Journal of Management
Vol. 46 No. 1, January 2020 70–99
DOI: 10.1177/0149206319868016

Bridging Methodological Divides Between Macro- and Microresearch: Endogeneity and Methods for Panel Data

Paul D. Bliese
Donald J. Schepker
Spenser M. Essman
Robert E. Ployhart
University of South Carolina

Both macro- and micro-oriented researchers frequently use panel data where the outcome of interest is measured repeated times. Panel data support at least five different modeling frameworks (within, between, incremental/emergent, cross-level, and growth). Researchers from macro- and micro-oriented domains tend to differentially use the frameworks and also use different analytic tools and terminology when using the same modeling framework. These differences have the potential to inhibit cross-discipline communication. In this review, we explore how macro- and microresearchers approach panel data with a specific emphasis on the theoretical implications of choosing one framework versus another. We illustrate how fixed-effects and random-effects models differ and how they are similar, and we conduct a thorough review of 142 articles that used panel data in leading management journals in 2017. Ultimately, our review identifies ways that researchers can better employ fixed- and random-effects models, model time as a meaningful predictor or ensure unobserved time heterogeneity is controlled, and align hypotheses to analytic choice. In the end, our goal is to help facilitate communication and theory development between macro- and micro-oriented management researchers.

Keywords: *methodology; endogeneity; panel data; random effects; fixed effects*

Acknowledgments: We thank the Editor and two anonymous reviewers for their helpful and constructive suggestions. We also thank Adam Steinbach for providing feedback on the paper, and Silvia Clark for her assistance with references.

Supplemental material for this article is available with the manuscript on the JOM *website.*

Corresponding author: Paul D. Bliese, Darla Moore School of Business, University of South Carolina, 1014 Greene St., Columbia, SC 29208, USA.

E-mail: paul.bliese@moore.sc.edu

In the last decade, there have been significant theoretical advances seeking to unite macro- and microdisciplines (e.g., Hutzschenreuter & Horstkotte, 2013; Wright, Coff, & Moliterno, 2014). Nonetheless, critical gaps remain, and one of the most persistent is how the disciplines analyze panel data. Panel data are often at the center of research bridging the macro-micro divide; however, they can be examined using at least five broad modeling frameworks (within, between, emergent, cross-level, and growth model) that vary in usage across disciplines. To further complicate communication, vocabulary and specific analytic tools (e.g., fixed- vs. random-effects models) also vary across disciplines. These differences have important theoretical implications that potentially limit conversations between macro- and microresearchers.

We review key vocabulary, broad modeling frameworks, and specific analytic tools associated with panel data. While many themes we address are methodological, our primary goal is to provide a conceptual review that helps facilitate cross-discipline communication. Methodology sets the boundaries for specifying theoretical questions—as such, it plays a central role in theory development. That is, methodological choices have consequences as to *what can be* (and sometimes actually *is*) tested and helps inform what we *know* about theory. A key part of our review centers on defining and describing terms common in one discipline that may be uncommon or misunderstood in the other. By drawing links between methodology and terminology, we help refine and enhance methodological understanding and theory specification across macro- and microdomains. In so doing, we also highlight new research avenues.

