

Organizational Research Methods

From Data to Causes II: Comparing Approaches to Panel Data Analysis

Journal:	<i>Organizational Research Methods</i>
Manuscript ID	ORM-16-0176.R2
Manuscript Type:	Regular Submission
Keywords:	Longitudinal data analysis < Quantitative Research, Time series < Research Design, Quantitative: Structural Equation Modeling

SCHOLARONE™
Manuscripts

Abstract

This paper compares a general cross-lagged model (GCLM) to other panel data methods based on their coherence with a causal logic and pragmatic concerns regarding modeled dynamics and hypothesis testing. We examine three ‘static’ models that do not incorporate temporal dynamics: random- and fixed-effects models that estimate contemporaneous relationships; and latent curve models. We then describe ‘dynamic’ models that incorporate temporal dynamics in the form of lagged effects: cross-lagged models estimated in a structural equation model (SEM) or multilevel model (MLM) framework; Arellano-Bond dynamic panel data methods; and autoregressive latent trajectory models. We describe the implications of overlooking temporal dynamics in static models and show how even popular cross-lagged models fail to control for stable factors over time. We also show that Arellano-Bond and autoregressive latent trajectory models have various shortcomings. By contrasting these approaches, we clarify the benefits and drawbacks of common methods for modeling panel data, including the GCLM approach we propose. We conclude with a discussion of issues regarding causal inference, including difficulties in separating different types of time-invariant and time-varying effects over time.

Note: All online supplementary materials are available for download here:

<https://doi.org/10.26188/5c9ec7295fef>

Key words: panel data model; causal inference; cross-lagged model; Granger causality; structural equation model; multilevel model; latent curve model; latent growth model; Arellano-Bond methods

From Data to Causes II: Comparing Approaches to Panel Data Analysis

Many methods exist for analyzing panel data (e.g., Arellano, 2003; Bollen & Curran, 2006; Box, Jenkins, & Reinsel, 2008). Yet, only some capitalize on the structure of panel data to offer a clear path to causal inferences. To justify such inferences, a theory of causality must be mapped onto a statistical model while addressing potential threats to causal inference (Granger, 1969, 1980). However, this is rarely done explicitly in most applications of panel data analysis.

We seek to promote a better understanding of causal modeling with panel data by showing the strengths and weaknesses of different panel data methods. For this, we use a coherence-based approach for comparing methods, while also being sensitive to their more pragmatic features. In terms of coherence, organization scholars note that it is crucial for developing and justifying theories by showing their link to existing logics and empirical findings (Locke & Golden-Biddle, 1997; Shepherd & Sutcliffe, 2011). In other words, a theory is justified if it fits with pre-existing ‘background systems’ or ‘webs of belief’ in a community of researchers (see philosophical thought by Davidson, 1986; Lehrer, 2000; Quine & Ullian, 1970).

This approach is often implicitly used to justify methods, such as by arguing for their coherence with a psychometric logic (Bagozzi & Edwards 1998; Edwards, 2011). We formally take this approach to evaluate panel data models based on their coherence with a typical view of causality, including: 1) a cause→effect temporal order; 2) possible bi-directional effects among variables; and 3) controls for potential confounders. Yet, we also recognize that models have a pragmatic character in terms of the range of dynamic processes they can capture, the richness of information from hypothesis tests, and their ease of use. We use these criteria rather than, for example, Monte Carlo simulations because the latter start by assuming that one model is true—which is not knowable in practice—to show the foregone conclusion that others are problematic.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

In what follows, we begin by describing the general cross-lagged model (GCLM) from our first paper, treating its relation to causal inference and the importance of accounting for unit effects (i.e., stable factors) and temporal dynamics (i.e., the dependence of the future on the past; Baltagi, 2013b; Hsiao, 2014). We also note the range of system dynamics and hypothesis tests associated with a GCLM, including short-run and long-run effects. Then, we contrast the GCLM against alternative panel data models, many of which are very common in organization research.

Throughout, we distinguish static and dynamic models, where only dynamic models treat dependence of the future on the past with lagged effects. We first present static models: random-effects models; fixed-effects models (i.e., group-mean centered or within-group approaches); and latent curve models. We then treat dynamic models: cross-lagged models as structural equation models (SEM) or multilevel models (MLM), including with group-mean centering; econometric Arellano-Bond dynamic panel data models; and autoregressive latent trajectory models. By evaluating these methods based on a logic of causality and pragmatic concerns, we show how: static models make causal inference problematic by excluding lagged effects; unit effects are left uncontrolled in cross-lagged models and group-mean centering produces dynamic panel bias; and Arellano-Bond and autoregressive latent trajectory methods have various shortcomings. Online materials available at <https://doi.org/10.26188/5c9ec7295fef> offer Mplus and Stata output, and the included Online Appendix A includes comparisons using data from Van Iddekinge et al. (2009) and Meier and Spector (2013), as in our first paper.

We conclude with recommendations for how to match panel data models to theory and the context of research, as well as some of the limitations for causal inference associated with controlling for confounds in panel data models. We also note that there are other methods for

analyzing panel data such as latent change score models (i.e., a model in differences; McArdle, 2001, 2009), but these can be seen as special cases of a GCLM (e.g., Voelkle & Oud, 2015).

The General Cross-Lagged Model (GCLM)

In our first paper, we started with a cross-lagged model for a unit i at an occasion t , for N units and T occasions, measured for two variables x_{it} and y_{it} . Parenthetic superscripts (x) and (y) indicate the dependent variable of the corresponding equation, and subscripts x and y indicate the predictor with which a coefficient may be associated. An h indicates a lag or lead, such as y_{it-h} for y at h occasions before t . With this, a familiar cross-lagged model is as follows:

$$x_{it} = \alpha_t^{(x)} + \beta_{x1}^{(x)} x_{it-1} + \beta_{y1}^{(x)} y_{it-1} + u_{it}^{(x)} \quad (1)$$

$$y_{it} = \alpha_t^{(y)} + \beta_{y1}^{(y)} y_{it-1} + \beta_{x1}^{(y)} x_{it-1} + u_{it}^{(y)} \quad (2)$$

wherein α_t is an occasion effect capturing factors that affect all units at a time t ; $\beta_{x1}^{(x)}$ and $\beta_{y1}^{(y)}$ are autoregressive (AR) terms, which can be thought of as the dependence of the future on the past (of the form $\beta_{yh}^{(y)}$, wherein h is a lag); $\beta_{y1}^{(x)}$ and $\beta_{x1}^{(y)}$ are cross-lagged (CL) terms, indicating short-run effects (of the form $\beta_{xh}^{(y)}$, wherein h is the lag); and u_{it} is an impulse capturing random factors with variances $\psi_{u_t}^{(x)}$ and $\psi_{u_t}^{(y)}$, and a covariance or ‘co-movement’ $\psi_{u_t}^{(xy)}$.

Here, AR and CL terms can be seen as indirect effects of past impulses on the future (seen by path tracing from an early u_{it}). Direct dependence on impulses can also be modeled, as can time-varying effects of stable factors or ‘unit effects’ to formulate a GCLM (see Figure 1):

$$x_{it} = \alpha_t^{(x)} + \lambda_t^{(x)} \eta_i^{(x)} + \beta_{x1}^{(x)} x_{it-1} + \delta_{x1}^{(x)} u_{it-1}^{(x)} + \beta_{y1}^{(x)} y_{it-1} + \delta_{y1}^{(x)} u_{it-1}^{(y)} + u_{it}^{(x)} \quad (1)$$

$$y_{it} = \alpha_t^{(y)} + \lambda_t^{(y)} \eta_i^{(y)} + \beta_{y1}^{(y)} y_{it-1} + \delta_{y1}^{(y)} u_{it-1}^{(y)} + \beta_{x1}^{(y)} x_{it-1} + \delta_{x1}^{(y)} u_{it-1}^{(x)} + u_{it}^{(y)} \quad (2)$$

wherein a unit effect η_i captures stable factors over time (e.g., individual personality or national culture), with variances $\psi_{\eta}^{(x)}$ and $\psi_{\eta}^{(y)}$, and covariance $\psi_{\eta}^{(xy)}$; λ_i is a time-varying unit effect at each occasion; and direct dependence of the future on past impulses exists as moving average (MA) terms $\delta_{x1}^{(x)}$ and $\delta_{y1}^{(y)}$, and cross-lagged moving average (CLMA) terms $\delta_{y1}^{(x)}$ and $\delta_{x1}^{(y)}$.

----- INSERT FIGURE 1 HERE -----

The GCLM strongly coheres with the logic of causality noted above: 1) causes precede effects via lagged predictors; 2) bi-directional effects are allowed by all variables predicting each other; 3) potential confounds are controlled as occasion effects and unit effects that can induce aggregate and unit-specific trends, respectively, while AR terms hold the past constant to assess predictors' unique effects. As for the pragmatic nature of a GCLM: 1) MA and CLMA terms enhance the range of dynamic processes it can model, with MA and CLMA terms that allow large temporary effects (e.g., small positive AR/CL effects and large positive MA/CLMA effects) or small persistent effects (e.g., large positive AR/CL effects and moderate negative MA/CLMA effects); 2) hypothesis tests offer rich information as short-run effects that take the form of CL + CLMA terms (e.g., $\beta_{y1}^{(x)} + \delta_{y1}^{(x)}$), whereas long-run effects can be estimated as impulse responses to show the effect of a random impulse on the future of each variable—a set of indirect effects; and 3) the GCLM is easily estimated with common SEM software.

To give some context to these assertions, we offer an abridged description of the panel data and results from our first paper. For this, it is important to keep in mind that estimating a GCLM requires choosing some number of unit effects (typically one η_i for each process variable) and a lag order for each term: p lags in an AR(p) model; q lags in an MA(q) model; c lags in an CL(c) model; and l lags in an CLMA(l) model. To choose a model, substantive and

statistical checking is used, favoring parsimonious models based on theory and contextual knowledge (Armstrong, Green, Graefe, 2015; Green & Armstrong, 2015). Our first paper illustrated this by reanalyzing income x_{it} and subjective well-being (SWB) y_{it} for $N = 135$ countries and $T = 6$ years (Diener, Tay, & Oishi, 2013). After model checking, we chose an AR(1)MA(2)CL(1)CLMA(1) model for income x_{it} and an AR(1)MA(1)CL(1)CLMA(1) model for SWB y_{it} (descriptive statistics are in our first paper; results are in Table 1 under Model 1).

----- INSERT TABLE 1 HERE -----

In terms of the dynamics implied by this model, income is highly persistent with an AR $\beta_{x1}^{(x)} = .958$, but a negative MA $\delta_{x.}^{(x)} = -.271$ reduces short-run persistence to $\beta_{x1}^{(x)} + \delta_{x.}^{(x)} = .686$. On the other hand, SWB has a smaller AR $\beta_{y1}^{(y)} = .390$, but a positive MA $\delta_{y1}^{(y)} = .191$ increases short-run persistence to $\beta_{y1}^{(y)} + \delta_{y1}^{(y)} = .581$. Also, estimated causal effects are not consistent with past findings that show significant positive effects (e.g., Diener et al., 2013). In a GCLM, the short-run income→SWB effect is $\beta_{x1}^{(y)} + \delta_{x1}^{(y)} = .139$, with long-run impulse responses showing this positive effect is highly persistent (due to income's large AR term), but CIs include zero. On the other hand, the short-run SWB→income effect is a negative $\beta_{y1}^{(x)} + \delta_{y1}^{(x)} = -.023$, and long-run impulse responses show this negative effect is highly persistent (due to income's large AR term), with CIs including zero. The point is that not only does the GCLM cohere with a logic of causality, but complex dynamics are accommodated while providing rich descriptions of long-run effects via impulse responses, all easily estimable in SEM software (in this case Mplus).

Alternative Approaches to Panel Data Analysis

Different researchers often use different methods for panel data analysis (contrast Baltagi, 2013b; Pitariu & Ployhart, 2010; Raudenbush, 2001). Yet, only some of these clearly map a causal logic to model parameters. To treat the coherence of different methods with a logic of causality and their pragmatic features, we treat common static and dynamic models. There are many variations on these models, but we offer typical specifications and descriptions. For clarity, in our text and figures u refers to a random impulse whereas we use ε as a residual that, as we note, may be conflated with unmodeled lagged effects, occasion effects, and/or unit effects.

Static Models

We define static models as those that do not specify dependence of the future on the past (i.e., excluding lagged effects; Hsiao, 2014). Thus, by ‘static’ we are referring to the nature of a statistical model rather than the data used for estimation. We begin with two common MLMs in the form of random- and fixed-effects specifications, and then discuss latent curve models.

Random-effects MLM. The MLM has gained substantial prominence in organizational science over the past 30 years (e.g., Hofmann, 1997). This approach recognizes that observations can be hierarchically structured, in our case $T = 6$ observations of SWB and income ‘nested’ in $N = 135$ countries. With this clustering, an MLM estimates relationships while modeling variation in outcomes due to lower-level factors across N and T versus higher-level factors across N . For example, Bloom (1999) predicted baseball player performance at multiple seasons with variables such as contemporaneous performance opportunity; Gulati (1995) predicted firm alliances at multiple years with contemporaneous measures of firm interdependence. Thus, inherent in this approach is treating the data as if they were a collection of T cross-sections (see Figure 2).

----- INSERT FIGURE 2 HERE -----

With SEM notation, we show this in an MLM that controls for occasion effects:

$$x_{it} = \alpha_t^{(x)} + \eta_i^{(x)} + \beta_{y0}^{(x)} y_{it} + \varepsilon_{it}^{(x)} \quad (3)$$

$$y_{it} = \alpha_t^{(y)} + \eta_i^{(y)} + \beta_{x0}^{(y)} x_{it} + \varepsilon_{it}^{(y)} \quad (4)$$

where a subscript ‘0’ on β indicates a contemporaneous relationship, η_i is a ‘random intercept’ at ‘Level-2’ or a between-unit term, and we treat ε_{it} as a ‘Level-1’ or within-unit residual.

As Figure 2 shows, each η_i is a latent variable meant to reflect stable unit-specific factors, but with one key assumption: each unit effect η_i is assumed to be uncorrelated with predictors. This is one potential source of the ‘random’ designation in a ‘random-effects’ model, where η_i is specified as uncorrelated with other variables, such as by specifying $\psi_{\eta}^{(xy)} = 0$ (see the dashed line in Figure 2)—this is often why the term ‘random effects’ is used for such models.

This approach causes concerns regarding causal inference and the pragmatic nature of the model. The first causal concern relates to the temporal nature of the effects. Without modeling temporal priority among the variables, effects like $\beta_{y0}^{(x)}$ and $\beta_{x0}^{(y)}$ require either mapping causality onto a logic of instantaneousness or researchers must accept that Eqs. 3 and 4 are misspecified by omitting lagged effects. If instantaneousness is accepted, it may be difficult to infer that x_{it} and y_{it} measure different things (or that one can be predictive of the other) because they are being modeled as immediately changing together (Gollob & Reichardt, 1987).

The next concern relates to possible bi-directional effects and controlling for confounds. Because $\beta_{y0}^{(x)}$ and $\beta_{x0}^{(y)}$ cannot both be estimated, estimating $\beta_{y0}^{(x)}$ or $\beta_{x0}^{(y)}$ requires assuming that either y or x is the cause of the other, so all x - y covariance (after accounting for α_t and η_i) is attributed to a causal effect of y on x or vice versa (as in Bloom, 1999; Gulati, 1995). This amounts to assuming that none of the following exist: a reversed causal direction; x - y feedback;

and common causes of x and y due to stable or time-varying factors, meaning no covariance due to unit effects or impulse co-movements, respectively—all of which are possible (Deaton, 2003).

Consider if $\beta_{y0}^{(x)}$ and $\beta_{x0}^{(y)}$ were separately estimated, which we do with two SEMs while imposing $\psi_{\eta}^{(xy)} = 0$. This results in an income→SWB term for Eq. 3: $\beta_{y0}^{(x)} = .162, t = 4.985, p < .001$; and an SWB→income term for Eq. 4: $\beta_{x0}^{(y)} = .799, t = 14.259, p < .001$. These cannot be decomposed into $x \rightarrow y$ and $y \rightarrow x$ causality, x - y feedback, or common causes (Geweke, 1982; Granger, 1980; Pierce & Haugh, 1977). One implication is that contemporaneous co-movements and unit effects are not controlled. This can be seen by comparing unit effect variances before and after prediction, showing $\psi_{\eta}^{(x)} = .806$ and $\psi_{\eta}^{(y)} = 1.096$ when the variable is a predictor, but $\psi_{\eta}^{(x)} = .582$ and $\psi_{\eta}^{(y)} = .369$ when the variable is an outcome, implying that $\beta_{y0}^{(x)}$ and $\beta_{x0}^{(y)}$ partly reflect unit effects due to stable factors, such as culture or various institutional characteristics. In other words, unit effects are confounded with the contemporaneous effects among the variables.

In terms of the pragmatic features of the model, although they are easy to estimate, the absence of lagged effects means that no dynamic processes can be accommodated, so short-run Granger-Sims tests and long-run impulse responses are precluded. This raises questions about the practical use of $\beta_{y0}^{(x)}$ or $\beta_{x0}^{(y)}$ for planning interventions or policy formation—this problem is symbolized by our use of ε to show that residuals are not designed to mimic a random impulse u . Although AR residuals can be specified in MLMs (Bolger & Laurenceau, 2013; Skrondal & Rabe-Hesketh, 2004, 2008), this only allows AR dynamics for ε rather than x - y effects.

In sum, random-effects MLM fails to adequately cohere with a logic of causality and suffers from pragmatic issues compared to a GCLM. By this, we do not mean that MLMs are wholly bad or wrong. They can be useful when directions of causal effects are known, unit

effects can be assumed uncorrelated or irrelevant, and lagged effects are irrelevant, possibly because of noisy data or because they are too distant in time. Thus, we do not categorically recommend against the model. Instead, we merely clarify issues associated with limited causal inference and pragmatic concerns. Of course, many researchers understand some of these issues, which leads us to the more common fixed-effects MLM specification for analyzing panel data.

Fixed-effects MLM. The fixed-effects MLM is equivalent to Eqs. 3 and 4, but it controls for unit effects—this is what econometricians often mean by ‘fixed effects’. For example, Judge, Ilies, and colleagues do this to eliminate stable individual differences to estimate within-person relationships among affect, job attitudes, work stressors, and the like (Ilies, Johnson, Judge, & Keeney, 2011; Ilies, Scott, & Judge, 2006; Judge & Ilies, 2004; Judge, Scott, & Ilies, 2006). This eliminates person-specific trends to estimate $\beta_{y0}^{(x)}$ and $\beta_{x0}^{(y)}$ using occasion-specific variation (for the prevalence of this, see Beal & Weiss, 2003; Beal, 2015; Bolger & Laurenceau, 2013; Curran & Bauer, 2011; Curran, Lee, Howard, Lane, & MacCallum, 2012; Hoffman, 2015; Hoffman & Stawski, 2009; Nezlek, 2012a, 2012b; Walls, Jung, & Schwartz, 2006; Wang & Maxwell, 2015).

This can be done in various ways (Halaby, 2004), classically with predictors to (dummy) code η_i . Equivalent SEMs allow $\eta_i^{(x)}$ or $\eta_i^{(y)}$ to covary with predictors (Allison, 2009; Bollen & Brand, 2010; Teachman, Duncan, Yeung, & Levy, 2001), or by estimating $\psi_{\eta}^{(xy)}$ as in Figure 2.

