relatively stable over the 12-month lag between occasions of measurement. The cross-lagged effect of maternal depressive symptoms on child internalizing behavior confirms previous findings that mothers with higher levels of depressive symptoms tend to have children with higher levels of internalizing symptoms, even when controlling for the child's previous levels of internalizing symptoms. There is also evidence of mother–child reciprocity because children with higher levels of internalizing symptoms tend to have mothers with higher levels of depressive symptoms, even when controlling for the mother's prior level of depressive symptoms.

The model accounts for over a third of the variance in maternal depressive symptoms at 36 months (R^2 = .37) with lagged maternal depressive symptoms accounting for 25% of the variance and child internalizing behavior accounting for just over 4%. More of the variance for child internalizing behavior at 36 months is accounted for by the model (R^2 = .59), with prior levels of child internalizing behavior accounting for 51% of the variance and prior levels of maternal depressive symptoms accounting for just over 1%.

TABLE 16.2. Structural Model Parameter Estimates, Standard Errors, and Standardized Estimates

Effect	Coefficient	*SE*	Standardized
Maternal depressive symptoms$_1 \rightarrow$ Maternal depressive symptoms$_2$	.48	.03	.50
Child internalizing behavior$_1 \rightarrow$ Child internalizing behavior$_2$	.75	.03	.71
Maternal depressive symptoms$_1 \rightarrow$ Child internalizing behavior$_2$	.13	.03	.12
Child internalizing behavior$_1 \rightarrow$ Maternal depressive symptoms$_2$	.20	.03	.20
Maternal depressive symptoms $_1 \leftrightarrow$ Child internalizing behavior$_1$	.39	.03	.39
Maternal depressive symptoms $_2 \leftrightarrow$ Child internalizing behavior$_2$	.09	.02	.18

Consistent with our previous discussion of the use of panel models for causal inference, we do not see these results as support for a causal effect of maternal depressive symptoms on child internalizing behavior or of child internalizing behavior on maternal depressive symptoms. The present analyses identify an interesting association that warrants further research, but with only two variables in the model and given the impossibility of manipulating either maternal depressive symptoms or child internalizing behavior, the results should not be used to bolster a causal claim without further supporting evidence.

Future Directions

Variable-Lag Models

Time plays a role in panel models in that associations are structured by time and variables are measured on multiple occasions. The role that time plays in panel models is characteristically different compared with the role of time in other models for longitudinal data that explicitly include time (e.g., LGCMs). In a panel model the autoregressive or cross-lagged effects are specific to the chosen lag between observations, but traditional panel models do not consider whether, or how, the autoregressive or cross-lagged effects may have differed if a different lag had been chosen. The use of fixed lags also complicates cross-study comparisons of similar effects. We have recently been exploring variations on the panel model in which time lags are included as predictors and the magnitude of autoregressive and cross-lagged effects can be moderated by lag (Selig, 2009; Selig, Preacher, & Little, 2009). This lag as moderator approach relies on a variable-lag design in which the lag between measurements is intentionally varied across participants (cf. McArdle & Woodcock, 1997, who used a variable lag design in conjunction with a latent growth model). We see these models as useful in that they allow one to examine how an effect changes with lag and may lead to a deeper understanding of the relations examined with panel models and a greater ability to make cross-study comparisons of effects.

Continuous Time Models

In line with our earlier comments about the role of time in panel models, we note that panel models use time in a discrete rather than a continuous fashion. A cross-lagged effect occurs across a discrete time lag. The effect is represented as static even though intuition may suggest that an effect of this sort unfolds over time. Models that use time in a continuous rather than a discrete fashion are not new. Coleman (1968) proposed continuous time models. Continuous time models would diminish the importance of choosing an appropriate lag and ease cross-study comparison of effects. Recent work by Oud (2002) demonstrates the use of continuous time models using data from a traditional panel design.

Random Effects Models

Traditional panel models treat all autoregressive and cross-lagged effects as fixed, meaning that the same regression weight is expected to apply to all members of the population. Random effects models, in contrast, hold that effects can vary across persons. Thus a particular cross-lagged effect, such as the previously described effect of maternal depressive symptoms on child internalizing symptoms, could vary across parent–child dyads. Such random effects models usually involve a time series or pooled cross-sectional design, and therefore issues of stationarity can arise (Rovine & Walls, 2006). However, the random effects approach to panel models hold great promise for examining how effects vary across persons and what person-level predictors may explain variation in these effects.

Planned Missing Designs

As mentioned earlier, many planned missing data designs exist that would be particularly useful for increasing the validity of inferences based on panel model data. For example, the three-forms questionnaire approach can reduce the fatigue associated with large protocols by randomly assigning participants to a questionnaire form. Each form would have approximately one-third of the item content randomly missing (the proportion missing is flexible). Because the random nature of the planned missing data is controlled by the experimenter, the missing data mechanism is missing completely at random and readily recoverable with modern missing data treatments (Enders, 2010; Graham, Taylor, Olchowski, & Cumsille, 2006). For controlling and estimating retest effects, participants can be randomly assigned to measurement occasion. The key here is, again, random assignment at the beginning of the study. By virtue of experimental (and statistical) control, the participants who were measured at the first assessment occasion are no different from those who did not experience the first occasion. At the second measurement occasion, any differences in the response patterns between first-time respondents and second-time respondents would be the amount and nature of the retest effects (see Graham, Olchowski, & Gilreath, 2007).