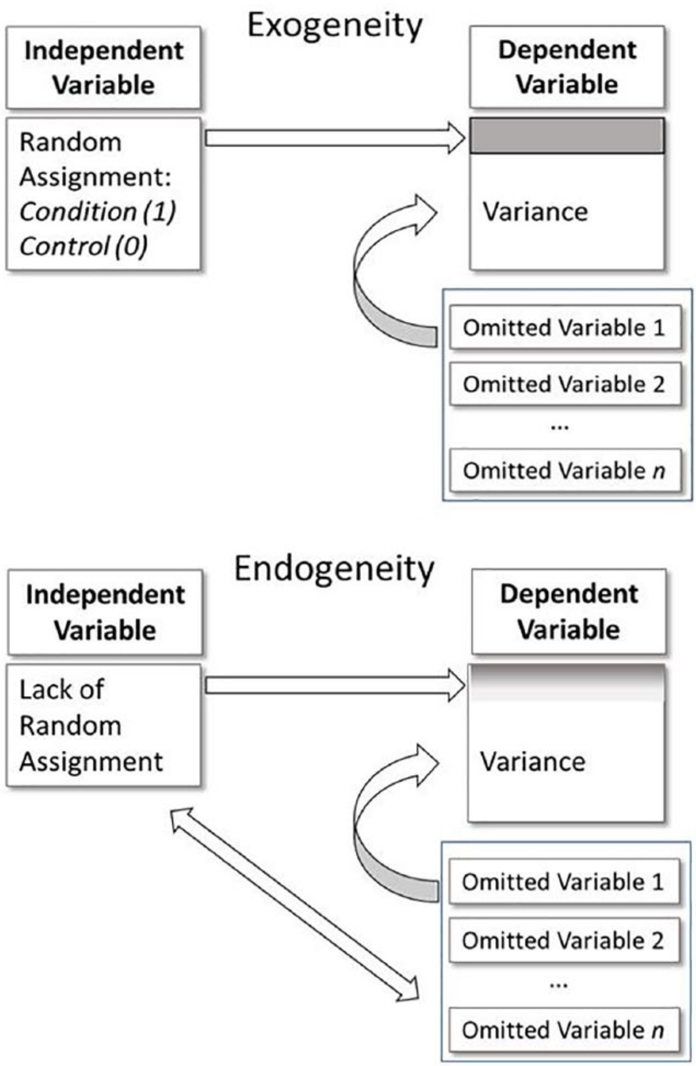
We seek to make five specific contributions. First, we define endogeneity and identify two forms associated with panel designs. We show, perhaps to the surprise of scholars on the "other side of the divide," that both macro- and microscholars seek to address similar concerns in panel data despite differences in terminology and analytic approaches. Second, we describe five prototypical modeling frameworks used to analyze panel data. Third, we discuss specific analytic tools (fixed- vs. random-effect models) and illustrate commonalities, differences, limitations, and implications for testing theory. Fourth, we conduct a focused (yet comprehensive) review of 142 articles that rely on panel data published in six leading macro- and microjournals. Our review identifies current practices in theory and hypothesis development, design, and analytic choices. Fifth, we highlight best practices and offer recommendations for future research.

Endogeneity Defined

Endogeneity is defined as *a condition where the independent variable (also a predictor or explanatory variable) is related to a model's error term* (Bailey, 2017; Hamilton & Nickerson, 2003; Semadeni, Withers, & Certo, 2014). To obtain an unbiased estimate of the relationship between an independent variable and the dependent variable (or outcome), the independent variable must be *un*related to other factors in the model's residual error. In many situations, however, one or more unmeasured variables are related to both the independent variable and the dependent variable. Under these circumstances, the estimated relationship between the independent variable and dependent variable does not equal the true value (Kennedy, 2008; Semadeni et al., 2014).[1]

Another way to understand endogeneity is to consider its opposite—exogeneity. Exogeneity occurs when the independent variable is *un*related to other factors captured by the error term. One way to ensure exogeneity is to randomly assign entities (e.g., employees, teams, firms) to

Figure 1
Exogeneity and Endogeneity



different levels of an independent variable, such as to treatment and control conditions. When levels of the independent variable are randomly assigned, residual errors are unrelated to values of the independent variable (Bailey, 2017; Bodner & Bliese, 2018; Rubin, 2008; Shadish, Cook, & Campbell, 2002). Thus, random assignment of the independent variable avoids endogeneity by equally distributing omitted causes across levels of the independent variable.

Figure 1 illustrates the difference between exogeneity and endogeneity. In the top graphic illustrating exogeneity, the independent variable predicts a small but clearly defined portion of variance in the dependent variable. Typically, the dependent variable also contains additional unexplained variance that in the exogeneous situation is unrelated to the levels of the independent variable. The bottom graphic illustrates endogeneity. The variance explained by the independent variable is not clearly delineated because part of the relationship could be due to omitted variables related to both the independent and dependent variable. When endogeneity is present, the challenge is to identify the portion of variance unique to the independent variable.

Exogeneity plays a key role in advancing theory because establishing causality requires eliminating alternative explanations (Antonakis, Bendahan, Jacquart, & Lalive, 2010; Bailey, 2017; Bodner & Bliese, 2018). Establishing causality also requires temporal precedence and evidence of relationships between variables (Kenny, 1979), but exogeneity is a necessary requirement. Calls to rely more heavily on randomized trials in both microresearch (e.g., Eden 2017; Highhouse, 2009) and macroresearch (Bailey, 2017; Chatterji, Findley, Jensen, Meier, & Nielson, 2016) reflect a push to strengthen researchers' ability to make causal claims.