A common MLM approach is to ‘group-mean’ or ‘within-group’ center predictors to de-mean them with unit averages \bar{x}_i and \bar{y}_i (Raudenbush & Bryk, 2002). Centering all variables can also be done (as in Stata’s ‘xtreg, fe’) or first-differences can be taken such as $x_{it} - x_{it-1}$ or $y_{it} - y_{it-1}$ and “no information about β is lost” (Arellano, 2003, p. 14). These methods are ‘fixed-effects’ models because it is as if unit effects are ‘fixed’ rather than allowed to vary as confounders—this

1
2
3 language is common in econometrics, but it is useful given its popularity for panel data models
4
5 and because it clarifies that the method by which stable confounders are eliminated is irrelevant.
6
7

8 However, in terms of the coherence of the model with a causal logic and its pragmatic
9
10 nature, the same problems exist as for the random-effects MLM, save for holding unit effects
11
12 constant. Estimated as two separate models, this produces a within-country income→SWB term
13
14 for Eq. 3: $\beta_{y0}^{(x)} = .095, t = 3.507, p < .001$; and an SWB→income term for Eq. 4: $\beta_{x0}^{(y)} = .387, t =$
15
16 $3.065, p = .002$. As before, these do not incorporate a temporal order and cannot be decomposed
17
18 into $x \rightarrow y$ versus $y \rightarrow x$ causality, x - y feedback, or common causes at a time t . Although “centering
19
20 removes any between-individual variance in estimates of within-individual relations among the
21
22 variables, meaning that the relations among the within-individual variables are unconfounded by
23
24 personality or other individual differences” (Judge et al., 2006, p. 130), this does little to help
25
26 interpret effects as causal in a short-run or long-run fashion associated with a dynamic process.
27
28
29
30

31 In sum, static fixed-effects MLMs are common and control for unit effects. Yet, they fail
32
33 to cohere with the logic of causality described previously and create pragmatic dilemmas for
34
35 modeling dynamic effects over time. Again, this does not mean they are wholly problematic and
36
37 may even be considered acceptable misspecifications when time lags are too distant or data are
38
39 too noisy to observe lagged effects, but compared to a GCLM they have multiple limitations.
40
41
42

43 ***Latent Curve Model (LCM)***. Another extremely common model is the LCM (i.e., latent
44
45 growth model; McArdle & Nesselrode, 2003), which estimates unit-specific trends over time.
46
47 However, these are static models because they omit lagged effects, even though researchers often
48
49 refer to them as indicating ‘dynamic’ relationships among trends. For example, Pitariu and
50
51 Ployhart (2010) illustrate this by predicting employee performance trends using trends in effort
52
53 over time, while predicting both of these trends with a time-invariant measure of team diversity.
54
55
56
57
58
59
60

To critically explore this logic, we start with a simple and familiar example of an LCM using SEM notation (see Figure 3; for alternatives, see Curran, 2003; Bollen & Curran, 2006):

$$x_{it} = \lambda_{1t}\eta_{1i}^{(x)} + \lambda_{2t}\eta_{2i}^{(x)} + \varepsilon_{it}^{(x)} \quad (5)$$

$$y_{it} = \lambda_{1t}\eta_{1i}^{(y)} + \lambda_{2t}\eta_{2i}^{(y)} + \varepsilon_{it}^{(y)} \quad (6)$$

with all terms as before except for two changes. First, $\lambda_{1t} = 1$ and time is treated as the predictor $\lambda_{2t} = t - 1$, implying an intercept η_{1i} at $t = 1$ and a unit-specific trend or slope factor η_{2i} (Bollen & Curran, 2006). Second, occasion effects α_t are set to zero, so observed means and covariances over T are modeled entirely by the trend variables' means and covariances: $\boldsymbol{\eta}_\eta \sim MVN(\boldsymbol{\mu}_\eta, \boldsymbol{\Psi}_\eta)$. For illustrative purposes, this LCM fits our data adequately (CFI = .96, TLI = .96, SRMR = .08, RMSEA = .09), with R^2 ranging from .90 to .99 across all T for both variables.

----- INSERT FIGURE 3 HERE -----

Results show positive growth for income and SWB: income intercept $\mu_{\eta_1}^{(x)} = 7.681$, $t = 96.095$, $p < .001$; income slope $\mu_{\eta_2}^{(x)} = .016$, $t = 2.579$, $p < .01$; SWB intercept $\mu_{\eta_1}^{(y)} = 5.331$, $t = 55.417$, $p < .001$; SWB slope $\mu_{\eta_2}^{(y)} = .026$, $t = 2.491$, $p < .013$. More importantly, the relationship among the slopes can be used for inference, which is a correlation $\psi_{\eta_2}^{(xy)} = .576$, $t = 3.862$, $p < .001$, indicating a positive relationship for income and SWB trends. With this, authors often draw causal inferences by treating $\eta_{2i}^{(x)}$ or $\eta_{2i}^{(y)}$ as an outcome of the other and/or by using time-invariant predictors (e.g., Pitariu & Ployhart, 2010; Raudenbush, 2001). Yet, this creates concerns about coherence with a logic of causality and the pragmatic nature of the LCM.

To first tackle the issue of causality as a temporal process, in LCM there seems to be 'an effect of time' by using it as a predictor, leading some researchers to treat trends as if time was

1
2
3 their cause (e.g., Curran & Bauer, 2011; Curran et al., 2012; Wang & Maxwell, 2015). Yet, ‘an
4 effect of time’ here is potentially misleading, as time defines causality rather than itself being a
5 causal factor (Pitariu & Ployhart, 2010; Voelkle & Oud, 2015). In turn, by conceptualizing time
6 as a cause, researchers can easily overlook the causal processes that may be of interest, such as
7 socialization, institutionalization, or maturation. Indeed, “although time is inextricably linked to
8 the concept of development, in itself it cannot explain any aspect of developmental change”
9 (Baltes, Reese, & Nesselrode, 1988, p. 108). In turn, perhaps the GCLM offers a better way to
10 treat trends as interactions among time- and unit-specific factors $\lambda_t \eta_i$, while acknowledging that
11 lagged effects may also exist (see Arellano, 2003; Bai, 2013; Boumahdi & Thomas, 2008).
12
13
14
15
16
17
18
19
20
21
22
23

24 This point leads to concerns over the LCMs static nature. As Eqs. 5 and 6 show, there are
25 no temporal dynamics as dependence of the future on the past, so causal effects are not modeled.
26 Instead, the variables available for causal inference are time-invariant factors like $\eta_{2i}^{(x)}$ or $\eta_{2i}^{(y)}$,
27 which omit a temporal dynamic needed to establish causality (Ahn, Lee, & Schmidt, 2001, 2013;
28 Bai, 2009; Nauges & Thomas, 2003). Thus, as with static MLMs, LCMs lead to the dilemma of
29 having to accept a theory of instantaneous causality or accept that any dynamic effects captured
30 by η_i are misspecified. Also, treating an intercept $\eta_{1i}^{(x)}$ or $\eta_{1i}^{(y)}$ as a predictor based on the fact
31 that it represents an initial occasion does not help, because trends are defined by all η_i terms.
32
33
34
35
36
37
38
39
40
41
42
43

44 Next, in terms of bi-directional effects and controlling for confounds, without relying on
45 a time-varying element, there is no way to know if $\eta_{2i}^{(x)} \rightarrow \eta_{2i}^{(y)}$ versus $\eta_{2i}^{(y)} \rightarrow \eta_{2i}^{(x)}$, or if feedback
46 exists, or if a third variable is causing both. Again, the point of using panel data is to establish a
47 temporal order, which LCMs do not account for. This is exacerbated by LCMs treating change as
48 being a function of only stable factors η_i (e.g., individual personality or culture), implying that
49
50
51
52
53
54
55
56
57
58
59
60

there are no population-level occasion effects α_t that are distinct from the trends of interest, and implying that the trends are not due to the persistence of random impulses u_{it} .

To explain, η_i is time-invariant and therefore implies a systematic, deterministic process over time, such as systematic changes in income due to stable cultural factors. Yet, variables like income often show a random or stochastic trend, with AR terms close to 1 (we find $\beta_{x1}^{(x)} = .958$). Thus, the trends in LCM may be due to population factors α_t or persistent random impulses rather than systematic, stable factors over time. Indeed, in the case of stochastic trends due to persistent impulses (e.g., AR = 1), time as a predictor can erroneously account for substantial variance (Nelson & Kang, 1984). Thus, a correlation $\psi_{\eta_2}^{(xy)} = .537$ among LCM slopes (and other time-invariant variables) may be spurious, reflecting shared stochastic trends (Braun, Kuljanin, & DeShon, 2013; Kuljanin, Braun, & DeShon, 2011; Nelson & Kang, 1981; Nelson & Plosser, 1982; Watson, 1986). Indeed, some researchers note that “the assumption of a stochastic trend is often more realistic than... a deterministic trend” (Box et al., 2008, p. 101), and “deterministic trend models are so implausible that they should never be imposed unless there is very strong supporting evidence” (Harvey, 1997, p. 196). Our view is more tempered, but we do think it is reasonable to give α_t , AR, MA, CL, and CLMA terms the opportunity to account for trends.

Finally, in terms of LCM’s pragmatic nature, it allows descriptive curve fitting, but it does not incorporate dynamic effects. In turn, it offers little help for planning interventions—how should this be done using $\eta_{2i}^{(x)}$ or $\eta_{2i}^{(y)}$ given that causal order is ignored and they are time-invariant? As Stock and Watson note for determining causal effects, “[t]he most reliable way to handle a trend in a series is to transform the series so that it does not have a trend” (2003, p. 466; see also Curran & Bauer, 2011; Curran et al., 2012; Wang & Maxwell, 2015). Although there is

no single way to de-trend data (Heckman, 1991; Stock & Watson, 1988; Yule, 1921, 1926), to assume that only deterministic trends $\eta_{2i}^{(x)}$ and $\eta_{2i}^{(y)}$ exist seems questionable.

Dynamic Models

Dynamic models are differentiated from static models by incorporating lagged effects such as AR, MA, CL, and CLMA terms. In what follows we discuss common dynamic models and explore their coherence with a causal logic along with their more pragmatic characteristics.

Cross-Lagged Models. As we have described, organizational researchers regularly use a cross-lagged model to analyze panel data as follows (see Figure 4a):

$$x_{it} = \alpha_t^{(x)} + \beta_{x1}^{(x)} x_{it-1} + \beta_{y1}^{(x)} y_{it-1} + \varepsilon_{it}^{(x)} \quad (7)$$

$$y_{it} = \alpha_t^{(y)} + \beta_{y1}^{(y)} y_{it-1} + \beta_{x1}^{(y)} x_{it-1} + \varepsilon_{it}^{(y)} \quad (8)$$

wherein all terms are as described previously, except we show a residual as ε_{it} for reasons that we discuss momentarily. Although higher-order lags for AR and CL terms are possible (Cole & Maxwell, 2003; Little, 2013; Maxwell & Cole, 2007; Maxwell, Cole, & Mitchell, 2011), in practice researchers often default to an AR(1)CL(1) model (e.g., Eby, Butts, Hoffman, & Sauer, 2015; Maynard et al., 2014; Nohe, Meier, Sonntag, & Michel, 2015). We estimate this model (see Table 1, Model 2.1), which follows the original analysis of our data in Diener et al. (2013).

----- INSERT FIGURE 4 HERE -----

Model fit is adequate (CFI = .96, TLI = .95, SRMR = .07, RMSEA = .10), with AR terms showing high persistence for income $\beta_{x1}^{(x)} = .952, t = 59.783, p < .001$, and SWB $\beta_{y1}^{(y)} = .877, t = 37.467, p < .001$. Also, CL effects are positive and statistically significant, with income→SWB $\beta_{x1}^{(y)} = .108, t = 3.813, p < .001$, and SWB→income $\beta_{y1}^{(x)} = .027, t = 2.324, p = .02$. These results

point to strong persistence of random impulses and large short-run effects among the variables. However, there are a few reasons to view these results cautiously.

In terms of a logic of causation, the cross-lagged model has the benefit of incorporating lagged effects and accounts for the possibility of bi-directional effects. However, compared to GCLM results, SWB has much stronger persistence over time, which seems misaligned with past findings of SWB being mean-reverting (e.g., Clark, Frijters, & Shields, 2008; Diener & Lucas, 1999). Similarly, CL effects are larger and statistically significant, and the SWB→income effect changes sign. One reason for these differences is that conventional cross-lagged models do not control for unit effects $\eta_i^{(x)}$ and $\eta_i^{(y)}$, which can bias β terms (Baltagi, 2013b; Hamaker, Kuiper, & Grasman, 2015; Hsiao, 2014). Specifically, in Eqs. 7 and 8, the $\varepsilon_{it}^{(x)}$ and $\varepsilon_{it}^{(y)}$ terms will reflect both a unit effect η_i and an impulse u_{it} . The result is that AR terms are larger due to the covariance over time associated with η_i (Arellano, 2003; Morgan & Winship, 2014; Vaisey & Miles, 2017), and CL terms are different for the same reason. Although using time-invariant controls is often recommended (e.g., Little, 2013), this rarely controls for unit effects entirely.

Similar problems arise in MLMs, which can be used to estimate similar models (Beal & Weiss, 2003; Bolger & Laurenceau, 2013; Griffin, 1997; Nezlek, 2001, 2008, 2011, 2012a, 2012b; see also Kling, Harvey, & Maclean, 2017). To start, a random-effects MLM also fails to control unit effects (e.g., Schonfeld & Rindskopf, 2007), which we show by estimating Eqs. 7 and 8 with Stata's 'xtreg, mle' (see Table 1, Model 2.2), resulting in similar AR terms of income $\beta_{x1}^{(x)} = .94, t = 55.01, p < .001$, and SWB $\beta_{y1}^{(y)} = .856, t = 29.17, p < .001$, and an income→SWB CL term $\beta_{x1}^{(y)} = .122, t = 3.39, p < .001$, and for SWB→income $\beta_{y1}^{(x)} = .036, t = 2.56, p < .020$.

This problem of bias in MLM has been recognized and many authors attempt to solve it by group-mean centering their data (i.e., a ‘within-group’ or WG model), as if to estimate a fixed-effects static model (e.g., Beal, Trougakos, Weiss, & Green, 2006; Bono, Foldes, Vinson, & Muros, 2007; Dalal et al., 2009; Fisher & Noble, 2004; Gielnik et al., 2015; Hoffman, 2015; Ilies, Johnson, Judge, & Keeney 2011; Ilies et al., 2006; Rovine & Walls, 2006). This is often done because researchers believe that this centering presents no issues beyond those of static MLMs (e.g., Beal, 2015; Duckworth, Tsukayama, & May, 2010; Enders & Tofghi, 2007; see also Uy, Foo, & Aguinis, 2010). However, this is not the case—it causes ‘dynamic panel bias’.

To explain, centering the variables by subtracting unit means produces:

$$x_{it} - \bar{x}_i = \alpha_t^{(x)} + \beta_{x1}^{(x)}(x_{it-1} - \bar{x}_i) + \beta_{y1}^{(x)}(y_{it-1} - \bar{y}_i) + \varepsilon_{it}^{(x)} \quad (9)$$

$$y_{it} - \bar{y}_i = \alpha_t^{(y)} + \beta_{y1}^{(y)}(y_{it-1} - \bar{y}_i) + \beta_{x1}^{(y)}(x_{it-1} - \bar{x}_i) + \varepsilon_{it}^{(y)} \quad (10)$$

which attempts to control for unit effects by assuming a unit average such as \bar{y}_i is a good estimate of $\eta_i^{(y)}$. Yet, just as a scale mean will correlate with true-score and item-specific variance, \bar{y}_i is a function of a unit effect $\eta_i^{(y)}$ and impulses $u_{it}^{(y)}$ (Bond, 2002). In turn, because AR and CL terms reflect persistent impulses, subtracting \bar{y}_i induces *negative* AR bias and also impacts CL terms.

To understand this point, consider the $T = 2$ case for y_{i1} and y_{i2} . Here, subtracting \bar{y}_i induces a negative AR bias because what results is a positive and negative observation $y_{i1} - \bar{y}_i$ and $y_{i2} - \bar{y}_i$. Because the covariance used to estimate \bar{y}_i is due to $\eta_i^{(y)}$ and AR terms due to $u_{it}^{(y)}$, removing \bar{y}_i eliminates an AR effect because this effect is a function of persistence in an

impulse $u_{it}^{(y)}$ (i.e., the impulse is subtracted, and thus so is the AR effect) resulting in a biased impulse ε_{it} in Eqs. 9 and 10, and altering AR and CL estimates (Alvarez & Arellano, 2003).

For over 40 years econometricians have known about this problem in models that include lagged effects, which they refer to as dynamic panel bias (Nerlove, 1967, 1971; Sevestre & Trognon, 1985). As Nickell described in 1981, the negative bias for AR terms take the form:

$$-\frac{1 + \beta_{y1}^{(y)}}{T - 1} \quad (11)$$

Although this bias is reduced as $T \rightarrow \infty$, it can be sizable at $T = 30$ or even 100 when AR (and CL) terms are large (Hamaker & Grasman, 2015; Jongerling, Laurenceau, & Hamaker, 2015; Judson & Owen, 1999; Nerlove, Sevestre, & Balestra, 2008). Unfortunately, the bias is caused by any method that attempts to subtract unit effects, including first-differencing and predictors that (dummy) code for η_i (Arellano, 2003), which seems of particular import when the direction of CL effects is of explicit interest (e.g., Vancouver, Thompson, & Williams, 2001; Vancouver, Thompson, Tischner, & Putka, 2002; Vancouver & Kendall, 2006; Yeo & Neal, 2006). Indeed, to our knowledge the only multilevel approach that accounts for unit effects without this bias is dynamic structural equation modeling or DSEM (see Asparouhov, Hamaker, & Muthén, 2018).

To show the problem, we estimate Eqs. 11 and 12 using Stata's 'xtreg, fe' (see Table 1, Model 2.3), which group-mean centers all variables—the same results emerge when group-mean centering in cross-lagged SEM. The result is very small AR estimates for income $\beta_{x1}^{(x)} = .079$, $t = 1.48$, $p = .139$, and for SWB $\beta_{y1}^{(y)} = .031$, $t = .49$, $p = .622$, with a larger CL term income \rightarrow SWB $\beta_{x1}^{(y)} = .275$, $t = 2.18$, $p = .03$, and smaller one for SWB \rightarrow income $\beta_{y1}^{(x)} = .024$, $t = .92$, $p = .358$.

In sum, cross-lagged SEM and random-effects MLM are similarly biased, and group-mean centering or other approaches to estimate fixed-effects models lead to dynamic panel bias. This said, “[t]he fact that these two estimators are likely to be biased in opposite directions [for AR effects] is useful. Thus we might hope that a candidate consistent estimator will lie between the... [two AR] estimates” (Bond, 2002, p. 144). There are various ways to do this (e.g., Allison, Williams, Moral-Benito, 2017; Asparouhov et al., 2018; Hamaker et al., 2015), such as a cross-lagged SEM with $\eta_i^{(x)}$ and $\eta_i^{(y)}$ having an unrestricted $\psi_{\eta}^{(xy)}$ for a ‘fixed-effects’ approach that avoids dynamic panel bias (see Figure 4b). This is possible because SEM treats $\eta_i^{(x)}$ and $\eta_i^{(y)}$ as missing data that have a relationship with lagged predictors (Bai, 2013; Moral-Benito, 2013). We show this with a cross-lagged SEM wherein $\eta_i^{(x)}$ and $\eta_i^{(y)}$ have unrestricted covariance $\psi_{\eta}^{(xy)}$, with time-varying unit effects $\lambda_i^{(x)}$ and $\lambda_i^{(y)}$ and co-movements $\psi_{u_t}^{(xy)}$ (see a variant in Hamaker et al., 2015). As Figure 4b shows, this is a GCLM with MA and CLMA terms eliminated.

This improves model fit over the cross-lagged SEM, ostensibly because AR and CL effects are no longer tasked with accounting for unit effect (co)variance (CFI = .98, TLI = .97, SRMR = .03, RMSEA = .09). Also, AR terms are between the cross-lagged SEM/random-effects MLM and group-mean centered MLM estimates (see Table 1, Model 2.4). The AR term for income is $\beta_{x1}^{(x)} = .583, t = 2.98, p = .003$, and for SWB it is $\beta_{y1}^{(y)} = .674, t = 7.43, p < .001$, which come close to total AR and MA effects $\beta_{x.}^{(x)} + \delta_{x.}^{(x)}$ and $\beta_{y1}^{(y)} + \delta_{y1}^{(y)}$ in the GCLM (in Table 1, Model 1). Also, CL terms are no longer significant, with SWB→income $\beta_{y1}^{(x)} = .004, t = .134, p = .89$, and income→SWB $\beta_{x1}^{(y)} = .014, t = .19, p = .85$, showing how unit effects can account for the same covariance producing CL effects in the cross-lagged SEM/random-effects MLM.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

With this in mind, it is notable that many published cross-lagged models assume η_i away or attempt to address it with control variables (e.g., Cole & Maxwell, 2003; du Toit & Browne, 2001; Little, 2013; Maxwell & Cole, 2007; Maxwell et al., 2011). Although theoretically arguing against η_i is possible, such arguments are typically absent and control variables will probably not account for unit effects entirely. Also, even if authors wanted to control for η_i , roughly half of cross-lagged models have only $T = 2$ occasions (Hamaker et al., 2015), making it impossible to control for unit effects and estimate AR terms, which requires $T \geq 3$ (Allison et al., 2017). Like recommendations for MLMs with lagged effects (e.g., Hoffman, 2015; Nezlek, 2012a, 2012b), the problem of small T and η_i persists in even leading organizational journals (e.g., Eby et al., 2015; Maynard et al., 2014; Nohe et al., 2015; Riketta, 2008).