Conclusion

In summary, the panel model is especially useful for identifying the relations between variables across time. Such models are useful for initial research into the effect of one variable on another and also when the researcher seeks to elaborate on how or for whom the effect occurs (e.g., mediation or moderation). Panel models are a poor choice when the goal is to show the functional form of systematic (mean-level) change in a variable over time or if the theoretical emphasis is on intraindividual change.

In common with many statistical models, the utility of panel model results for making causal claims is much more a consideration of the data and research design than of any particular quality of the statistical model. Results from panel models using quasi-experimental or observational data should not be used as the sole support for a causal claim. However, such results could be used in conjunction with theory and other empirical results as one element in a larger argument in favor of a causal relationship. Although we acknowledge critiques of the use of the panel model for longitudinal data, we believe it has the potential to shed light on longitudinal associations between variables that can further our understanding of developmental processes.

References

Achenbach, T. M. (1992). *Manual for the Child Behavior Checklist 2–3 and 1992 Profile*. Burlington: University of Vermont, Department of Psychiatry.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Monterey, CA: Brooks/Cole.
Belsky, J., Fearon, R. M. P., & Bell, B. (2007). Parenting, attention and externalizing problems: Testing mediation longitudinally, repeatedly and reciprocally. *Journal of Child Psychology and Psychiatry, 48*, 1233–1242.
Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology, 31*, 419–456.
Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach*. Hoboken, NJ: Wiley.
Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.
Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558–577.
Coleman, J. S. (1968). The mathematical study of change. In H. M. Blalock & A. B. Blalock (Eds.), *Methodology in social research* (pp. 428–478). New York: McGraw-Hill.
Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology, 57*, 505–528.
Cummings, E. M., & Davies, P. T. (1994). Maternal depression and child development. *Journal of Clinical Psychology and Psychiatry, 35*, 73–112.
Downey, G., & Coyne, J. C. (1990). Children of depressed parents: An integrative review. *Psychological Bulletin, 108*, 50–76.
Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.
Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development, 58*, 80–92.
Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed?: Some practical clarifications of multiple imputation theory. *Prevention Science, 8*, 206–213.
Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11*, 323–343.
Greenberg, D. F., & Kessler, R. C. (1982). Equilibrium and identification in linear panel models. *Sociological Methods and Research, 10*, 435–451.
Hertzog, C., & Nesselroade, J. R. (1987). Beyond autoregressive models: Some implications of the trait–state distinction for the structural modeling of developmental change. *Child Development, 58*, 93–109.
Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley-Interscience.
Little, T. D. (in press). *Longitudinal SEM*. New York: Guilford Press.
Little, T. D., Card, N. A., Slegers, D. W., & Ledford, E. C. (2007). Representing contextual effects in multiple-group MACS models. In T. D. Little, J. A., Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 121–147). Mahwah, NJ: Erlbaum.
Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question and weighing the merits. *Structural Equation Modeling, 9*, 151–173.
Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analysis of longitudinal data. *International Journal of Behavioral Development, 31*, 357–365.
Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling, 13*, 59–72.
MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ: Erlbaum.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*, 23–44.

Mayer, L. S. (1986). On cross-lagged panel models with serially correlated errors. *Journal of Business and Economic Statistics, 4*, 347–357.

McArdle, J. J., & Woodcock, R. W. (1997). Expanding test–retest designs to include developmental time-lag components. *Psychological Methods, 2*, 403–435.

Meade, A. W., Johnson, E. C., & Brady, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*, 568–592.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543.

Muthén, L. K., & Muthén, B. O. (2009). *Mplus user's guide* (5th ed.). Los Angeles: Authors.

Oud, J. H. L. (2002). Continuous time modeling of the cross-lagged panel design. *Kwantitatieve Methoden, 69*, 1–26.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*, 552–566.

Rogosa, D. (1987). Causal models do not support scientific conclusions: A comment in support of Freedman. *Journal of Educational Statistics, 12*, 185–195.

Rovine, M. J., & Walls, T. A. (2006). A multilevel autoregressive model to describe interindividual differences in the stability of a process. In J. L. Schafer & T. A. Walls (Eds.), *Models for intensive longitudinal data* (pp. 124–147). New York: Oxford University Press.

Sameroff, A. J. (1983). Developmental systems: Contexts and evolution. In W. Kessen (Ed.), *Handbook of child psychology: Vol. 1. History, theory, and methods* (pp. 237–294). New York: Wiley.

Selig, J. P. (2009). *Where has the time gone?: The role of time lags in models for longitudinal data*. Unpublished doctoral dissertation, University of Kansas.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development, 6*, 144–164.

Selig, J. P., Preacher, K. J., & Little, T. D. (2009). Lag as moderator models for longitudinal data [Abstract]. *Multivariate Behavioral Research, 44*, 853.