Realities in both macro- and microresearch make it unrealistic to expect researchers to abandon nonexperimental research, so both disciplines have developed methodological approaches to help minimize the effects of endogeneity. In microresearch, concerns about "common method variance," "reverse causality," and "omitted variables" are but a few of the terms used in lieu of "endogeneity." For many microresearchers, though, the term endogeneity, and the methods used to address it, are unfamiliar. Antonakis and colleagues (2010) provide a comprehensive overview of relevant methods, but even with such resources, the double hit of unfamiliar terms and solutions gives the impression endogeneity is a novel econometric concept.

Endogeneity has a specific definition but occurs in numerous forms, and different analytic approaches are used to address each form (Antonakis et al., 2010). Here we focus on two forms—unobserved heterogeneity across the units of observation (e.g., firms) and unobserved heterogeneity over time. These forms of endogeneity are pervasive in panel data, yet macro- and microresearchers tend to approach them differently, which has implications for drawing inferences and testing conceptual questions. We also explore how both macro- and micro-oriented research can benefit from adopting approaches used in their nonrespective areas. We recognize that macroresearchers often focus on other forms of endogeneity (e.g., Antonakis et al., 2010; Shaver, 1998), but understanding how the two basic forms are addressed in panel data is important for facilitating cross-discipline communication.

Modeling Frameworks for Panel Data

Panel data are characterized by "repeated observations on the same cross section of, say, individuals, households, firms, or cities, over time" (Wooldridge, 2010: 6). Panel data are also referred to as "longitudinal" or "multilevel" among micro-oriented researchers. Panel data are widely used in macroresearch, and key econometric texts devote entire chapters to the analysis of panel data (e.g., Greene, 2018; Verbeek, 2008). Certo, Withers, and Semadeni (2017) noted that by 2014, panel data represented over half of all published articles in *Strategic Management Journal*. Our review finds similar usage in the macro-oriented literature.

To illustrate, the first five columns of Table 1 provide 6 years of hypothetical panel data on annual turnover and return on investment (ROI) for three firms. In microterminology, the

Table 1
Example Panel Data

Firm	Industry	Year	Turnover	ROI	Average Turnover	Average ROI	Demeaned Turnover	Demeaned ROI
1	Healthcare	2012	12	7	9.50	8.57	2.50	−1.57
1	Healthcare	2013	11	8	9.50	8.57	1.50	−0.57
1	Healthcare	2014	10	8.8	9.50	8.57	0.50	0.23
1	Healthcare	2015	9	9	9.50	8.57	−0.50	0.43
1	Healthcare	2016	8	9.1	9.50	8.57	−1.50	0.53
1	Healthcare	2017	7	9.5	9.50	8.57	−2.50	0.93
2	Other	2012	33	10	30.50	11.15	2.50	−1.15
2	Other	2013	32	10.5	30.50	11.15	1.50	−0.65
2	Other	2014	31	11	30.50	11.15	0.50	−0.15
2	Other	2015	30	11.5	30.50	11.15	−0.50	0.35
2	Other	2016	29	11.9	30.50	11.15	−1.50	0.75
2	Other	2017	28	12	30.50	11.15	−2.50	0.85
3	Healthcare	2012	6	2	4.67	2.93	1.33	−0.93
3	Healthcare	2013	5.5	2.5	4.67	2.93	0.83	−0.43
3	Healthcare	2014	5	2.9	4.67	2.93	0.33	−0.03
3	Healthcare	2015	4.5	3.1	4.67	2.93	−0.17	0.17
3	Healthcare	2016	4	3.5	4.67	2.93	−0.67	0.57
3	Healthcare	2017	3	3.6	4.67	2.93	−1.67	0.67
Correlations			Raw = .80		Between = .85[a]		Within = −.95[a]	
Slope			Simple pooled = 0.25		Between = 0.26[a]		Within = −0.45[a]	
Firm fixed effects			–0.45					
Firm/year fixed effects			–0.28					

Note: ROI = return on investment.
[a]Estimates in the text and table are based on nonrounded variables from raw data. Using rounded values in the table may result in estimates that differ from text.

yearly observations represent Level 1 or lower-level variables and the firms represent Level 2 or higher-level entities (Level 1 is nested within Level 2). Level 2 variables describe the firm (e.g., “industry” is a Level 2 variable, as is “average turnover”). We use the “Level 1” and “Level 2” terminology to provide generalizable language across contexts and designs.