In sum, classic cross-lagged SEMs and MLMs should be avoided when seeking to make causal inferences in panel data. Conveniently, with $T \geq 3$ a cross-lagged SEM can be modified to account for η_i (Allison et al., 2017; Bollen & Brand, 2010; Hamaker et al., 2015; Teachman et al., 2001). Pragmatically, this is useful because cross-lagged models allow estimating short-run and long-run effects using impulse responses. Yet, dynamics are more limited than a GCLM by excluding MA and CLMA terms. Conveniently, just as cross-lagged models can be modified to account for η_i , they can also be modified to include MA and CLMA terms, as we have shown.

Arellano-Bond Methods

There are many econometric approaches to panel data analysis found in organizational research, but a popular example is the Arellano-Bond (AB) method (see overviews in Bond, 2002; Bun & Sarafidis, 2015; for foundational work see also Arellano & Bond, 1991; Arellano & Bover, 1995; Blundell & Bond, 1998; Holtz-Eakin, Newey, & Rosen, 1988). For example,

Piening Baluch, and Salge (2013) used the AB method to show a positive effect of HR practices on organizational performance (for other examples, see Barkema & Schivjen, 2008; Foster, 2010; George, 2005; Goldstein, 2012). To we examine this method by briefly outlining its logic and some of its causality-oriented and pragmatic dilemmas, with more details in Online Appendix B (see also Arellano, 2003; Baltagi, 2013b; Roodman, 2009a, 2009b).

The problem that AB methods address is that a unit effect η_i is unobserved, but fixed-effects methods that eliminate it cause dynamic panel bias. To overcome this, AB methods use variables in differences (e.g., $y_{it} - y_{it-1}$) and levels (e.g., y_{it}). Consider a classic AB model:

$$x_{it} - x_{it-1} = \alpha_t^{(x)} - \alpha_{t-1}^{(x)} + \beta_{x1}^{(x)}(x_{it-1} - x_{it-2}) + \beta_{y1}^{(x)}(y_{it-1} - y_{it-2}) + \varepsilon_{it}^{(x)} - \varepsilon_{it-1}^{(x)} \quad (12)$$

$$y_{it} - y_{it-1} = \alpha_t^{(y)} - \alpha_{t-1}^{(y)} + \beta_{y1}^{(y)}(y_{it-1} - y_{it-2}) + \beta_{x1}^{(y)}(x_{it-1} - x_{it-2}) + \varepsilon_{it}^{(y)} - \varepsilon_{it-1}^{(y)} \quad (13)$$

wherein this subtracts η_i because it exists at both current and past occasions. As we noted above, “no information about β is lost” when differencing in static models (Arellano, 2003, p. 14), but in the presence of lagged predictors this induces dynamic panel bias (Bond, 2002).

AB methods attempt to eliminate η_i and dynamic panel bias based on the fact that: 1) a first-difference $y_{it} - y_{it-1}$ eliminates η_i , but causes dynamic panel bias; and 2) a lagged level y_{it-1} includes η_i , it but does not cause dynamic panel bias. In turn, because neither η_i nor dynamic panel bias exist in the overlap of differences and levels, AB methods can be thought of as using lagged levels to predict differences, and then substituting the predicted part for the differences in Eqs. 12 and 13. This is ‘instrumenting’ or an ‘instrumental variable’ procedure, with lagged levels being instruments for differences. With this logic, a generalized methods of moments (GMM) estimator is used that we treat in Online Appendix B, but here note that this

1
2
3 ‘difference-GMM’ approach is often augmented by a ‘system-GMM’ approach that incorporates
4 variables in levels using differences as instruments (Blundell & Bond, 1998; Roodman, 2009b).
5
6

7
8 To evaluate this method, we first note that it allows a coherent temporal order for causal
9 effects. Also, bi-directional causality and potential confounds are addressed by using lagged
10 instruments, such that the past of a predictor should cause the future of an outcome, but the
11 reverse should not be true. In turn, this is meant to eliminate reverse causation and confounding
12 by common causes. Also, although the model excludes MA and CLMA terms and thus limits the
13 range of potential dynamic processes it can model, there is the pragmatic benefit of a particular
14 long-run effect that can be estimated (Baltagi, 2013b). Specifically, by estimating each model for
15 an outcome separately, this allows a thought experiment wherein a predictor is increased by 1-
16 unit and this is maintained over time, with effects ‘aggregating’ at each occasion via AR terms.
17 This allows computing a long-run effect shown here for the $x \rightarrow y$ case as $\beta_{x1}^{(y)} / (1 - \beta_{y1}^{(y)})$.
18
19
20
21
22
23
24
25
26
27
28
29
30

31 Yet, using the AB method requires checking its assumption that the information from
32 instruments is unrelated to residuals (i.e., instruments should be related to outcomes only via
33 predictors). This is checked by residual autocorrelation and Sargan/Hansen tests with a null
34 hypothesis of no instrument-residual covariance, meaning that small p -values entail rejecting the
35 assumption of valid instruments (i.e., large p -values imply assumptions are met). If these tests
36 show small p -values, instruments can be lagged further until a valid set of instruments is found.
37
38
39
40
41
42
43
44

45 Unfortunately, this approach leads to concerns related to controlling for confounds and
46 the method’s practical implementation. First, using too few instruments causes inefficiency (i.e.,
47 large SEs), but using too many causes overfitting that reintroduces unit effects. Also, if changes
48 over time are systematic, differences can correlate with η_i , rendering instruments invalid. This
49 problem is compounded by the fact that with many instruments, Sargan/Hansen tests will have
50
51
52
53
54
55
56
57
58
59
60

1
2
3 large p -values that suggest valid instruments even if they are not (see Online Appendix B;
4
5 Roodman, 2009b). Thus, as the amount of information instruments offer for GMM estimation
6
7 approaches N , both overfitting and invalid Sargan/Hansen tests are expected. This makes it hard
8
9 to ensure confounds are controlled, while leading to pragmatic concerns over how to conduct
10
11 analyses given that as T increases so does the number of potential instrument sets.
12
13

14
15 To show this, we first took a system-GMM approach using Stata's 'xtabond2', with all
16
17 available lags to instrument the equations in differences and a single lag for equations in levels
18
19 (see Table 1, Model 3.1; see Online Appendix B and Stata output). The AR terms show a larger
20
21 estimate for income when compared to the previous cross-lagged SEM controlling for η_i , $\beta_{x1}^{(x)} =$
22
23 $.982$, $t = 14.81$, $p < .001$, but a more similar estimate for SWB $\beta_{y1}^{(y)} = .66$, $t = 5.94$, $p < .001$,
24
25 with CL terms showing no SWB→income effect $\beta_{y1}^{(x)} = .011$, $t = .28$, $p = .782$, nor an
26
27 income→SWB effect $\beta_{x1}^{(y)} = .184$, $t = 1.55$, $p = .121$. Here, the effects of SWB are similar to
28
29 what we expect, but the effects of income are similar to those from typical cross-lagged
30
31 SEM/random-effects MLM, suggesting overfitting and/or a correlation among η_i and
32
33 differences.
34
35
36
37
38
39

40
41 When checking the assumption of instrument validity, the Sargan/Hansen tests show $p <$
42
43 $.5$ in levels equations. Yet, with many instruments this p -value is biased towards zero, and thus
44
45 even this large p -value suggests a potential correlation among η_i and differenced instruments
46
47 because, as noted in Online Appendix B, the amount of information gained by the instruments is
48
49 close to our sample size of $N = 135$. Thus, large p -values are expected for Sargan/Hansen tests
50
51 and therefore $p < .05$ may indicate invalid instruments—but we cannot know this with certainty.
52
53
54
55
56
57
58
59
60

To address these issues, we estimate a second model in differences only to reduce the instrument count, which shows better Sargan/Hansen test results (the smallest $p = .683$). Here, AR terms somewhat acceptable but are much less efficient (see Table 1, Model 3.2), with income $\beta_{x1}^{(x)} = .458, t = 1.47, p = .141$, and SWB $\beta_{y1}^{(y)} = .411, t = 2.12, p = .034$, with CL terms showing an SWB→income effect $\beta_{y1}^{(x)} = .138, t = 1.94, p = .054$, but no income→SWB effect $\beta_{x1}^{(y)} = .063, t = .08, p = .938$. However, we are wary of these results because changing the instrument set changes them dramatically. Indeed, by varying the instrument sets, we observed AR effects for income ranging from -.3 to 1.2, as well as large ranges for other coefficients. This shows the added problem of having to correctly specify both a substantive model *and* instrument sets.

To overcome the issues of GMM, maximum-likelihood approaches in SEM exist (Allison et al., 2017; Bai, 2013; Moral-Benito, 2013), such as Stata's 'xtpdml' tool (Williams, Allison, & Moral-Benito, 2018). To emphasize the pragmatic value of SEM for panel data, we illustrate a similar approach using separate models for income and SWB, with a first occasion $t = 1$ allowed to freely correlate with η_i and all predictors (for the income→SWB case, see Figure 5). Also, predictors are allowed to covary with η_i to produce a fixed-effects model that avoids dynamic panel bias, and predictors covary with past and contemporaneous residuals to account for potential reverse-causation and common causes (see Allison et al., 2017; Bollen & Brand, 2010).

----- INSERT FIGURE 5 HERE -----

The models we estimate show adequate fit (for income, CFI = .99, TLI = .97, SRMR = .02, RMSEA = .07; for SWB, CFI = .97, TLI = .94, SRMR = .03, RMSEA = .11), and results are consistent with the cross-lagged SEM with covariance among unit effects in Figure 4b (see Table 1, Model 3.3 and compare with Model 2.4). The AR term for income is $\beta_{x1}^{(x)} = .59, t = 2.95, p =$

.003, and for SWB $\beta_{y1}^{(y)} = .62, t = 7.44, p < .001$, with CL effects SWB→income $\beta_{y1}^{(x)} = .053, t = .877, p = .38$ and income→SWB $\beta_{x1}^{(y)} = .076, t = .275, p = .783$. These results have the benefit of maximum-likelihood, which addresses some issues of GMM (Hsiao, Pesaran, & Tahmiscioglu, 2002). This shows the benefits of SEM in general, but also Figure 5 suggests a GCLM extension that has causal and pragmatic benefits: control variables can be included without being specified as outcomes by allowing them to covary with η_i as well as past and current occasions of other variables—although there are issues to consider with this approach, noted in Online Appendix B.

In sum, AB methods have issues in terms of controlling for confounds while also having the pragmatic problem of being difficult to use. This may be why researchers caution that “[w]here system GMM offers the most hope, it may offer the least help” (Roodman, 2009a). To overcome the problems of AB methods, an SEM approach using either our GCLM or separate models for dependent variables can be used. In our view, this shows the benefits of our SEM framework in general, even for those trained in an econometric tradition.

Autoregressive Latent Trajectories

The autoregressive latent trajectory (ALT) model combines cross-lagged and LCM methods (Curran & Bollen, 2001; Bollen & Curran, 2004; 2006; Bollen & Zimmer, 2010). Although ALT is not common in organizational research, we include it here because it models both trends and lagged effects, which addresses many causal and pragmatic concerns we have with other methods. We show an ALT as follows (Figure 6a):

$$x_{it} = \lambda_{1i} \eta_{1i}^{(x)} + \lambda_{2i} \eta_{2i}^{(x)} + \beta_{x1}^{(x)} x_{it-1} + \beta_{y1}^{(x)} y_{it-1} + \epsilon_{it}^{(x)} \quad (15)$$

$$y_{it} = \lambda_{1i} \eta_{1i}^{(y)} + \lambda_{2i} \eta_{2i}^{(y)} + \beta_{y1}^{(y)} y_{it-1} + \beta_{x1}^{(y)} x_{it-1} + \epsilon_{it}^{(y)} \quad (16)$$

1
2
3 wherein all terms are as before for the LCM, with $\lambda_{1t} = 1$ and $\lambda_{2t} = t - 1$ implying a unit-specific
4
5
6 linear trend with intercept η_{1i} and slope η_{2i} , which have an estimated mean structure (involving
7
8
9 μ_η) as in an LCM. With this generality, the ALT is similar to various approaches that detrend
10
11
12 observed data in a model-based way while estimating AR and CL effects (see Ahn et al., 2001,
13
14 2013; Bai, 2009; Nauges & Thomas, 2003). However, like the LCM, occasion effects α_i are
15
16
17 constrained to zero so that a mean structure must be accounted for by the η_i trend terms.
18

19 ----- INSERT FIGURE 6 HERE -----
20

21
22 Although ALT models are often given substantive interpretations for trends associated
23
24 with η_i , we treat the ALT as a model that estimates lagged effects while accounting for η_i and
25
26 avoiding dynamic panel bias. To show this, we estimate Eqs. 15 and 16 with a linear trend as in
27
28 Figure 6a (Table 1, Model 4.1). This model has unrestricted covariance among initial occasions
29
30 and unit effects, which helps avoid assumptions of about η_i at the first occasion (for insight, see
31
32 Allison et al., 2017; Hamaker, 2005). Yet, this ALT imposes a functional form for η_i (here a
33
34 linear trajectory) that can strongly impact estimates of AR and CL terms. We show this with the
35
36 second ALT in Figure 6b, which as only one effect for each of x and y , but with λ_i unrestricted
37
38 as in our GCLM (Table 1, Model 4.2). This differs from other models above by constraining α_i
39
40 to zero, which identifies a mean structure for η_i and clarifies how LCMs and ALTs ignore
41
42 occasion effects (i.e., global shocks), which is a strong assumption of that all trends are
43
44 systematic rather than also being due to time-specific effects that the ALT does not control.
45
46
47
48
49
50

51
52 In our demonstration, maximum-likelihood estimation of the first ALT failed, as we often
53
54 encounter in the presence of missing data in early occasions that are not treated as dependent
55
56
57
58
59
60

variables—a potential pragmatic issue of the ALT. To solve this, we used a Bayes procedure with Markov Chain Monte Carlo estimation with ‘uninformative’ or ‘diffuse’ priors to approximate maximum-likelihood results (Muthén & Asparouhov, 2012; see Online Appendix C for more details). For consistency, we report results using frequentist concepts such as t -values and p -values, but rely on their Bayesian analogues that are based on posterior distributions.

Results for the first model in Figure 6a show effects that are different from those above (Table 1, Model 4.1). The AR term for income is much smaller than that found in our GCLM, with $\beta_{x1}^{(x)} = .083$, $t = .703$, $p = .112$, but SWB is in a more similar range, with $\beta_{y1}^{(y)} = .451$, $t = 4.465$, $p < .001$. The CL effects also change but remain non-significant, with an SWB→income effect $\beta_{y1}^{(x)} = .019$, $t = .5$, $p = .156$, and an income→SWB effect $\beta_{x1}^{(y)} = .067$, $t = .335$, $p = .186$.

Alternatively, the second ALT in Figure 6b with a single term η_i and an unrestricted λ_i (and α_i constrained to zero) for each series also has adequate fit to the data (CFI = .97, TLI = .96, SRMR = .06, RMSEA = .09) but shows different results (see Table 1, Model 4.2). Here, AR effects are more reasonable for income $\beta_{x1}^{(x)} = .561$, $t = 2.042$, $p = .041$, and for SWB $\beta_{y1}^{(y)} = .597$, $t = 5.369$, $p < .001$. Although the CL effects show no significant SWB→income effect $\beta_{y1}^{(x)} = .046$, $t = .898$, $p = .369$, the income→SWB effect is $\beta_{x1}^{(y)} = .318$, $t = 3.662$, $p < .001$. This difference between these ALT models occurs because the trends compete with lagged effects to explain observed covariance (Heckman, 1991; Voelkle, 2008; Wang & Maxwell, 2015). Indeed, especially if $t = 1$ occasions are treated as independent, ALT trends can easily account for too much covariance, which may be exacerbated in multivariate cases (Sims, 1996, 2000).

Although the second ALT with time-varying unit effects shows AR terms that appear more reasonable, it has an income→SWB effect that is not supported in our GCLM or the cross-

1
2
3 lagged model that controls for η_i (Figure 4b and Table 1, Model 2.4), and is unsupported in the
4
5
6 previous AB alternative (Figure 5 and Table 1, Model 3.3). The key difference between the
7
8 second ALT and the other models is that occasion effects are constrained to zero with a mean
9
10 structure imposed for unit effects—so that if an occasion effect α_t increases, a unit effect η_i
11
12 increases—which illustrates the importance of estimating occasion effects rather than imposing a
13
14 mean structure that is necessarily incorporated into a systematic trend associated with η_i .
15
16
17

18
19 Perhaps what is most important about these comparisons is that we do not find major
20
21 differences between AR and CL effects when using models that account for time-varying unit
22
23 effects terms and allow occasion effects in SEM (Figure 4b, Table 1, Model 2.4; Figure 6b,
24
25 Table 1, Model 3.3). Indeed, these models show effects that are similar to those from our full
26
27 GCLM (Figure 1, Table 1, Model 1), unlike models that: ignore η_i (i.e., cross-lagged models;
28
29 Figure 4a, Table 1, Models 2.1 and 2.2); attempt eliminate η_i as if it were observed (i.e., within-
30
31 group centering; Table 1, Model 2.3); or over-specify systematic trends (i.e., the first ALT;
32
33 Figure 6a, Table 1, Model 4.1). In sum, the ALT we describe above—with AR and CL terms
34
35 constrained to equality over time—may be a valuable tool for researchers who believe that LCM
36
37 is appropriate and want to estimate lagged effects. However, LCMs are potentially problematic
38
39 because they ignore occasion effects, which can drastically impact the lagged effects estimated in
40
41 an ALT. Also, in addition to being difficult to estimate in some cases, the ALT excludes MA and
42
43 CLMA terms that increase the range of dynamic processes that the model can incorporate.
44
45
46
47
48

49 Discussion

50
51 We have compared various static and dynamic panel data models to our GCLM. The
52
53 static models we treat are random- and fixed-effect models and LCMs, which offer no clear path
54
55 to causal inference as a temporal process, with LCMs assuming all trends are systematic rather
56
57
58
59
60

1
2
3 than having elements of randomness. The dynamic models have other problems: typical cross-
4
5 lagged models fail to control for η_i , and group-mean centering to account for η_i induces
6
7 dynamic panel bias. Alternatively, Arellano-Bond methods with instruments and GMM may be
8
9 useful in some cases, but they lead to dilemmas that SEM overcomes. Finally, the ALT combines
10
11 LCM and lagged terms, which can overfit systematic trends, and even in the simplest case forces
12
13 occasion effects to take part in these trends. In all cases, MA, CLMA, and time-varying unit
14
15 effects are typically not specified in these models, giving the GCLM an advantage for estimating
16
17 effects that do not follow the structure implied by AR and CL effects, such as small but
18
19 persistent lagged effects or large but temporary lagged effects.
20
21
22
23

24 Although we do not compare all panel data models that appear in organization science,
25
26 other models can often be understood in ways that are consistent with the kinds of comparisons
27
28 we draw above (e.g., Chow, Ho, Hamaker, & Dolan, 2010; Hamaker & Dolan, 2009; Hamaker,
29
30 Nesselroade, & Molenaar, 2007; Nesselroade, McArdle, Aggen, & Meyers, 2002). For example,
31
32 latent change models and related approaches merely estimate effects among variables in
33
34 differences (e.g., Box et al., 2008; McArdle, 2001, 2009; McArdle & Hamagami, 2001, 2004).
35
36 These models can be reparameterized as cross-lagged SEMs or panel vector autoregressive
37
38 models (see Allen & Fildes, 2001; Arellano, 2003; Bai, 2013; Baltagi, 2013a, 2013b; McArdle &
39
40 Nesselroade, 2014; Moral-Benito, 2013; Usami, Hayes, & McArdle, 2015; Voelkle & Oud,
41
42 2015). Thus, such models will encounter the same issues we describe when they do not properly
43
44 control for η_i or when they overlook the possibility of meaningful MA and CLMA effects.
45
46
47
48
49

50 When considering such parameters, as we have noted, a process should be mapped onto a
51
52 statistical model using theory and previous findings, as well as substantive and statistical
53
54 checking. For this purpose, not all of the terms that the GCLM includes need to be specified, but
55
56
57
58
59
60

1
2
3 researchers should know that they are available in SEM if they are deemed to be of interest.
4
5 Indeed, past research has recognized all of the terms included in the GCLM in various ways
6
7 (e.g., du Toit & Browne, 2001, 2007; Hamaker, Dolan, & Molenaar, 2002; Hamaker & Grasman,
8
9 2015; Hamaker et al., 2015), and therefore a GCLM can be seen as bringing these terms together
10
11 in a coherent and easy-to-implement SEM framework—facilitated by our online supplemental
12
13 material that allows automatically generating Mplus program code using an Excel file.
14
15