According to context-emergent turnover theory, one might expect that greater collective turnover will contribute to lower ROI as a result of the depletion of human capital resources and disruptions to productivity (Nyberg & Ployhart, 2013). Within each firm, higher turnover is associated with lower ROI. In the example provided in Table 1, however, firms with high average turnover also have high average ROI—Firm 2 has 30.5% average turnover but the highest average ROI at 11.15, while Firm 3 has low average turnover (4.67%) and low average ROI (2.93). Columns eight and nine provide firm demeaned (i.e., group-mean centered) variables calculated by subtracting the firm mean from each time observation. The rows below the raw data provide relevant correlations and regression slopes. Notice the correlations for the raw data and that based on firm means are positive while the within correlation based on the demeaned variable is negative.

Both macro- and microresearchers recognize that analyzing panel data using a simple pooled regression model that ignores heterogeneity associated with firm and time will likely return ambiguous parameter estimates and erroneous standard errors of within-firm relationships (e.g., Bailey, 2017; Bliese & Hanges, 2004; Pinheiro & Bates, 2000; Raudenbush & Bryk, 2002). For instance, a simple ordinary least squares (OLS) regression model regressing ROI on turnover for the data in Table 1 returns a positive (and problematic) slope estimate of 0.247 ($p < .001$).

What may be less obvious is that panel data can be analyzed using five alternative modeling frameworks, each of which tests different, but potentially complementary, theoretical propositions. The five modeling frameworks involve testing relationships (1) within entities, (2) between entities, (3) as incremental or emergent relationships, (4) as cross-level interactions, or (5) as a growth model.[2] To be clear, we are *not* implying researchers apply these five to see what "works," nor should researchers default to only one of the five frameworks under the belief that only one is appropriate. As we discuss, the theoretical question of interest underlies the choice.

Within-Entity Relationships

Within-entity relationships involve theoretical questions about time-varying (Level 1) variables. In our running example, a within-entity hypothesis could be phrased as "within-firm turnover is negatively related to within-firm ROI" or "a relative increase in a firm's turnover is associated with a relative decrease in a firm's ROI." To establish a causal claim, we would ideally randomly assign firms to either increase or decrease turnover and examine the impact on ROI, but this is obviously not feasible. Therefore, we strengthen our ability to make causal inferences by controlling for unobserved heterogeneity.

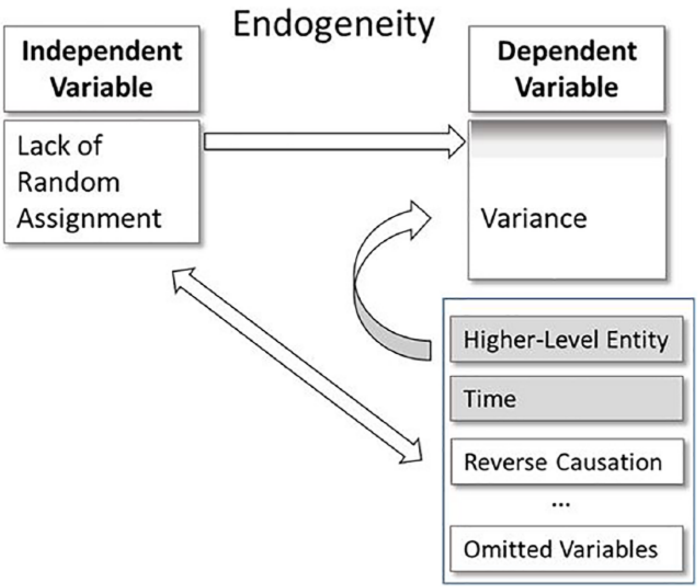
With respect to heterogeneity due to time, turnover and ROI may both change as a function of an exogeneous shock. A temporal event such as an economic recession may cause both variables to covary as a function of the event rather than as a function of each other. By removing variance due to time, we have a better estimate of how turnover and ROI are related. With respect to the firm-level heterogeneity, any statistical model that fails to account for the higher-level unit suffers from a specific form of endogeneity where the model error is related to characteristics of higher-level entities (Bailey, 2017; Certo et al., 2017). Characteristics include any time-invariant characteristic of the firm (i.e., Level 2 characteristics), such as industry or whether the firm is privately held (assuming this status does not change over time).[3]

Figure 2 illustrates the two common forms of endogeneity in panel data and also shows that many additional forms exist. As we explain later, analytic tools routinely used by macro-oriented researchers are extremely effective in controlling for unobserved heterogeneity related to time and firm. As a consequence, macro-oriented researchers tend to model these two forms as "a given" and focus on other forms of endogeneity (e.g., Shaver, 1998). In contrast, micro-oriented researchers use analytic tools in ways that are not always as effective in addressing these forms of endogeneity, which leads to differences in the baseline conversation.