16
17 This said, especially the comparison of our model with the ALT brings up important
18
19 issues regarding competition, so to speak, among unit effects and the AR, MA, CL, and CLMA
20
21 terms that we use for causal inference. The problem is that unit effects can be parameterized in a
22
23 wide variety of ways (just as AR, MA, CL, and CLMA terms can be), and these specifications
24
25 will produce different kinds of competition among parameters to explain auto- and cross-
26
27 covariation (as illustrated by the ALT models we estimate). This issue, in a very general sense,
28
29 was first discussed by “Student” (1914), who treated time as a predictor in order to detrend
30
31 longitudinal data (for historical developments, see Hooker, 1905; Yule, 1921, 1926; Tintner,
32
33 1940). As Yule (1921) noted in commenting on Student’s approach, the problem was that:
34
35

36
37 *“if ‘Student’ [1914] desires to remove from his figures secular movements, periodic*
38
39 *movements, uniform movements, and accelerated movements—well the reader is left*
40
41 *wondering with what sort of movements he does desire to deal... He desires to find the*
42
43 *correlation between x and y when every component in each of the variables is eliminated*
44
45 *which can well be called a function of the time, and nothing is left but residuals such that*
46
47 *the residual of a given year is uncorrelated with those that precede or that follow it...*
48
49 *[However], [t]he only residuals which it is easy to conceive as being totally uncorrelated*
50
51 *with one another in the manner supposed are errors of observation”* (p. 502-504).
52
53
54
55
56
57
58
59
60

1
2
3 The fact is that any part of a process can be theorized either as a systematic trend of some
4 duration that should be removed or as an AR, MA, CL, or CLMA effect that can have a causal
5 interpretation. Unfortunately, there is no single best way to differentiate trends from such causal
6 effects across all contexts, and attempts to do so often amount to exercises in aesthetic judgment
7 (Heckman, 1991)—consider ‘smoothing’ techniques that are often done in relation to graphical
8 figures. The problem is that “[t]ime series variables can, without further restrictions, be thought
9 of as composed of a part with a variable trend, plus a part that is not the trend. But a moment’s
10 reflection indicates that this decompositions lacks content... one economist’s ‘trend’ can be
11 another’s ‘cycle’” (Stock & Watson, 1988, p. 150). The fact is that all real-world data are messy
12 and never offer certainty regarding how a model should be constructed or effects conceptualized.
13
14
15
16
17
18
19
20
21
22
23
24
25

26 Again, Yule pointed this out long ago, noting that, “it is not my view alone but the view
27 of most writers on the subject up to 1914, that the essential difficulty of the time-correlation
28 problem is the difficulty of isolating for study different components in the total movement of
29 each variable” (1921, p. 501). Unfortunately, no single solution to this problem exists—or can
30 exist—that is applicable to all research contexts. Given the uncertainties here, we recommend
31 including a single time-varying unit effect term $\lambda_i \eta_i$, but other specifications are possible, such
32 as the ALTs that we estimate, which can fundamentally alter the lagged effects used for causal
33 inference. Therefore, we suggest caution when specifying and trying to interpret trends, and note
34 that researchers may wish to conduct sensitivity analyses that look at AR, MA, CL, and CLMA
35 effects under different parameterizations. As we have noted, we also encourage relying on theory
36 to decide how to specify a parsimonious model that closely targets hypotheses of interest.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 In conclusion, as we noted in our first paper, panel data models are not a panacea for
53 unconditional causal inference, just as randomized controlled trials are not (Cartwright & Hardie,
54
55
56
57
58
59
60

1
2
3 2012). From a practical perspective, causal inference under any approach is meant to allow using
4
5 past observations to plan and execute actions such as interventions or policy changes that are
6
7 designed to work for a set of specific purposes (Heckman, 2003, 2005). This practical orientation
8
9 should be kept in mind when both accounting for trends with unit effects and interpreting AR,
10
11 MA, CL, and CLMA terms along with their Granger-Sims and impulse-response counterparts.
12
13
14 No empirical method secures the future against uncertainty, but panel models like the GCLM can
15
16 be a useful complement to other methods for making plans and acting in the face of uncertainty.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Ahn, S. C., Lee, Y. H., & Schmidt, P. (2001). GMM estimation of linear panel data models with time-varying individual effects. *Journal of Econometrics*, *101*(2), 219-255. DOI: 10.1016/S0304-4076(00)00083-X
- Ahn, S. C., Lee, Y. H., & Schmidt, P. (2013). Panel data models with multiple time-varying individual effects. *Journal of Econometrics*, *174*(1), 1-14. DOI: 10.1016/j.jeconom.2012.12.002
- Allen, P. G., & Fildes, R. (2001). Econometric forecasting. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers.
- Allison, P. D. (2009). *Fixed effects regression models*. Thousand Oaks, CA: Sage Publications.
- Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum likelihood for cross-lagged panel models with fixed effects. *Socius*, *3*, 1-17. DOI: 10.1177/2378023117710578
- Alvarez, J., & Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica*, *71*(4), 1121-1159. DOI: 10.1111/1468-0262.00441
- Arellano, M. (2003). *Panel data econometrics*. Oxford: Oxford University Press.
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, *58*(2), 277-297. DOI: 10.2307/2297968
- Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29-51. DOI: 10.1016/0304-4076(94)01642-D
- Armstrong, J. S., Green, K. C., & Graefe, A. (2015). Golden rule of forecasting: Be conservative. *Journal of Business Research*, *68*(8), 1717-1731. DOI: 10.1016/j.jbusres.2015.03.031
- Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling*, *25*, 359-388. DOI: 10.1080/10705511.2017.1406803
- Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, *1*(1), 45-87. DOI: 10.1177/109442819800100104
- Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, *77*(4), 1229-1279. DOI: 10.3982/ECTA6135
- Bai, J. (2013). Likelihood approach to dynamic panel models with interactive effects. MRPA Paper No. 50267. Retrieved from: <http://mrpa.ub.uni-muenchen.de/50267/> DOI: 10.2139/ssrn.2332992
- Baltagi, B. H. (2013a). Dynamic panel data models. In N. Hashimzade & M. Thornton (Eds.), *Handbook of research methods and applications on empirical macroeconomics* (pp. 229-248). Cheltenham: Edward Elgar.
- Baltagi, B. (2013b). *Econometric analysis of panel data*. John Wiley & Sons.
- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1988). *Life span developmental psychology: Introduction to research methods*. Hillsdale, NJ: Erlbaum.
- Barkema, H. G., & Schijven, M. (2008). Toward unlocking the full potential of acquisitions: The role of organizational restructuring. *Academy of Management Journal*, *51*(4), 696-722. DOI: 10.5465/amr.2008.33665204
- 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*(1), 383-407. DOI: 10.1146/annurev-orgpsych-032414-111335

- 1
2
3 Beal, D. J., Trougakos, J. P., Weiss, H. M., & Green, S. G. (2006). Episodic processes in
4 emotional labor: Perceptions of affective delivery and regulation strategies. *The Journal of*
5 *Applied Psychology*, *91*(5), 1053-1065.
6
7 Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in
8 organizational research. *Organizational Research Methods*, *6*(4), 440-464. DOI:
9 10.1177/1094428103257361
10 Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations.
11 *Academy of Management Journal*, *42*(1), 25-40. DOI: 10.5465/256872
12 Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel
13 data models. *Journal of Econometrics*, *87*(1), 115-143. DOI: 10.1016/S0304-
14 4076(98)00009-8
15 Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary*
16 *and experience sampling research*. New York: Guilford Press.
17 Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A
18 structural equations approach. *Social Forces*, *89*(1), 1-34. DOI: 10.1353/sof.2010.0072
19 Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis
20 of two traditions. *Sociological Methods & Research*, *32*(3), 336-383. DOI:
21 10.1177/0049124103260222
22 Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*.
23 New York: Wiley.
24 Bollen, K. A., & Zimmer, C. R. (2010). An overview of the autoregressive latent trajectory
25 (ALT) model. In K. van Montfort, J. H. L. Oud, & A. Satorra (Eds.), *Longitudinal research*
26 *with latent variables* (pp. 153-176). Berlin: Springer-Verlag.
27 Bond, S. R. (2002). Dynamic panel data models: A guide to micro data methods and practice.
28 *Portuguese Economic Journal*, *1*(2), 141-162. DOI: 10.1007/s10258-002-0009-9
29 Bono, J. E., Foldes, H. J., Vinson, G., & Muros, J. P. (2007). Workplace Emotions: The Role of
30 Supervision and Leadership. *Journal of Applied Psychology*, *92*(5), 1357-1367.
31 Boumahdi, R., & Thomas, A. (2008). Endogenous regressors and correlated effects. In L. Mátyás
32 & P. Sevestre (Eds.), *The econometrics of panel data: Fundamentals and recent*
33 *developments in theory and practice* (pp. 89-112). Berlin: Springer-Verlag.
34 Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis: Forecasting and*
35 *control* (4th ed.). New York: Wiley.
36 Braun, M. T., Kuljanin, G., & DeShon, R. P. (2013). Spurious results in the analysis of
37 longitudinal data in organizational research. *Organizational Research Methods*, *16*(2), 302-
38 330. DOI: 10.1177/1094428112469668
39 Bun, M., & Sarafidis, V. (2015). Dynamic panel data models. In B. Baltagi (Ed.), *Oxford*
40 *handbook of panel data* (pp. 76-110). Oxford: Oxford University Press.
41 Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*.
42 Oxford: Oxford University Press.
43 Chow, S.-M., Ho, M.-h. R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and differences
44 between structural equation modeling and state-space modeling techniques. *Structural*
45 *Equation Modeling*, *17*, 303-332. DOI:10.1080/10705511003661553
46 Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An
47 explanation for the Easterlin paradox and other puzzles. *Journal of Economic literature*,
48 *46*(1), 95-144. DOI: 10.1257/jel.46.1.95
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data:
4 Questions and tips in the use of structural equation modeling. *Journal of Abnormal*
5 *Psychology, 112*(4), 558-577.
- 6
7 Curran, P. J. (2003). Have multilevel models been structural equation models all along?
8 *Multivariate Behavioral Research, 38*(4), 529-569. DOI: 10.1207/s15327906mbr3804_5
- 9
10 Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person
11 effects in longitudinal models of change. *Annual Review of Psychology, 62*, 583-619. DOI:
12 10.1146/annurev.psych.093008.100356
- 13
14 Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and
15 latent curve models. In L. Collins & A. G. Sayer (Eds.), *New methods for the analysis of*
16 *change* (pp. 107-135). Washington, DC: American Psychological Association.
- 17
18 Curran, P. J., Lee, T., Howard, A. L., Lane, S., & MacCallum, R. A. (2012). Disaggregating
19 within-person and between-person effects in multilevel and structural equation growth
20 models. In J. R. Harring & G. R. Hancock (Eds.), *Advances in longitudinal methods in the*
21 *social and behavioral sciences* (pp. 217-253). Charlotte, NC: Information Age.
- 22
23 Dalal, R. S., Lam, H., Weiss, H. M., Welch, E. R., & Hulin, C. L. (2009). A within-person
24 approach to work behavior and performance: Concurrent and lagged citizenship-
25 counterproductivity associations, and dynamic relationships with affect and overall job
26 performance. *Academy of Management Journal, 52*(5), 1051-1066. DOI:
27 10.5465/amj.2009.44636148
- 28
29 Davidson, D. (1986). A coherence theory of knowledge and truth. In E. LePore (Ed.), *Truth and*
30 *interpretation* (pp. 307-319). Oxford: Blackwell.
- 31
32 Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic*
33 *Literature, 41*(1), 113-158. DOI: 0.1257/002205103321544710
- 34
35 Diener, E., & Lucas, R. E. (1999). Personality and subjective well-being. In D. Kahneman, E.
36 Diener, & N. Schwarz (Eds.), *Well-being: Foundations of hedonic psychology* (pp. 213-
37 229). New York: Sage.
- 38
39 Diener, E., Tay, L., & Oishi, S. (2013). Rising income and the subjective well-being of nations.
40 *Journal Of Personality & Social Psychology, 104*(2), 267-276.
- 41
42 Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using longitudinal
43 hierarchical linear modeling: An illustration predicting achievement from self-control.
44 *Social Psychological and Personality Science, 1*(4), 311-317. DOI:
45 10.1177/1948550609359707
- 46
47 du Toit, S. H. C., & Browne, M. W. (2001). The covariance structure of a vector ARMA time
48 series. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), *Structural equation modeling:*
49 *Present and future. A festschrift in honor of Karl Jöreskog* (pp. 279-314). Chicago, IL:
50 Scientific Software International.
- 51
52 du Toit, S. H. C., & Browne, M. W. (2007). Structural equation modeling of multivariate time
53 series. *Multivariate Behavioral Research, 42*(1), 67-101. DOI:
54 10.1080/00273170701340953
- 55
56 Eby, L. T., Butts, M. M., Hoffman, B. J., Sauer, J. B. (2015). Cross-lagged relations between
57 mentoring received from supervisors and employee OCBs: Disentangling causal direction
58 and identifying boundary conditions. *Journal of Applied Psychology, 100*(4), 1275-1285.
59 DOI: 10.1037/a0038628
- 60
61 Edwards, J. R. (2011). The fallacy of formative measurement. *Organizational Research*
62 *Methods, 14*(2), 370-388. DOI: 10.1177/1094428110378369

- 1
2
3 Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel
4 models: A new look at an old issue. *Psychological Methods, 12*(2), 121-138. DOI:
5 10.1037/1082-989X.12.2.121
6
7 Fisher, C. D., & Noble, C. S. (2004). A within-person examination of correlates of performance
8 and emotions while working. *Human Performance, 17*(2), 145-168. DOI:
9 10.1207/s15327043hup1702_2
10
11 Foster, E. M. (2010). Causal inference and developmental psychology. *Developmental*
12 *Psychology, 46*(6), 1454.
13
14 George, G. (2005). Slack resources and the performance of privately held firms. *Academy of*
15 *Management Journal, 48*(4), 661-676. DOI: 10.5465/amj.2005.17843944
16
17 Geweke, J. F. (1982). Measurement of linear dependence and feedback between multiple time
18 series. *Journal of the American Statistical Association, 77*(378), 304-324. DOI:
19 10.1080/01621459.1982.10477803
20
21 Gielnik, M. M., Spitzmuller, M., Schmitt, A., Klemann, D. K., & Frese, M. (2015). “I put in
22 effort, therefore I am passionate”: Investigating the path from effort to passion in
23 entrepreneurship. *Academy of Management Journal, 58*(4), 1012-1031. DOI:
24 10.5465/amj.2011.0727
25
26 Goldstein, A. (2012). Revenge of the managers labor cost-cutting and the paradoxical resurgence
27 of managerialism in the shareholder value era, 1984 to 2001. *American Sociological*
28 *Review, 77*(2), 268-294. DOI: 10.1177/0003122412440093
29
30 Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child*
31 *Development, 58*, 80-92.
32
33 Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-
34 spectral method. *Econometrica, 3*, 424-438. DOI: 10.2307/1912791
35
36 Granger, C. W. J. (1980). Testing for causality, a personal viewpoint. *Journal of Economic*
37 *Dynamics and Control, 2*, 329-352. DOI: 10.1016/0165-1889(80)90069-X
38
39 Green, K. C., & Armstrong, J. S. (2015). Simple versus complex forecasting: The evidence.
40 *Journal of Business Research, 68*(8), 1678-1685. DOI: 10.1016/j.jbusres.2015.03.026
41
42 Griffin, M. A. (1997). Interaction between individuals and situations: Using HLM procedures to
43 estimate reciprocal relationships. *Journal of Management, 23*(6), 759-773. DOI:
44 10.1177/014920639702300604
45
46 Gulati, R. (1995). Social structure and alliance formation patterns: A longitudinal analysis.
47 *Administrative Science Quarterly, 40*, 619-652. DOI: 10.2307/2393756
48
49 Halaby, C. N. (2004). Panel models in sociological research: Theory into practice. *Annual*
50 *Review of Sociology, 30*, 507-544. DOI: 10.1146/annurev.soc.30.012703.110629
51
52 Hamaker, E. L. (2005). Conditions for the equivalence of the autoregressive latent trajectory
53 model and a latent growth curve model with autoregressive disturbances. *Sociological*
54 *Methods & Research, 33*, 404-416. DOI: 10.1177/0049124104270220
55
56 Hamaker, E. L., & Dolan, C. V. (2009). Idiographic data analysis: Quantitative methods—from
57 simple to advanced. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra, & N. Chaudhary
58 (Eds.), *Dynamic process methodology in the social and developmental sciences* (pp. 191-
59 216). New York: Springer.
60
61 Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2002). On the nature of SEM estimates of
62 ARMA parameters. *Structural Equation Modeling, 9*, 347-368. DOI:
63 10.1207/S15328007SEM0903_3
64
65
66
67
68
69
70

- 1
2
3 Hamaker, E. L. & Grasman, R. P. (2015). To center or not to center? Investigating inertia with a
4 multilevel autoregressive model. *Frontiers in Psychology*, 5, 1492. DOI:
5 10.3389/fpsyg.2014.01492
6
7 Hamaker, E. L., Kuiper, R. M., Grasman, R. P. (2015). A critique of the cross-lagged panel
8 model. *Psychological Methods*, 20(1), 102-116. DOI: 10.1037/a0038889
9
10 Hamaker, E. L., Nesselroade, J. R., & Molenaar, P. C. (2007). The integrated trait–state model.
11 *Journal of Research in Personality*, 41(2), 295-315. DOI: 10.1016/j.jrp.2006.04.003
12
13 Harvey, A. C. (1997). Trends, cycles and autoregressions. *The Economic Journal*, 107(444),
14 192-201.
15
16 Heckman, J. (1991). Identifying the hand of past: Distinguishing state dependence from
17 heterogeneity. *The American Economic Review*, 81(2), 75-79.
18
19 Heckman, J. (2003). Conditioning, causality and policy analysis. *Journal of Econometrics*,
20 112(1), 73-78.
21
22 Heckman, J. J. (2005). Rejoinder: Response to Sobel. *Sociological Methodology*, 35(1), 135-162.
23 DOI: 10.1111/j.0081-1750.2006.00166.x
24
25 Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuations and change*.
26 New York: Routledge.
27
28 Hoffman, L., & Stawski, R. S. (2009). Persons as contexts: Evaluating between-person and
29 within-person effects in longitudinal analysis. *Research in Human Development*, 6(2-3),
30 97-120. DOI: 10.1080/15427600902911189
31
32 Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models.
33 *Journal of Management*, 23(6), 723-744. DOI: 10.1177/014920639702300602
34
35 Holtz-Eakin, D., Newey, W., & Rosen, R. (1988). Estimating vector autoregressions with panel
36 data. *Econometrica*, 56, 1371-1395.
37
38 Hooker, R. H. (1905). On the correlation of successive observations illustrated by corn-prices.
39 *Journal of the Royal Statistical Society*, 68(4), 696-703. DOI: 10.2307/2339422
40
41 Hsiao, C. (2014). *Analysis of panel data* (3rd ed.). Cambridge: Cambridge University Press.
42
43 Hsiao, C., Pesaran, M. H., & Tahmiscioglu, A. K. (2002). Maximum likelihood estimation of
44 fixed effects dynamic panel data models covering short time periods. *Journal of*
45 *Econometrics*, 109(1), 107-150. DOI: 10.1016/S0304-4076(01)00143-9
46
47 Ilies, R., Johnson, M. D., Judge, T. A., & Keeney, J. (2011). A within-individual study of
48 interpersonal conflict as a work stressor: Dispositional and situational moderators. *Journal*
49 *of Organizational Behavior*, 32(1), 44-64. DOI: 10.1002/job.677
50
51 Ilies, R., Scott, B. A., & Judge, T. A. (2006). The interactive effects of personal traits and
52 experienced states on intraindividual patterns of citizenship behavior. *Academy of*
53 *Management Journal*, 49(3), 561-575. DOI: 10.5465/amj.2006.21794672
54
55 Jongerling, J., Laurenceau, J. P., & Hamaker, E. L. (2015). A multilevel AR(1) model: Allowing
56 for inter-individual differences in trait-scores, inertia, and innovation variance.
57 *Multivariate Behavioral Research*, 50(3), 334-349. DOI: 10.1080/00273171.2014.1003772
58
59 Judge, T. A., & Ilies, R. (2004). Affect and job satisfaction: A study of their relationship at work
60 and at home. *Journal of Applied Psychology*, 89(4), 661-673.
61
62 Judge, T. A., Scott, B. A., & Ilies, R. (2006). Hostility, job attitudes, and workplace deviance:
63 test of a multilevel model. *Journal of Applied Psychology*, 91(1), 126-138.
64
65 Judson, R. A., & Owen, A. L. (1999). Estimating dynamic panel data models: A guide for
66 macroeconomists. *Economics Letters*, 65(1), 9-15.

- 1
2
3 Kling, G., Harvey, C., & Maclean, M. (2017). Establishing causal order in longitudinal studies
4 combining binary and continuous dependent variables. *Organizational Research Methods*,
5 20(4), 770-799. DOI: 10.1177/1094428115618760
6
7 Kuljanin, G., Braun, M. T., & DeShon, R. P. (2011). A cautionary note on modeling growth
8 trends in longitudinal data. *Psychological Methods*, 16(3), 249-264.
9
10 Lehrer, K. (2000). *Theory of knowledge*. New York: Routledge.
11
12 Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford Press.
13
14 Locke, K., & Golden-Biddle, K. (1997). Constructing opportunities for contribution: Structuring
15 intertextual coherence and “problematizing” in organizational studies. *Academy of*
16 *Management Journal*, 40(5), 1023-1062. DOI: 10.5465/256926
17
18 Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation.
19 *Psychological Methods*, 12(1), 23-44. DOI: 10.1037/1082-989X.12.1.23
20
21 Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of
22 longitudinal mediation: Partial and complete mediation under an autoregressive model.
23 *Multivariate Behavioral Research*, 46(5), 816-841. DOI: 10.1080/00273171.2011.606716
24
25 Maynard, M. T., Luciano, M. M., D’Innocenzo, L., Mathieu, J. E., & Dean, M. D. (2014).
26 Modeling time-lagged reciprocal psychological empowerment–performance relationships.
27 *Journal of Applied Psychology*, 99(6), 1244.
28
29 McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural
30 analysis. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), *Structural equation modeling:*
31 *Present and future* (pp. 342-380). Lincolnwood, IL: Scientific Software International.
32
33 McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal
34 data. *Annual Review of Psychology*, 60, 577-605.
35
36 McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear
37 dynamic analyses with incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *New*
38 *methods for the analysis of change* (pp. 139–175). Washington, DC: APA Press.
39
40 McArdle, J. J., & Hamagami, F. (2004). Methods for dynamic change hypotheses. In K. van
41 Montfort, J. H. L. Oud, & A. Satorra (Eds.), *Recent developments on structural equation*
42 *models* (pp. 295-335). Dordrecht: Kluwer.
43
44 McArdle, J. J., & Nesselroade, J. R. (2003). Growth curve analysis in contemporary
45 psychological research. In J. Schinka & W. Velicer, (Eds.), *Comprehensive handbook of*
46 *psychology* (Vol. 2, pp. 447-480). New York: Pergamon Press.
47
48 McArdle, J. J., & Nesselroade, J. R. (2014). *Longitudinal data analysis using structural equation*
49 *models*. Washington, DC: American Psychological Association.
50
51 Meier, L. L., & Spector, P. E. (2013). Reciprocal effects of work stressors and counterproductive
52 work behavior: A five-wave longitudinal study. *Journal of Applied Psychology*, 98(3),
53 529-539. DOI: 10.1037/a0031732
54
55 Moral-Benito, E. (2013) Likelihood-based estimation of dynamic panels with predetermined
56 regressors. *Journal of Business & Economic Statistics*, 31(4), 451-472. DOI:
57 10.1080/07350015.2013.818003
58
59 Morgan, S. L., & Winship, C. (2014). *Counterfactuals and causal inference*. Cambridge:
60 Cambridge University Press.
61
62 Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible
63 representation of substantive theory. *Psychological Methods*, 17(3), 313.
64
65
66
67
68
69
70

- 1
2
3 Nauges, C., & Thomas, A. (2003). Consistent estimation of dynamic panel data models with
4 time-varying individual effects. *Annales d'Economie et de Statistique*, 70, 53-75. DOI :
5 10.2307/20076374
- 6 Nelson, C. R., & Kang, H. (1981). Spurious periodicity in inappropriately detrended time series.
7 *Econometrica*, 49(3), 741-751. DOI: 10.2307/1911520
- 8 Nelson, C. R., & Kang, H. (1984) Pitfalls in the use of time as an explanatory variable in
9 regression. *Journal of Business & Economic Statistics*, 2, 73-82. DOI:
10 10.1080/07350015.1984.10509371
- 11 Nelson, C. R., & Plosser, C. R. (1982). Trends and random walks in macroeconomic time series:
12 Some evidence and implications. *Journal of Monetary Economics*, 10, 139-162. DOI:
13 10.1016/0304-3932(82)90012-5
- 14 Nerlove, M. (1967). Experimental evidence on the estimation of dynamic economic relations
15 from a time series of cross-sections. *Economic Studies Quarterly*, 18, 42-74. DOI:
16 10.11398/economics1950.18.3_42
- 17 Nerlove, M. (1971). Further evidence on the estimation of dynamic economic relations from a
18 time series of cross sections. *Econometrica*, 39(2), 359-387. DOI: 10.2307/1913350
- 19 Nerlove, M., Sevestre, P., & Balestra, P. (2008). Introduction. In L. Mátyás & P. Sevestre (Eds.),
20 *The econometrics of panel data: Fundamentals and recent developments in theory and*
21 *practice* (pp. 3-22). Berlin: Springer-Verlag.
- 22 Nesselrode, J. J., McArdle, J. J., Aggen, S. H., & Meyers, J. (2002). Dynamic factor analysis
23 models for multivariate time series analysis. In D. M. Moskowitz & S. L. Hershberger
24 (Eds.), *Modeling individual variability with repeated measures data: Advances and*
25 *techniques* (pp. 233-266). Mahwah, NJ: Erlbaum.
- 26 Nezlek, J. B. (2001). Multilevel random coefficient analyses of event-and interval-contingent
27 data in social and personality psychology research. *Personality and Social Psychology*
28 *Bulletin*, 27(7), 771-785. DOI: 10.1177/0146167201277001
- 29 Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality
30 psychology. *Social and Personality Psychology Compass*, 2(2), 842-860. DOI:
31 10.1111/j.1751-9004.2007.00059.x
- 32 Nezlek, J. B. (2011). *Multilevel modeling for social and personality psychology*. London: Sage.
- 33 Nezlek, J. B. (2012a). *Diary methods for social and personality psychology*. London: Sage.
- 34 Nezlek, J. B. (2012b). Multilevel modeling of diary-style data. In M. R. Mehl & T. S. Conner
35 (Eds.) *Handbook of research methods for studying daily life* (pp. 357-383). New York:
36 Guilford Press.
- 37 Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49, 1417-1426.
38 DOI: 10.2307/1911408
- 39 Nohe, C., Meier, L. L., Sonntag, K., & Michel, A. (2015). The chicken or the egg? A meta-
40 analysis of panel studies of the relationship between work-family conflict and strain.
41 *Journal of Applied Psychology*, 100(2), 522-536.
- 42 Piening, E. P., Baluch, A. M., & Salge, T. O. (2013). The relationship between employees'
43 perceptions of human resource systems and organizational performance: Examining
44 mediating mechanisms and temporal dynamics. *Journal of Applied Psychology*, 98(6), 926.
- 45 Pierce, D. A., & Haugh, L. D. (1977). Causality in temporal systems: Characterization and a
46 survey. *Journal of Econometrics*, 5(3), 265-293. DOI: 10.1016/0304-4076(77)90039-2
- 47
48
49
50
51
52
53
54
55
56
57
58
59
60

- Pitariu, A., & Ployhart, R. E. (2010). Explaining change: Theorizing and testing dynamic mediated longitudinal relationships. *Journal of Management*, *36*(2), 405-429. DOI: 10.1177/0149206308331096
- Quine, W., & Ullian, J. (1970). *The web of belief*. New York: Random House.
- Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*, *52*(1), 501-525. DOI: 0.1146/annurev.psych.52.1.501
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- Ricketta, M. (2008). The causal relation between job attitudes and performance: A meta-analysis of panel studies. *Journal of Applied Psychology*, *93*(2), 472-481. DOI: 10.1037/0021-9010.93.2.472
- Roodman, D. (2009a). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, *71*(1), 135-158. DOI: 10.1111/j.1468-0084.2008.00542.x
- Roodman, D. (2009b). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, *9*(1), 86-136. DOI: 10.2139/ssrn.982943
- Rovine, M. J., & Walls, T. A. (2006). Multilevel autoregressive modeling of interindividual differences in the stability of a process. In T. A. Walls & J. L. Schafer (Eds.), *Models for Intensive Longitudinal Data* (pp. 124-147). New York: Oxford University Press.
- Schonfeld, I. S., & Rindskopf, D. (2007). Hierarchical linear modeling in organizational research: Longitudinal data outside the context of growth modeling. *Organizational Research Methods*, *10*(3), 417-429. DOI: 10.1177/1094428107300229
- Sevestre, P., & Trognon, A. (1985). A note on autoregressive error components models. *Journal of Econometrics*, *28*(2), 231-245. DOI: 10.1016/0304-4076(85)90122-8
- Shepherd, D. A., & Sutcliffe, K. M. (2011). Inductive top-down theorizing: A source of new theories of organization. *Academy of Management Review*, *36*(2), 361-380. DOI: 10.5465/amr.2009.0157
- Sims, C. A. (1996). *Inference for multivariate time series models with trend*. Princeton, NJ: Princeton University.
- Sims, C. (2000). Using a likelihood perspective to sharpen econometrics discourse: Three examples. *Journal of Econometrics*, *95*(2), 443-462. DOI: 10.1016/S0304-4076(99)00046-9
- Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.
- Skrondal, A., & Rabe-Hesketh, S. (2008). Multilevel and related models for longitudinal data. In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 275-299). New York: Springer.
- Stock, J. H., & Watson, M. W. (1988). Variable trends in economic time series. *The Journal of Economic Perspectives*, *2*(3), 147-174. DOI: 10.1257/jep.2.3.147
- Stock, J. H., & Watson, M. W. (2003). *Introduction to econometrics* (Vol. 104). Boston, MA: Addison Wesley.
- Teachman, J., Duncan, G. J., Yeung, W. J., & Levy, D. (2001). Covariance structure models for fixed and random effects. *Sociological Methods & Research*, *30*(2), 271-288. DOI: 10.1177/0049124101030002005
- Tintner, G. (1940). *The variate difference method*. Bloomington, IN: Principia Press.

- 1
2
3 Usami, S., Hayes, T., & McArdle, J. J. (2015). On the mathematical relationship between latent
4 change score and autoregressive cross-lagged factor approaches: Cautions for inferring
5 causal relationship between variables. *Multivariate Behavioral Research*, *50*(6), 676-687.
6 DOI: 10.1080/00273171.2015.1079696
7
8 Uy, M. A., Foo, M. D., & Aguinis, H. (2010). Using experience sampling methodology to
9 advance entrepreneurship theory and research. *Organizational Research Methods*, *13*(1),
10 31-54. DOI: 10.1177/1094428109334977
11
12 Vaisey, S., & Miles, A. (2017). What you can—and can't—do with three-wave panel
13 data. *Sociological Methods & Research*, *46*(1), 44-67. DOI: 10.1177/0049124114547769
14
15 Van Iddekinge, C. H., Ferris, G. R., Perrewé, P. L., Perryman, A. A., Blass, F. R., & Heetderks,
16 T. D. (2009). Effects of selection and training on unit-level performance over time: A
17 latent growth modeling approach. *Journal of Applied Psychology*, *94*(4), 829-843. DOI:
18 10.1037/a0014453
19
20 Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation
21 and performance in a learning context. *The Journal of Applied Psychology*, *91*(5), 1146-
22 1153.
23
24 Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies
25 examining the negative effect of self-efficacy on performance. *The Journal of Applied*
26 *Psychology*, *87*(3), 506-516.
27
28 Vancouver, J. B., Thompson, C. M., & Williams, A. A. (2001). The changing signs in the
29 relationships among self-efficacy, personal goals, and performance. *Journal of Applied*
30 *Psychology*, *86*(4), 605.
31
32 Voelkle, M. C. (2008). Reconsidering the use of autoregressive latent trajectory (ALT) models.
33 *Multivariate Behavioral Research*, *43*, 564-591. DOI: 10.1080/00273170802490665
34
35 Voelkle, M. C., & Oud, J. H. L. (2015) Relating latent change score and continuous time models.
36 *Structural Equation Modeling*, *22*(3), 366-381. DOI: 10.1080/10705511.2014.935918
37
38 Walls, T. A., Jung, H., & Schwartz, J. E. (2006). Multilevel models for intensive longitudinal
39 data. In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 3-
40 37). New York: Oxford University Press.
41
42 Wang, L., & Maxwell, S. E. (2015). On disaggregating between-person and within-person
43 effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63.
44 DOI: 10.1037/met0000030
45
46 Watson, M. W. (1986). Univariate detrending methods with stochastic trends. *Journal of*
47 *Monetary Economics*, *18*, 49-75.
48
49 Williams, R., Allison, P. D., & Moral-Benito, E. (2018). Linear dynamic panel-data estimation
50 using maximum likelihood and structural equation modeling. *The Stata Journal*, *18*, 293-
51 326.
52
53 Yeo, G. B., & Neal, A. (2006). An examination of the dynamic relationship between self-
54 efficacy and performance across levels of analysis and levels of specificity. *Journal of*
55 *Applied Psychology*, *91*(5), 1088-1101.
56
57 Yule, G. U. (1921). On the time-correlation problem, with special reference to the variate-
58 difference correlation method. *Journal of the Royal Statistical Society*, *84*, 497-526. DOI:
59 10.2307/2341101
60
61 Yule, G. U. (1926). Why do we sometimes get nonsense correlations between time series? A
62 study in sampling and the nature of time series. *Journal of the Royal Statistical Society*,
63 *89*(1), 1-64. DOI: 10.2307/2341482

Table 1

Results for Dynamic Models

Parameter	Model									
	1	2.1	2.2	2.3	2.4	3.1	3.2	3.3	4.1	4.2
SWB→SWB AR/MA Terms $\beta_{y1}^{(y)}$ and $\delta_{y1}^{(y)}$										
$\beta_{y1}^{(y)}$.390	.877*	.856*	.031	.674*	.66*	.441*	.62*	.451*	.597*
$\delta_{y1}^{(y)}$.191	---	---	---	---	---	---	---	---	---
$\beta_{y1}^{(y)} + \delta_{y1}^{(y)}$.580*	.877*	.856*	.031	.674*	.66*	.441*	.62*	.451*	.597*
Income→Income AR/MA Terms $\beta_{x1}^{(x)}$ and $\delta_{x1}^{(x)}$										
$\beta_{x1}^{(x)}$.958*	.952*	.944*	.079	.583*	.982*	.458	.59*	.083	.561*
$\delta_{x1}^{(x)}$	-.326	---	---	---	---	---	---	---	---	---
$\delta_{x2}^{(x)}$.055	---	---	---	---	---	---	---	---	---
$\beta_{x1}^{(x)} + \delta_{x1}^{(x)}$.686*	.952*	.944*	.079	.583*	.982*	.458	.59*	.083*	.561*

Table 1 (continued)

Results for Dynamic Models

Parameter	Model									
	1	2.1	2.2	2.3	2.4	3.1	3.2	3.3	4.1	4.2
Income→SWB CL/CLMA Terms $\beta_{x1}^{(y)}$ and $\delta_{x1}^{(y)}$										
$\beta_{x1}^{(y)}$.131	.108*	.122*	.275*	.014	.184	.063	.076	.069	.318*
$\delta_{x1}^{(y)}$.009	---	---	---	---	---	---	---	---	---
$\beta_{x1}^{(y)} + \delta_{x1}^{(y)}$.139	.108*	.122*	.275*	.014	.184	.063	.076	.069	.318*
SWB→Income CL/CLMA Terms $\beta_{y1}^{(x)}$ and $\delta_{y1}^{(x)}$										
$\beta_{y1}^{(x)}$	-.103	.027*	.036*	.024	.004	.011	.138	.053	.019	.046
$\delta_{y1}^{(x)}$.080	---	---	---	---	---	---	---	---	---
$\beta_{y1}^{(x)} + \delta_{y1}^{(x)}$	-.023	.027*	.036*	.024	.004	.011	.138	.053	.019	.046

Note. Columns are named after the models described in the text as follows: Model 1 is our full AR(1)MA(2)CL(1)CLMA(1) GCLM; 2.1 is a cross-lagged SEM; 2.2 is a cross-lagged MLM; 2.3 is a group-mean or within-group centered cross-lagged model; 2.4 is a cross-lagged SEM with time-varying unit effects; 3.1 is a ‘system-GMM’ model in differences and levels; 3.2 is a difference-GMM;

1
2
3 *3.3 is an SEM version of a latent fixed-effects model (see Allison, 2014); 4.1 is an ALT with a linear trend to represent unit effects; 4.2*
4 *is an ALT with a single time-varying unit effect; SWB = subjective well-being; AR = autoregressive; MA = moving average; CL =*
5 *cross-lagged; CLMA = cross-lagged moving average.*
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

For Peer Review

RUNNING HEAD: From Data to Causes II

Figure 1

A Full GCLM with AR(1)MA(1)CL(1)CLMA(1) Effects

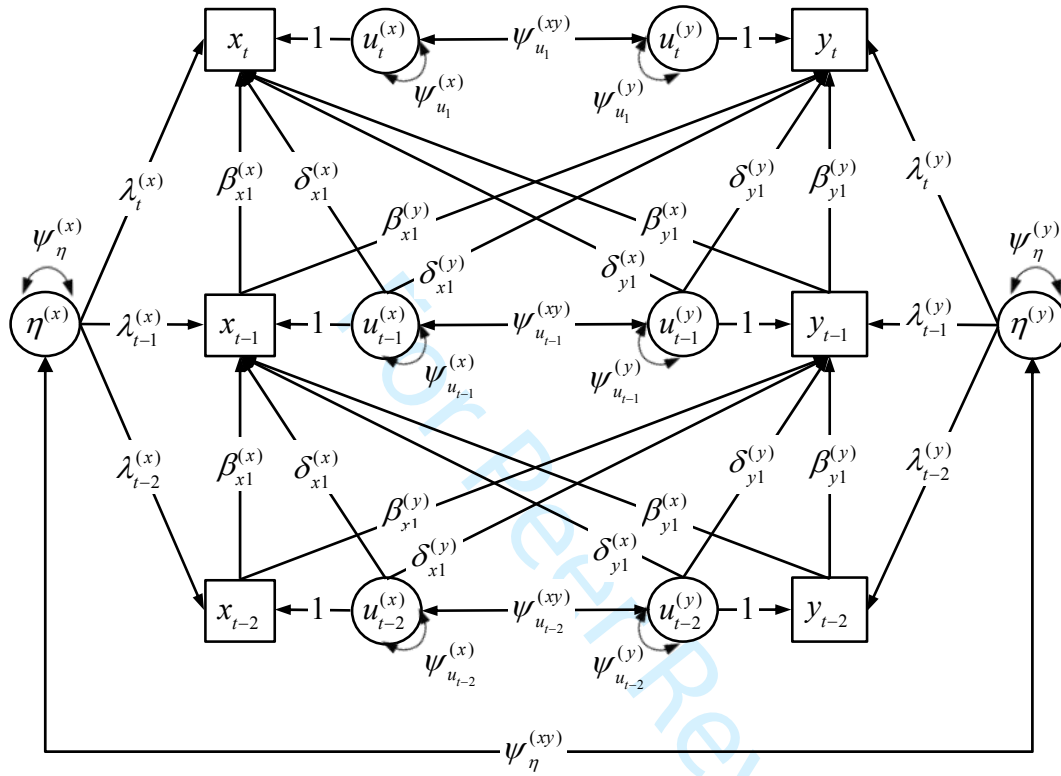
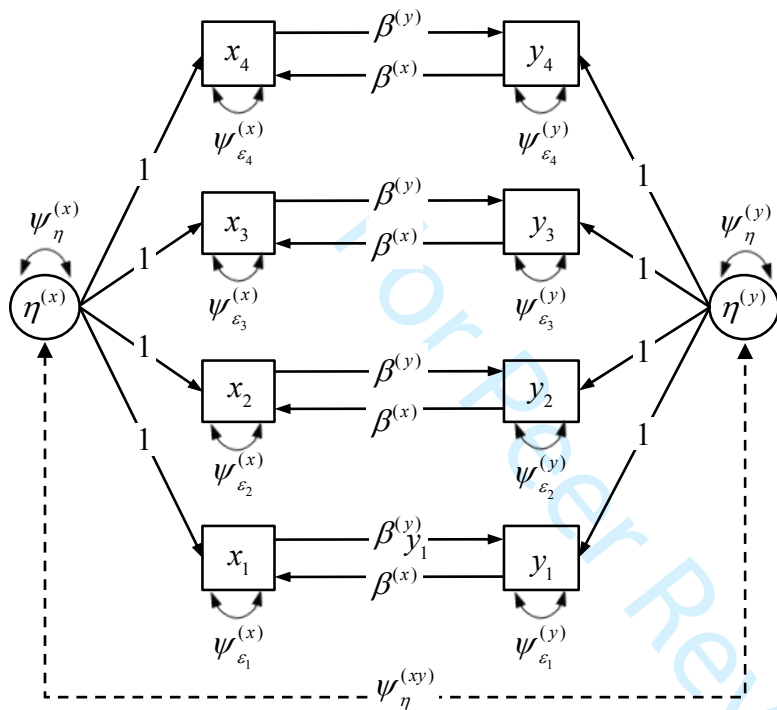


Figure 2

Random-Effects Model when Latent Unit Effect Covariance Restricted to Zero (Dashed Line),

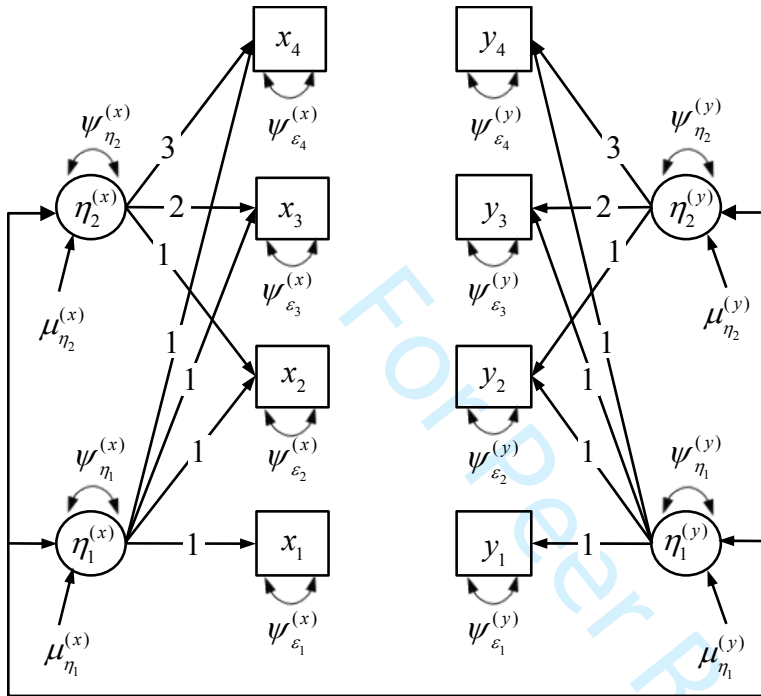
but Fixed-Effects Model Accounting for Unit Effects when Latent Covariance Estimated



RUNNING HEAD: From Data to Causes II

Figure 3

Latent Curve Model Showing Latent Means as μ (Excluding Covariance Labels for Concision)



RUNNING HEAD: From Data to Causes II

48

Figure 4a

Typical Cross-Lagged Model

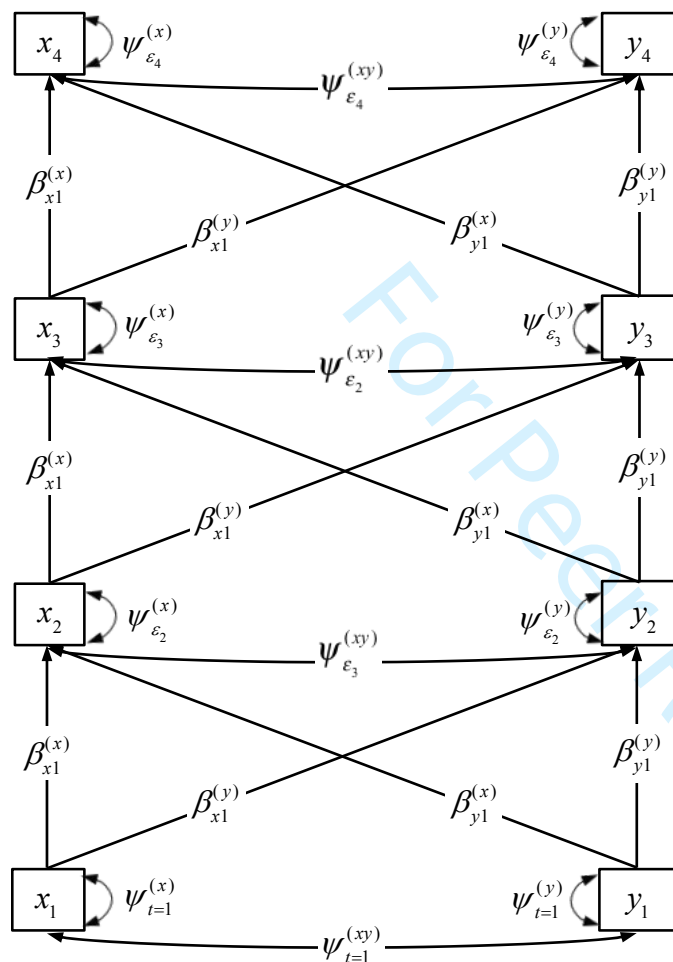


Figure 4b

Cross-Lagged Model Modified to Account for Unit Effects η

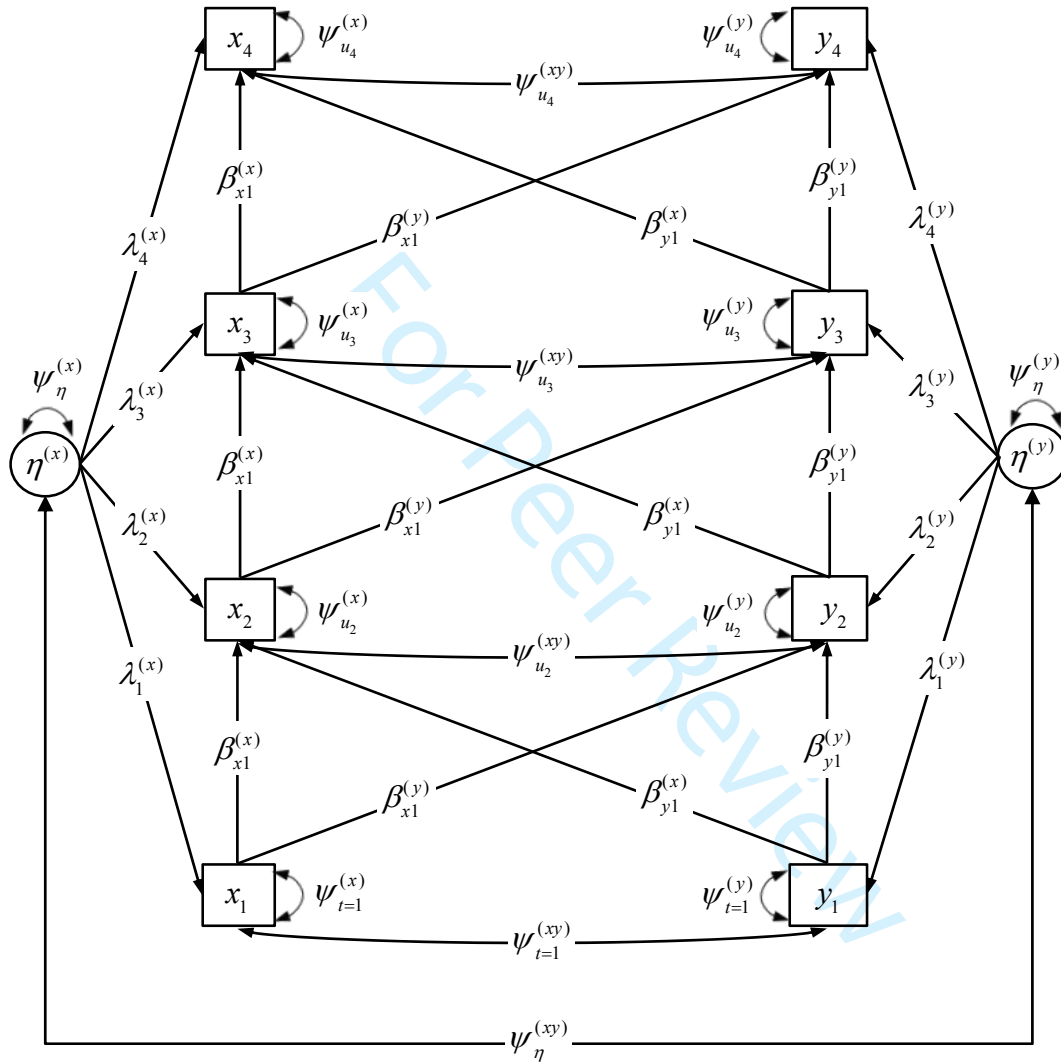


Figure 5

Alternative to Arellano-Bond Methods Using SEM (Showing an Income→SWB Model and Excluding Covariance Labels for Concision)

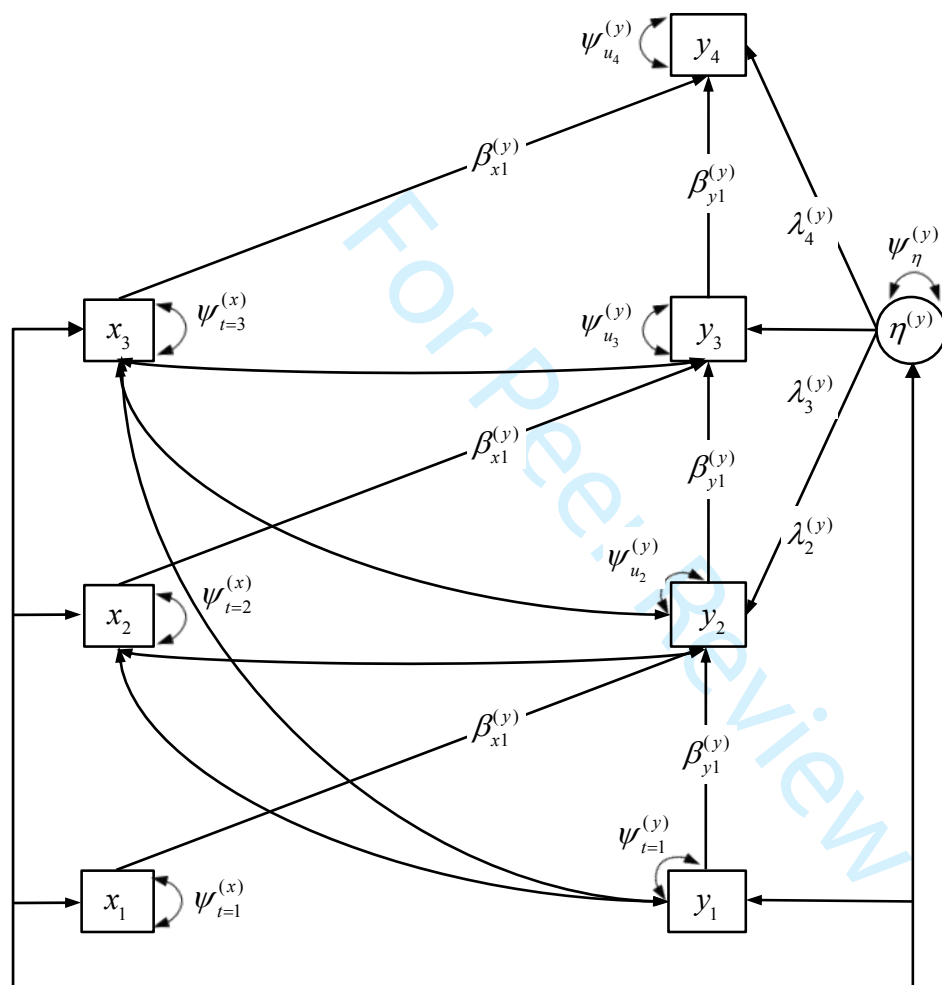


Figure 6a

Autoregressive Latent Trajectory Model with A Linear Trajectory and Showing Latent Means as μ Terms (Excluding Covariance Labels for Concision)

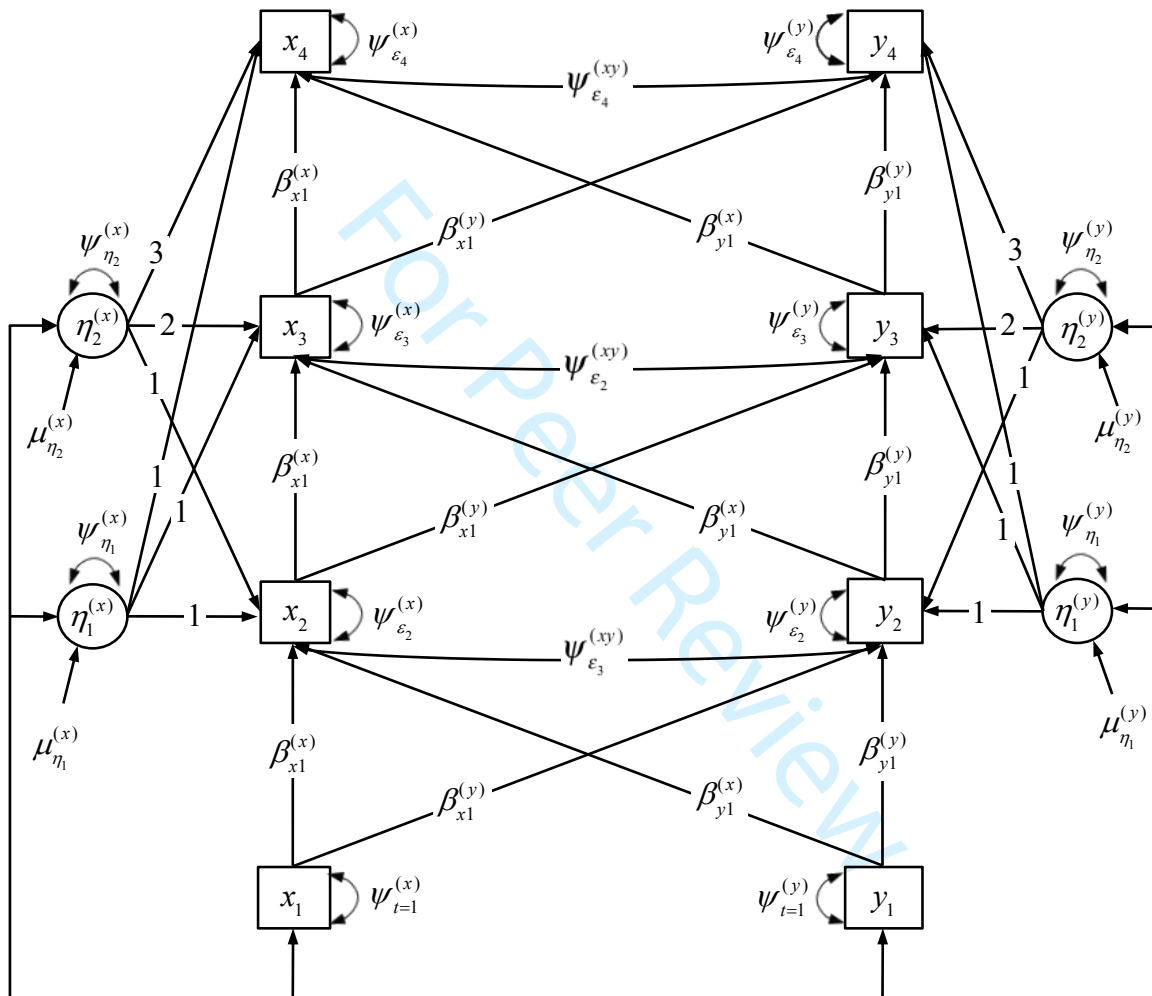
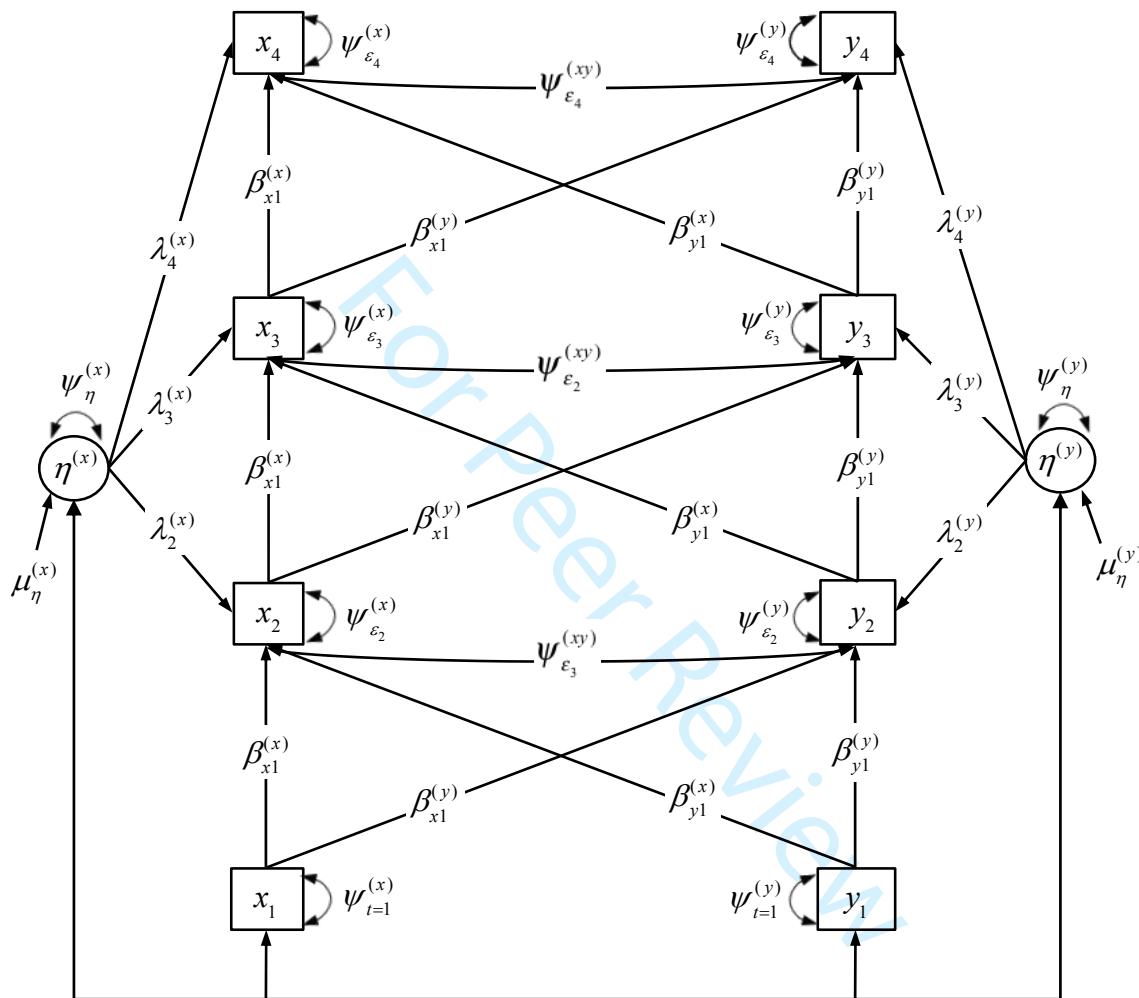


Figure 6b

Autoregressive Latent Trajectory Model with A Model-Estimated Trajectory and Showing Latent Means as μ Terms (Excluding Covariance Labels for Concision)



Online Appendix A

As with our first paper, we reanalyzed a subset of the data from Van Iddekinge et al. (2009) and Meier and Spector (2013). We offer results for each dataset in a similar but abridged fashion as in our main text (see Tables A1 and A2). Specifically, we compare findings using the approaches offered in our paper, with some exceptions because we are working with summary data and therefore we could not estimate the following models: a cross-lagged MLM (Model 2.2); a group-mean or within-group centered cross-lagged model (Model 2.3); a ‘system-GMM’ in differences and levels (Model 3.1); and a difference-GMM (Model 3.2). All other models are estimated, starting with our full GCLM (Model 1 in Tables A1 and A2), and proceeding to the three static models: random- and fixed-effects MLM with contemporaneous relationships specified as SEMs; and an LCM. We then estimate the following dynamic models: a traditional cross-lagged SEM (Model 2.1); a cross-lagged SEM with time-varying unit effects (Model 2.4); an SEM version of two Arellano-Bond models (one for each outcome variable) that are latent fixed-effects models (Model 3.3; see Allison, Williams, & Moral-Benito, 2017); an ALT with a linear trend to represent unit effects (Model 4.1); an ALT with a single time-varying unit effect (Model 4.2). All data, input, and output can be found in a .zip file as additional online material:

<https://www.dropbox.com/s/vxwbz7q4qiq0u87/Online%20Supplemental%20Material%20for%20Paper%20II.zip?dl=0>

As before, Van Iddekinge et al.’s (2009) data are $N = 861$ locations of an organization at $T = 6$ occasions, and from this sample we used entry-level employee retention as a measure of performance as y and employee training procedures as a measure of HR practices as x in the equations that follow. Meier and Spector’s data are $N = 663$ people at $T = 5$ occasions, and from this sample we used counterproductive work behaviors (CWB) as y and experienced incivility

(EI) as a work stressor x in the equations that follow—both of these variables are interpersonal in nature and therefore were appropriate for an individual-level example. We first present results for the Van Iddekinge et al. (2009) data and then Meier and Spector (2013). Throughout, any equations referenced are from the main text.

Table A1 shows results for entry-level employee retention and HR training from Van Iddekinge et al. (2009). The first model is a GCLM including time-varying unit effects and AR(1)MA(1)CL(1)CLMA(1) lagged effects. This model shows good fit to the data (CFI = .989, TLI = .981, SRMR = .023, RMSEA = .052). In terms of dynamics, training is moderately persistent with an AR $\beta_{x1}^{(x)} = .545$, but a small negative MA $\delta_{x1}^{(x)} = -.048$, which modestly reduces short-run persistence with $\beta_{x1}^{(x)} + \delta_{x1}^{(x)} = .497$. On the other hand, retention has a smaller AR $\beta_{y1}^{(y)} = .301$, and a positive MA $\delta_{y1}^{(y)} = .019$ slightly increases short-run persistence to $\beta_{y1}^{(y)} + \delta_{y1}^{(y)} = .320$. In a GCLM, the short-run training→retention effect is $\beta_{x1}^{(y)} + \delta_{x1}^{(y)} = .043$, with long-run impulse responses showing this positive effect is only slightly persistent, with CIs not including zero. On the other hand, the short-run retention→training effect is $\beta_{y1}^{(x)} + \delta_{y1}^{(x)} = .076$, and long-run impulse responses show this positive effect is moderately persistent, with CIs not including zero.

For comparison, we begin with static models to analyze the relationship between training, which is being treated as x , and retention, which is being treated as y in the equations referenced here. The results of two separate static random-effects MLMs estimated as SEMs (with $\psi_{\eta}^{(xy)} = 0$ imposed on the model) show a training→retention term for Eq. 3: $\beta_{y0}^{(x)} = .008$, $t = 1.674$, $p = .094$; and a retention→training term for Eq. 4: $\beta_{x0}^{(y)} = .191$, $t = 5.859$, $p < .001$. We then estimated two static fixed-effects models (by freeing $\psi_{\eta}^{(xy)}$) to investigate the same relationships. This

1
2
3 produces a training→retention term for Eq. 3: $\beta_{y0}^{(x)} = -.015, t = -2.824, p < .005$; and a
4
5
6 retention→training term for Eq. 4: $\beta_{x0}^{(y)} = -.037, t = -1.003, p = .316$. While the outcomes of the
7
8 random-effects models would seem to indicate a retention→training relationship, the positive
9
10 covariance among unit effects appears to be driving these results. This is shown in the results of
11
12 the fixed-effects models, which imply a negative within-unit relationship among the variables.
13
14 This is in contrast to conventional knowledge (e.g., Arthur, 1994) and the findings of Van
15
16 Iddekinge et al. (2009), who demonstrated a small but positive relationship from training to
17
18 retention. Of course, neither of these models can assist with causal inference to the extent that bi-
19
20 directional→lagged effects are excluded.
21
22
23

24
25 Finally, results of LCM show negative growth for training and retention: training
26
27 intercept $\mu_{\eta_1}^{(x)} = 7.571, t = 136.213, p < .001$; training slope $\mu_{\eta_2}^{(x)} = -.025, t = -1.971, p < .049$;
28
29 retention intercept $\mu_{\eta_1}^{(y)} = 8.809, t = 335.825, p < .001$; retention slope $\mu_{\eta_2}^{(y)} = -.051, t = -10.596, p$
30
31 $< .001$. Using the relationship among the slopes for inference, which is a correlation $\psi_{\eta_2}^{(xy)} = .252,$
32
33 $t = 4.693, p < .001$, indicating a positive relationship for training and retention trends. Yet, of
34
35 course, this relationship among these time-invariant variables does not help with determining the
36
37 direction of any causal effects among them as they change over time.
38
39
40
41
42

43
44 Moving on to the dynamic models (see Table A1), the outcomes of a traditional cross-
45
46 lagged SEM similar to the original Van Iddekinge et al. (2009) analyses (Model 2.1)
47
48 demonstrates adequate fit (CFI = .932, TLI = .921, SRMR = .159, RMSEA = .107), with AR
49
50 terms showing moderate persistence for training $\beta_{x1}^{(x)} = .66, t = 58.997, p < .001$, and retention
51
52 $\beta_{y1}^{(y)} = .719, t = 94.248, p < .001$. Also, CL effects are positive and statistically significant, with
53
54
55
56
57
58
59
60

1
2
3 training→retention $\beta_{x1}^{(y)} = .059, t = 15.152, p < .001$, and retention→training $\beta_{y1}^{(x)} = .198, t =$
4
5
6 9.305, $p < .001$. However, the results for these coefficients are confounded by unit effects. To
7
8 control for these effects, we analysed a cross-lagged SEM with time-varying unit effects (Model
9
10 2.4). The data fit the model well (CFI = .988, TLI = .982, SRMR = .022, RMSEA = .051), with
11
12 smaller AR terms for training $\beta_{x1}^{(x)} = .491, t = 18.712, p < .001$, and retention $\beta_{y1}^{(y)} = .326, t =$
13
14 22.558, $p < .001$, as expected after controlling for unit effects. Similar to the example in the main
15
16 text, these estimates come close to total AR and MA effects $\beta_x^{(x)} + \delta_x^{(x)}$ and $\beta_{y1}^{(y)} + \delta_{y1}^{(y)}$ in the
17
18 GCLM (Model 1 in Table A1). The CL term is also significant for training→retention $\beta_{x1}^{(y)} =$
19
20 .041, $t = 7.168, p < .001$, but it is no longer significant for retention→training $\beta_{y1}^{(x)} = .016, t = .5,$
21
22
23
24
25
26
27
28
29 $p = .617$, which again is not surprising given that unit effects are now controlled.

30 The results of the remaining dynamic models remain consistent with the cross-lagged
31
32 SEM that includes covariance among time-varying unit effects (in Table 1, compare Model 2.4
33
34 with Models 3.3, 4.1, 4.2). First we estimated an SEM version of two Arellano-Bond models
35
36 (one for each outcome variable), each of which is a latent fixed-effects model (Model 3.3) and
37
38 shows good fit to the data (for training, CFI = .994, TLI = .987, SRMR = .012, RMSEA = .035;
39
40 for retention, CFI = .988, TLI = .973, SRMR = .015, RMSEA = .063), with: an AR term for
41
42 training $\beta_{x1}^{(x)} = .467, t = 17.267, p < .001$; an AR term for retention $\beta_{y1}^{(y)} = .323, t = 22.141, p <$
43
44 .001; a CL effect for retention→training $\beta_{y1}^{(x)} = .001, t = .022, p = .982$; and a CL effect for
45
46
47
48
49 training→retention $\beta_{x1}^{(y)} = .051, t = 4.981, p < .001$.

50
51
52 The results of ALT models are also similar with both showing good fit. Here, the first
53
54 traditional ALT model showed good fit to the data (Model 4.1, CFI = .955, TLI = .933, SRMR =

.040, RMSEA = .099), with an AR term for training as $\beta_{x1}^{(x)} = .445, t = 10.006, p < .001$, and an AR term for retention as $\beta_{y1}^{(y)} = .270, t = 11.975, p < .001$. CL terms are also similar to Model 2.4, with CL effects for retention→training $\beta_{y1}^{(x)} = .070, t = 1.077, p = .282$, and a CL effect for training→retention $\beta_{x1}^{(y)} = .019, t = 2.101, p < .036$. Likewise, Model 4.2 (an ALT with a single time-varying unit effect) also fits the data well (CFI = .982, TLI = .977, SRMR = .036, RMSEA = .058), with: an AR effect for training $\beta_{x1}^{(x)} = .461, t = 18.972, p < .001$; an AR effect for retention $\beta_{y1}^{(y)} = .333, t = 22.731, p < .001$; CL effects for retention→training $\beta_{y1}^{(x)} = .049, t = 1.235, p = .217$; and CL effects for training→retention $\beta_{x1}^{(y)} = .045, t = 7.626, p < .001$. These results are interesting because, unlike the ALT results in the main text, the Van Iddekinge et al. (2009) results appear to robustly show a training→retention effect across all models.

Moving on to Meier and Spector (2013), we used EI, treated as x , and CWB, treated as y in the equations referenced below. Table A2 shows results for all models. Because some data were missing due to attrition, we used the minimum $N = 382$ (at the last occasion $t = 5$), because “if all indicators have missing values, it would appear that [minimum N per covariance] is the best choice of N . In contrast, [minimum N per variance] appears to be the best choice for N if the distribution of missingness is not uniform across the indicators” (Enders & Peugh, 2004, p. 17). In the panel data case, both conditions are met at the minimum N .

The first model in Table A2 is our GCLM—including time-varying unit effects, with AR(1)MA(1)CL(1)CLMA(1) lagged effects—that fits the data well (CFI = .981, TLI = .960, SRMR = .024, RMSEA = .073). In terms of dynamics, EI is moderately persistent with a negative AR $\beta_{x1}^{(x)} = -.563$, but a positive MA $\delta_{x1}^{(x)} = .617$ decreases short-run persistence to

1
2
3 $\beta_{x1}^{(x)} + \delta_{x1}^{(x)} = .054$. On the other hand, CWB has a small AR $\beta_{y1}^{(y)} = .139$, and a negative MA $\delta_{y1}^{(y)} =$
4
5
6 $-.059$ decreases short-run persistence to $\beta_{y1}^{(y)} + \delta_{y1}^{(y)} = .080$. In a GCLM, the short-run EI→CWB
7
8
9 effect is $\beta_{x1}^{(y)} + \delta_{x1}^{(y)} = -.036$, with long-run impulse responses showing this negative effect is only
10
11
12 slightly persistent, but CIs include zero. On the other hand, the short-run CWB→EI effect is
13
14 $\beta_{y1}^{(x)} + \delta_{y1}^{(x)} = .007$, and long-run impulse responses not only does not show a positive persistence,
15
16
17 but it also moderately persists in the opposite direction, with CIs including zero.

18
19
20 Next, similar to the previous example, we report the outcomes of the static models first.
21
22 The results of two separate static random-effects MLMs estimated as SEMs show positive and
23
24 significant effects for EI→CWB effect $\beta_{y0}^{(x)} = .201, t = 15.564, p < .001$, and a CWB→EI effect
25
26 $\beta_{x0}^{(y)} = .557, t = 15.670, p < .001$. The results of static fixed-effects MLMs are close to this
27
28 model, where they produce an EI→CWB $\beta_{y0}^{(x)} = .156, t = 10.715, p < .001$, and CWB→EI $\beta_{x0}^{(y)} =$
29
30
31 $.428, t = 10.331, p < .001$.

32
33
34
35
36 Next, an LCM shows negative growth over time for both variables: EI intercept $\mu_{\eta_1}^{(x)} =$
37
38 $1.800, t = 50.942, p < .001$; EI slope $\mu_{\eta_2}^{(x)} = -.040, t = -4.233, p < .001$; CWB intercept $\mu_{\eta_1}^{(y)} =$
39
40
41 $1.355, t = 58.898, p < .001$; CWB slope $\mu_{\eta_2}^{(y)} = -.012, t = -2.268, p < .023$. Here, there exists a
42
43
44 positive relationship for EI and CWB trends, with the correlation $\psi_{\eta_2}^{(xy)}$ slightly exceeding and
45
46
47 therefore we report it as $\psi_{\eta_2}^{(xy)} \approx 1, t = 3.984, p < .001$.

48
49
50
51 Reemphasizing our arguments in the text about the problems involved with any
52
53 assumptions of causality based on the statistically significant relationships estimated by each one
54
55 of these models, we move onto dynamic models. Understandably, due to their similar cross-
56
57
58
59
60

1
2
3 lagged design, Model 2.1 (See Table A2) is the only model showing a similar result to Meier and
4 Spector (2013). The data adequately fits to the model (CFI = .851, TLI = .818, SRMR = .12,
5
6 RMSEA = .154), with AR terms showing high persistence for EI $\beta_{x1}^{(x)} = .568, t = 26.745, p <$
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
This is curious given the similarity of these models.

To account for unit effects, we estimated a cross-lagged SEM with time-varying unit effects (Model 2.4). The model shows good fit statistics (CFI = .982, TLI = .968, SRMR = .029, RMSEA = .065), with substantially smaller AR terms for EI $\beta_{x1}^{(x)} = .096, t = 2.641, p < .008$, and for CWB $\beta_{y1}^{(y)} = .082, t = 2.377, p < .017$. The CL effect for CWB→EI is nonsignificant $\beta_{y1}^{(x)} = .025, t = .540, p = .589$, and for EI→CWB $\beta_{x1}^{(y)} = -.038, t = -2.242, p < .025$. The startling finding here is the latter negative relationship from EI to CWB, which suggests that more uncivilized behaviours (such as being interrupted by someone while talking), the less counterproductive behaviours will be manifested in the future.

The SEM versions of two Arellano-Bond models with latent fixed-effects model (Model 3.3) shows satisfactory fit (for EI, CFI = .993, TLI = .980, SRMR = .014, RMSEA = .047; for CWB, CFI = .985, TLI = .956, SRMR = .015, RMSEA = .071), and results are consistent with the cross-lagged SEM that controls for time-varying unit effects, with an AR term for EI $\beta_{x1}^{(x)} = .115, t = 3.054, p < .002$, and AR term for CWB $\beta_{y1}^{(y)} = .090, t = 2.50, p < .012$. Also, the CL

effect for CWB→EI $\beta_{y1}^{(x)} = .082, t = 1.013, p = .311$, with a key difference being a non-significant CL effect for EI→CWB $\beta_{x1}^{(y)} = -.038, t = -1.349, p = .177$.

Next, we move to ALT models with a linear trend for unit effects (Model 4.1) and one time-varying unit effect (Model 4.2). The data fits the first model well (CFI = .991, TLI = .982, SRMR = .029, RMSEA = .048). The AR term for EI is much smaller than that in the GCLM, with $\beta_{x1}^{(x)} = -.020, t = -.395, p = .693$, but CWB is in a more similar range, with $\beta_{y1}^{(y)} = .156, t = 2.273, p < .023$. The CL effects also change, with a CWB→EI effect $\beta_{y1}^{(x)} = -.016, t = -.199, p = .842$, and an unexpected EI→CWB effect $\beta_{x1}^{(y)} = -.081, t = -2.686, p < .007$. Model 4.2 also has adequate fit to the data (CFI = .979, TLI = .968, SRMR = .038, RMSEA = .065) but shows significant AR effects for EI $\beta_{x1}^{(x)} = .097, t = 2.637, p < .008$, and similar AR effects for CWB $\beta_{y1}^{(y)} = .088, t = 2.576, p < .01$. CL effects show no significant CWB→EI effect $\beta_{y1}^{(x)} < .001, t = -.004, p = .997$, but the EI→CWB effect remains significant $\beta_{x1}^{(y)} = -.037, t = -1.987, p < .047$.

In sum, these findings suggest similar problems as those offered in our main text, which emphasizes the benefits of the GCLM both on causal and pragmatic grounds, including controlling for unit effects and allowing a more general dynamical structure to be represented with MA and CLMA terms. When considering some of the reversed-directions for CL terms, we are reticent to offer substantive conclusions about these effects, but note that such reversals may be experienced when between- and within-unit covariances differ (i.e., when unit-effect covariances differ in sign from lagged effects).

Note: References are at the end of this document.

Table A1
Results for Dynamic Models, Using Data from Van Iddekinge et al. (2009)

Parameter	Model					
	1	2.1	2.4	3.3	4.1	4.2
Retention→Retention AR/MA Terms $\beta_{y1}^{(y)}$ and $\delta_{y1}^{(y)}$						
$\beta_{y1}^{(y)}$.301*	.719*	.326*	.323*	.270*	.333*
$\delta_{y1}^{(y)}$.019	---	---	---	---	---
$\beta_{y1}^{(y)} + \delta_{y1}^{(y)}$.320*	.719*	.326*	.323*	.270*	.333*
Training→Training AR/MA Terms $\beta_{x1}^{(x)}$ and $\delta_{x1}^{(x)}$						
$\beta_{x1}^{(x)}$.545*	.660*	.491*	.467*	.445*	.461*
$\delta_{x1}^{(x)}$	-.048	---	---	---	---	---
$\beta_{x1}^{(x)} + \delta_{x1}^{(x)}$.497*	.660*	.491*	.467*	.445*	.461*
Training→Retention CL/CLMA Terms $\beta_{x1}^{(y)}$ and $\delta_{x1}^{(y)}$						
$\beta_{x1}^{(y)}$.066*	.059*	.041*	.051*	.019*	.045*
$\delta_{x1}^{(y)}$	-.023	---	---	---	---	---
$\beta_{x1}^{(y)} + \delta_{x1}^{(y)}$.043*	.059*	.041*	.051*	.019*	.045*
Retention→Training CL/CLMA Terms $\beta_{y1}^{(x)}$ and $\delta_{y1}^{(x)}$						
$\beta_{y1}^{(x)}$	-.058	.198*	.016	.001	.070	.049
$\delta_{y1}^{(x)}$.135	---	---	---	---	---
$\beta_{y1}^{(x)} + \delta_{y1}^{(x)}$.076*	.198*	.016	.001	.070	.049

Note. Columns are named after the models described in the text as follows: Model 1 is our full AR(1)MA(1)CL(1)CLMA(1) GCLM; 2.1 is a cross-lagged SEM; 2.4 is a cross-lagged SEM with time-varying unit effects; 3.3 is an SEM version of a latent fixed-effects model (see Allison et al., 2017); 4.1 is an ALT with a linear trend to represent unit effects; 4.2 is an ALT with a single time-varying unit effect; and 4.3 is a multivariate extension of 3.3, similar to the ALT in 4.2 but with occasion effects α_i estimated and a mean structure for unit effects restricted to zero; AR = autoregressive; MA = moving average; CL = cross-lagged; CLMA = cross-lagged moving average.

Table A2

Results for Dynamic Models, Using Data from Meier & Spector (2013)

Parameter	Model					
	1	2.1	2.4	3.3	4.1	4.2
CWB→CWB AR/MA Terms $\beta_{y1}^{(y)}$ and $\delta_{y1}^{(y)}$						
$\beta_{y1}^{(y)}$.139	.606*	.082*	.090*	.156*	.088*
$\delta_{y1}^{(y)}$	-.059	---	---	---	---	---
$\beta_{y1}^{(y)} + \delta_{y1}^{(y)}$.080*	.606*	.082*	.090*	.156*	.088*
EI→EI AR/MA Terms $\beta_{x1}^{(x)}$ and $\delta_{x1}^{(x)}$						
$\beta_{x1}^{(x)}$	-.563*	.568*	.096*	.115*	-.020	.097*
$\delta_{x1}^{(x)}$.617*	---	---	---	---	---
$\beta_{x1}^{(x)} + \delta_{x1}^{(x)}$.054*	.568*	.096*	.115*	-.020	.097*
EI→CWB CL/CLMA Terms $\beta_{x1}^{(y)}$ and $\delta_{x1}^{(y)}$						
$\beta_{x1}^{(y)}$	-.073	.045*	-.038*	-.038	-.081*	-.037*
$\delta_{x1}^{(y)}$.037	---	---	---	---	---
$\beta_{x1}^{(y)} + \delta_{x1}^{(y)}$	-.036	.045*	-.038*	-.038	-.081*	-.037*
CWB→EI CL/CLMA Terms $\beta_{y1}^{(x)}$ and $\delta_{y1}^{(x)}$						
$\beta_{y1}^{(x)}$.297*	.224*	.025	.082	-.016	.000
$\delta_{y1}^{(x)}$	-.290*	---	---	---	---	---
$\beta_{y1}^{(x)} + \delta_{y1}^{(x)}$.007	.224*	.025	.082	-.016	.000

Note. Columns are named after the models described in the text as follows: Model 1 is our full AR(1)MA(1)CL(1)CLMA(1) GCLM; 2.1 is a cross-lagged SEM; 2.4 is a cross-lagged SEM with time-varying unit effects; 3.3 is an SEM version of a latent fixed-effects model (see Allison et al., 2017); 4.1 is an ALT with a linear trend to represent unit effects; 4.2 is an ALT with a single time-varying unit effect; and 4.3 is a multivariate extension of 3.3, similar to the ALT in 4.2 but with occasion effects α_i estimated and a mean structure for unit effects restricted to zero; CWB

1
2
3 = *counterproductive work behaviors*; *EI* = *experienced incivility*; *AR* = *autoregressive*; *MA* =
4
5 *moving average*; *CL* = *cross-lagged*; *CLMA* = *cross-lagged moving average*.
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Online Appendix B

Arellano-Bond and System-GMM for Dynamic Panel Data Models

The AB method treats a single dependent variable using a generalized method of moments (GMM) estimator (see Arellano, 2003; Roodman, 2009b). Conceptually, the logic of this approach can be shown with the lagged first differences from Eqs. 12 and 13 in the main text being instrumented by levels as follows (with predicted components shown with a $\hat{}$ term). We start by repeating Eqs. 12 and 13 in the main text as Eqs. B1 and B2 here for consistency:

$$x_{it} - x_{it-1} = \alpha_t^{(x)} - \alpha_{t-1}^{(x)} + \beta_{x1}^{(x)}(x_{it-1} - x_{it-2}) + \beta_{y1}^{(x)}(y_{it-1} - y_{it-2}) + \varepsilon_{it}^{(x)} - \varepsilon_{it-1}^{(x)} \quad (\text{B1})$$

$$y_{it} - y_{it-1} = \alpha_t^{(y)} - \alpha_{t-1}^{(y)} + \beta_{y1}^{(y)}(y_{it-1} - y_{it-2}) + \beta_{x1}^{(y)}(x_{it-1} - x_{it-2}) + \varepsilon_{it}^{(y)} - \varepsilon_{it-1}^{(y)} \quad (\text{B2})$$

and then show the prediction of first differences using lagged levels as follows:

$$\widehat{x_{it-1} - x_{it-2}} = a_{1t}^{(x)}x_{it-2} + \dots + a_{h-1t}^{(x)}x_{it-h} \quad (\text{B3})$$

$$\widehat{y_{it-1} - y_{it-2}} = a_{1t}^{(y)}y_{it-2} + \dots + a_{h-1t}^{(y)}y_{it-h} \quad (\text{B4})$$

wherein predicted differences can be thought of as taking the place of the first differences in Eqs. B1 and B2 (including predictors to account for α_t). In other words, the predictor $y_{it-1} - y_{it-2}$ in Eqs. B1 and B2 would be substituted by $\widehat{y_{it-1} - y_{it-2}}$ from Eq. B4. Here, it is important to notice that instruments (e.g., y_{it-2}) are not contemporaneous with error terms (e.g., $\varepsilon_{it}^{(y)} - \varepsilon_{it-1}^{(y)}$), which is meant to allow valid instrumentation—under conditions we cover shortly—as noted in the main text (i.e., instruments should not correlate with residuals). This is akin to the condition that the instruments (e.g., the predictors in Eqs. B3 and B4) are uncorrelated with the outcomes in Eqs. 1 and 2 except through the predictors in Eqs. 1 and 2. Organizational researchers may know this better as a kind of ‘full mediation’, meaning no ‘direct effect’ of instruments on outcomes and, instead, only an ‘indirect effect’ from instruments to outcomes via predictors—in the context of instruments, this will be described below as a condition for instrument validity.

This classic ‘difference-GMM’ approach is augmented by a ‘system-GMM’ method that is meant to improve efficiency by including equations in both first-differences and levels (see Blundell & Bond, 1998; Roodman, 2009b). Conceptually, this can be shown as follows:

$$x_{it} = \alpha_t^{(x)} + \beta_{x1}^{(x)}x_{it-1} + \beta_{y1}^{(x)}y_{it-1} + \varepsilon_{it}^{(x)} \quad (\text{B5})$$

$$y_{it} = \alpha_t^{(y)} + \beta_{y1}^{(y)}y_{it-1} + \beta_{x1}^{(y)}x_{it-1} + \varepsilon_{it}^{(y)} \quad (\text{B6})$$

wherein, conceptually, the lagged predictors are instrumented by first-differences as follows:

$$\widehat{x}_{it-1} = b_{1t}^{(x)}(x_{it-1} - x_{it-2}) \quad (\text{B7})$$

$$\widehat{y}_{it-1} = b_{1t}^{(y)}(y_{it-1} - y_{it-2}) \quad (\text{B8})$$

such that lagged levels in Eqs. B5 and B6 would be replaced by the predicted values in Eqs. B7 and B8. Again, note that instruments in B7 and B8 (i.e., the predictors) do not temporally overlap with the residuals from Eqs. B5 and B6. The idea is that there is a disjunction between the presence of η_i and the dynamic panel bias due to first-differencing, so that levels and first differences can instrument each other. Also of note is that longer lags of the differenced instruments in Eqs. B7 and B8 are redundant with the information in lags from Eqs. B3 and B4, and may therefore be unnecessary (Roodman, 2009b).

To frame our discussion, the GMM ‘moments’ that are used for estimation are typically described as j ‘moment conditions’ in a covariance matrix \mathbf{Z} , with $j(j+1)/2$ unique elements (see Arellano, 2003; Roodman, 2009b). Each element represents covariance among instruments and residuals, which is minimized to estimate β terms—the unfamiliar reader can imagine an analogue in the form of ordinary least squares, where residuals (or residual covariance among predictors and outcomes) are minimized to estimate β terms. For example, together, Eqs. B1 and B3 imply a moment condition $E[x_{it-2}(\varepsilon_{it}^{(x)} - \varepsilon_{it-1}^{(x)})] = 0$, which would be an on-diagonal term in \mathbf{Z} and minimized to estimate $\beta_{x1}^{(x)}$ (akin to minimizing $COV(x_{it-2}, (\varepsilon_{it}^{(x)} - \varepsilon_{it-1}^{(x)}))$). In turn, a moment condition covariance matrix \mathbf{Z} can be used for model checking against observed data.

This checking is done by recognizing that moment conditions are criteria for instrument validity vis-à-vis the assumption that instruments are uncorrelated with residuals (if this does not hold, β terms are biased). For example, $E[x_{it-2}(\varepsilon_{it}^{(x)} - \varepsilon_{it-1}^{(x)})] = 0$ in Eqs. B1 and B3 only if x_{it-2} is unrelated to x_{it} except via $x_{it-1} - x_{it-2}$ (i.e., the full mediation or no indirect effect idea described above). This type of assumption can be checked by residual autocorrelation tests in difference equations at two lags and beyond (because differencing creates autocorrelation at a single lag), as well as Sargan/Hansen tests with a null hypothesis of no instrument-residual covariance (i.e., moment conditions with zero expectations, such that small p -values indicate rejecting the assumption of valid instruments and thus large p -values imply valid assumptions; Roodman, 2009b). This latter test is done based on j moment conditions, which grows at a rate T^2 . In turn, the \mathbf{Z} matrix can become very large unless instrument lags are limited or are

1
2
3 ‘collapsed’ with multiple moment conditions as a single term for estimation/hypothesis tests
4
5 (Roodman, 2009a).
6

7
8 When taken as a whole, this approach to panel data analysis produces many dilemmas
9
10 related to: choosing the number of instruments and their range in T ; using difference- versus
11 system-GMM; and interpreting Sargan/Hansen tests. To explain, first, keeping instruments
12 independent of residuals requires making assumptions about predictors because they instrument
13 themselves (Roodman, 2009a, 2009b). The instrument specifications in Eqs. B1-B8 imply an
14 assumption that predictors are ‘predetermined’ (i.e., weakly or sequentially exogenous), meaning
15 they are influenced by past residuals but uncorrelated with contemporaneous residuals (e.g., in
16 Eqs. B1 and B3, x_{it-2} may be related to ε_{it-3} but not ε_{it-2} ; Arellano, 2003). If this does not hold,
17 an instrument is endogenous and should be lagged further because, otherwise, it will correlate
18 with residuals (e.g., Eq. B3 could start with x_{it-3} instead of x_{it-2}). Alternatively, if a predictor is
19 assumed to be ‘strictly exogenous’ and uncorrelated with all residuals, it will instrument itself
20 without lags. These possibilities create the problem of not knowing how to treat instruments—
21 unlike our GCLM that assumes endogeneity and allows predetermination or exogeneity to exist
22 as null covariances or effects (if researchers want to go down this route for any reason).
23
24 Although tests of instrument validity may be used to choose an instrument specification, as we
25 describe shortly these tests often fail to detect assumption violations. Also, erring on the side of
26 caution by assuming endogeneity increases the minimum T required to 4 for model estimation,
27 and reduces the number of available instruments, which can create other issues we now discuss.
28
29

30
31 The second problem with GMM is that too few or too many instruments causes problems
32 because too few reduces efficiency but too many overfits a model and biases parameters (as is
33 well-known in regression, $R^2 \rightarrow 1$ as predictors $k \rightarrow N$). Thus, few instruments are needed to
34 avoid overfitting but this reduces efficiency. Indeed, as T increases, the potential instrument
35 count can approach or exceed N very quickly, which leads to overfitting that, in turn, causes
36 dynamic panel bias in Eqs. B5 and B6 by failing to expunge unit effects in Eqs. B7 and B8 (i.e.,
37 the results will start to resemble cross-lagged SEM/random-effects MLM). To avoid this,
38 instrument counts can be reduced by using fewer lags or collapsing them (Roodman, 2009a,
39 2009b), but how best to do this for achieving good estimates is never certain because parameters
40 and the validity of assumptions are, by definition, unknown.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Third, problems are encountered as AR terms approach 1 (i.e., a random walk), which is
4 especially true in difference-GMM because high persistence reduces the unique information in
5 levels as instruments, making system-GMM better for variables like income that often exhibit a
6 random walk (Blundell & Bond, 1998). Yet, system-GMM makes moment conditions harder to
7 satisfy because residuals contain η_i , as in Eqs. B5 and B6, which can correlate with differenced
8 instruments, as in Eqs. B7 and B8 (i.e., a correlation between a time-invariant factor such as $\eta_i^{(y)}$
9 and the magnitude of occasion-to-occasion change such as $y_{i2} - y_{i1}$; Baltagi, 2013b; Blundell &
10 Bond, 1998; Roodman, 2009b). Indeed, it is easy to imagine scenarios wherein time-invariant
11 factors such as personality or culture systematically correlate with change over time. Thus,
12 “[w]here system GMM offers the most hope, it may offer the least help” (Roodman, 2009a).
13
14

15 Fourth, compounding the previous problem, system-GMM increases j , which causes the
16 validity of the Sargan/Hansen test to become biased towards accepting the alternate hypothesis
17 of valid instruments as the size of \mathbf{Z} grows in relation to N (i.e., p -values for these tests become
18 inappropriately large). For difference- and system-GMM, this puts researchers in the awkward
19 position of having to choose instruments based on tests that make it more difficult to detect
20 invalid instruments precisely when using more instruments, which is ironic because having more
21 instruments increases the chances that instruments are being used which are invalid (Roodman,
22 2009a). Again, this motivates reducing the number of instruments and/or collapsing them to aid
23 Sargan/Hansen test validity, and in general being very cautious interpreting test p -values, such as
24 by rejecting instrument validity when $p < .25$ or even higher. More generally, this illustrates the
25 dilemma of choosing difference- versus system-GMM, the latter of which imposes more
26 assumptions while at the same time reducing the ability to test these assumptions.
27
28

29 Fifth, another dilemma arises when autocorrelation and Sargan/Hansen tests indicate the
30 violation of assumptions (i.e., small p -values). One way to address this is to substantively alter a
31 model by increasing AR and/or CL lag orders to reduce autocorrelation and/or the correlation
32 among instruments and future residuals. However, another option is increasing the initial lag of
33 instruments, so that their temporal distance from residuals increases (e.g., the right-hand side of
34 Eq. B3 might begin with $a_1^{(x)}x_{it-3}$). The issue is that autocorrelation and Sargan/Hansen tests can
35 be construed as substantive tests of model misspecification or of instrument misspecification.
36
37 The result is that, unlike SEM which only allows substantive model respecifications, researchers
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 must choose substantive model changes and/order changes to instrument sets that merely serve
4
5 to avoid violating assumptions, the latter of which can call into question substantive inferences
6
7 (Hendry, 2002). Indeed, for a β term of any magnitude with any p -value, it is impossible to
8
9 know whether results are due to the choice of predictors or instruments and either can be
10
11 modified, which adds to the complexity of using this approach and interpreting hypothesis tests.
12
13

14
15 In sum, “[GMM] estimators carry a great and under-appreciated risk: the capacity *by*
16
17 *default* to generate results that are invalid and appear valid... One leading estimator, difference
18
19 GMM, often suffers from weak instrumentation. The favored alternative, system GMM, works
20
21 only under arguably special circumstances. Perhaps, the lesson to be drawn is that [the methods]
22
23 have serious limitations” (Roodman, 2009a, p. 156). With this in mind, it is convenient that an
24
25 SEM approach allows overcoming many of these difficulties (Allison et al., 2017; Bai, 2013;
26
27 Moral-Benito, 2013; Williams, Allison, & Moral-Benito, 2018).
28
29
30

31
32 In our main text we illustrate this by estimating Eqs. B1-B8 using Stata’s ‘xtabond2’ with
33
34 robust *SEs* using a Windmeijer (2005) correction and a twostep system-GMM estimator (see
35
36 Roodman, 2009b). This would normally motivate control variables (e.g., unemployment) and
37
38 specification searches to choose a number of substantive model lags, which we avoid to keep the
39
40 models consistent with those already explored. To increase the validity of Sargan/Hansen tests,
41
42 we collapse instruments but still have trouble keeping j (and \mathbf{Z}) small (with $j = 15$ and thus 120
43
44 unique elements in \mathbf{Z} , which approaches our N), so we retain instrument sets with $p \geq .5$ because
45
46 very large p -values are expected in such cases. Alternatively, in the second model run with a
47
48 model in differences only, $j = 12$ and therefore Sargan/Hansen tests are more trustworthy. To see
49
50 how each instrument set contributed to misfit, we separately tested the validity of the differences
51
52 and levels equations for each predictor, allowing differences-in-Sargan/Hansen tests that indicate
53
54 whether overall fit improves when removing a set of instruments (see Roodman, 2009a, 2009b).
55
56
57
58
59

60
In terms of moving beyond AB methods using SEM, when considering the maximum-likelihood based alternative by Allison and colleagues (see Allison et al., 2017; Moral-Benito, 2013; Williams et al., 2018), it is notable that this method allows the introduction of control variables in a way that is potentially very useful. In econometric parlance, if a control variable is exogenous it would show no covariance with other modeled variables; if it is predetermined it would only correlated with η_i and past observations; if it is endogenous it would also correlate

with contemporaneous observations. Thus, unlike AB methods that rely on GMM, the status of predictors as predetermined, exogenous, or endogenous can be specified and model-estimated in SEM rather than imposed and tested with tools like Sargan/Hansen tests.

This said, it is relevant to consider various issues related to estimating separate models for each dependent variable, and including control variables. Consider that control variables are not formally included in a cross-lagged structure with AR terms. Instead, many covariances are specified among controls and residuals as well as η_i . This allows missing indirect effects that function via control variables (see Hsiao, 1982; Lütkepohl, 1993; Dufour & Renault, 1998), which can be understood by recognizing that, without AR terms, predictor and control variables cannot participate in impulse responses. Thus, long-run effects that involve predictors and control variables as parts of a causal chain cannot be tested. For this and other reasons, Allison (2009) recommends models that treat predictors as outcomes in subsequently estimated models (as we did in the main text), which is common for MLM and AB methods. In this tradition, “[a]n attractive feature of this approach is that it does not require models for [predictors in a] series to be specified in order to estimate the parameters” (Bond, 2002, p. 151), avoiding potential misspecification of a model for predictors. However, alternative perspectives exist that motivate models like ours, such as that “[t]he x and y equations should be modeled jointly even [if] the parameters of interest are those in the y equation only” Bai (2013; see also Ashley, 1983; Atukeren, 2008; Sims, 1980; Stock & Watson, 2001).

Reference List

- Allison, P. D. (2009). *Fixed effects regression models*. Thousand Oaks, CA: Sage Publications.
- Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum likelihood for cross-lagged panel models with fixed effects. *Socius*, 3, 1-17. DOI: 10.1177/2378023117710578
- Arellano, M. (2003). *Panel data econometrics*. Oxford: Oxford University Press.
- Arthur, J. B. (1994). Effects of human resource systems on manufacturing performance and turnover. *Academy of Management Journal*, 37(3), 670-687. DOI: 10.5465/256705
- Ashley, R. (1983). On the usefulness of macroeconomic forecasts as inputs to forecasting models. *Journal of Forecasting*, 2(3), 211-223. DOI: 10.1002/for.3980020304
- Atukeren, E. (2008). Christmas cards, Easter bunnies, and Granger-causality. *Quality & Quantity*, 42(6), 835-844. DOI: 10.1007/s11135-007-9078-2
- Bai, J. (2013). Likelihood approach to dynamic panel models with interactive effects. MRPA Paper No. 50267. Retrieved from: <http://mrpa.ub.uni-muenchen.de/50267/> DOI: 10.2139/ssrn.2332992
- Baltagi, B. (2013b). *Econometric analysis of panel data*. John Wiley & Sons.
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143. DOI: 10.1016/S0304-4076(98)00009-8
- Bond, S. R. (2002). Dynamic panel data models: A guide to micro data methods and practice. *Portuguese Economic Journal*, 1(2), 141-162. DOI: 10.1007/s10258-002-0009-9

- 1
2
3 Dufour, J. M., & Renault, E. (1998). Short run and long run causality in time series: Theory.
4 *Econometrica*, 66, 1099-1125. DOI: 10.2307/2999631
- 5 Enders, C. K., & Peugh, J. L. (2004). Using an EM covariance matrix to estimate structural
6 equation models with missing data: Choosing an adjusted sample size to improve the
7 accuracy of inferences. *Structural Equation Modeling*, 11(1), 1-19. DOI:
8 10.1207/S15328007SEM1101_1
- 9 Hendry, D. F. (2002). Applied econometrics without sinning. *Journal of Economic Surveys*,
10 16(4), 591-604. DOI: 10.1111/1467-6419.00180
- 11 Hsiao, C. (1982). Autoregressive modeling and causal ordering of economic variables. *Journal*
12 *of Economic Dynamics and Control*, 4, 243-259.
- 13 Lütkepohl, H. (1993). Testing for causation between two variables in higher-dimensional VAR
14 models. In H. Schneeweib & K. F. Zimmermann (Eds.), *Studies in Applied Econometrics*
15 (pp. 75-91). Heidelberg: Physica-Verlag HD.
- 16 Meier, L. L., & Spector, P. E. (2013). Reciprocal effects of work stressors and counterproductive
17 work behavior: A five-wave longitudinal study. *Journal of Applied Psychology*, 98(3),
18 529-539. DOI: 10.1037/a0031732
- 19 Moral-Benito, E. (2013) Likelihood-based estimation of dynamic panels with predetermined
20 regressors. *Journal of Business & Economic Statistics*, 31(4), 451-472. DOI:
21 10.1080/07350015.2013.818003
- 22 Roodman, D. (2009a). A note on the theme of too many instruments. *Oxford Bulletin of*
23 *Economics and Statistics*, 71(1), 135-158. DOI: 10.1111/j.1468-0084.2008.00542.x
- 24 Roodman, D. (2009b). How to do xtabond2: An introduction to difference and system GMM in
25 Stata. *Stata Journal*, 9(1), 86-136. DOI: 10.2139/ssrn.982943
- 26 Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48, 1-48. DOI:
27 10.2307/1912017
- 28 Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. *The Journal of Economic*
29 *Perspectives*, 15(4), 101-115. DOI: 10.1257/jep.15.4.101
- 30 Van Iddekinge, C. H., Ferris, G. R., Perrewé, P. L., Perryman, A. A., Blass, F. R., & Heetderks,
31 T. D. (2009). Effects of selection and training on unit-level performance over time: A
32 latent growth modeling approach. *Journal of Applied Psychology*, 94(4), 829-843. DOI:
33 10.1037/a0014453
- 34 Williams, R., Allison, P. D., & Moral-Benito, E. (2018). Linear dynamic panel-data estimation
35 using maximum likelihood and structural equation modeling. *The Stata Journal*, 18, 293-
36 326.
- 37 Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step
38 GMM estimators. *Journal of Econometrics*, 126(1), 25-51. DOI:
39 10.1016/j.jeconom.2004.02.005
- 40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60