Between-Entity Relationships

Between-entity relationships focus on theoretical questions involving characteristics of the higher-level entities (Level 2 variables). For example, a between-entity hypothesis could

Figure 2
Endogeneity in Panel Data



be "firms that (on average) have high turnover will (on average) have low ROI." By aggregating the data (across years) in Table 1 and obtaining firm-level means for turnover and ROI, we can estimate how average turnover is related to average ROI. In Table 1, the raw between-level correlation is .848. Reverse causation and unmeasured firm-level variables represent important forms of endogeneity in aggregate models.

While it may seem straightforward to develop and test theory about aggregate-level relationships, it can be deceptively complex. Researchers (e.g., Bliese, Maltarich, & Hendricks, 2018; Lincoln & Zeitz, 1980) have shown that relationships between group means will mirror raw relationships even if group membership is arbitrarily assigned. For instance, in Table 1 the raw correlation is .797. If we (a) randomly assign rows to create three new "pseudofirms," (b) calculate three pseudofirm means for turnover and ROI, and (c) estimate the between-group correlation, the between-group correlation would be .797.[4] That is, even if there is no meaningful variance associated with the higher-level entity (which we approximate by randomly assigning rows to "pseudofirms"), the magnitude of within-firm relationships will be mirrored in between-firm relationships. From a theory development perspective, this inheritance across levels requires researchers to carefully consider the origin of the mechanism underlying the group-mean correlation (Lincoln & Zeitz, 1980) and suggests that interpretations of relationships between group means need to be considered in light of within-firm relationships.

Incremental or Emergent Relationships

A third variant formally tests whether the within-firm and between-firm relationships differ from each other. This variant is called incremental (Hofmann & Gavin, 1998) or emergent

(Bliese et al., 2018) because a different relationship emerges at the aggregate level. For example, we might hypothesize that “the firm-level relationship between average turnover and average ROI significantly differs in magnitude from the within-firm relationship between turnover and ROI.”

Differences in causal mechanisms across levels drive differences in relationship strength (see Bliese et al., 2018; Certo et al., 2017; Chan, 1998; Diez-Roux, 1998; Hofmann & Gavin, 1998; Kozlowski & Klein, 2000), so emergent models can be a rich source of theoretical development. It is important to obtain a “pure” estimate of the within-firm relationship between turnover and ROI when testing whether the between-firm relationships significantly differ from the within-firm relationships. As we elaborate shortly, unobserved heterogeneity due to the nested data and time should be removed from the within-firm estimates in these models.

Cross-Level Interactions

The fourth modeling framework involves cross-level interactions, where a lower-level relationship (Level 1 relationship) is moderated by a characteristic of the higher-level entity (Level 2 variable). In our running example, one might hypothesize that “the within-firm turnover-ROI relationship is moderated by industry.” In Table 1, we might test whether the relationship between turnover and ROI is stronger in some firms than others and whether the firm-level attribute of industry moderates the relationships. As with tests involving incremental relationships, it is important to remove unobserved heterogeneity due to the entity and time from the within-firm estimates when testing cross-level interactions.

Growth Model (Temporal Hypotheses)

The final variant involves treating the repeated observations as a substantive variable to examine trajectories of change over time (Singer & Willett, 2003). For example, a growth model hypothesis might state that “ROI significantly increases over time and that rates of change over time are related to industry.” In a typical growth model, a new variable would be created indexing time as a vector from 0:(T–1), where T is the number of measurement occasions. In our example, time $T = 0$ would be 2012 and time $T = 5$ would be 2017. A growth model regresses within-firm ROI on the time vector using 1 degree of freedom rather than five time dummy-coded variables. Growth models can test whether ROI increases or decreases for the entire sample and whether the time slope varies among firms. In a growth model, variance associated with time moves from being a source of error to a variable of substantive interest.

Summary of Models

The modeling frameworks are more complex than presented here, and the broader literature contains numerous variants of these prototypes. Clearly, however, panel data are rich in terms of their ability to test different theoretical hypotheses. Critical to our review is that there is no methodological or empirical reason to select one modeling framework over another. Rather, the choice is driven by the theoretical question of interest. In addition to choices about modeling frameworks, researchers in both disciplines tend to use different analytic tools. In our experience, researchers in both macro- and microareas often know one method well but are unfamiliar with alternative models that could be equally appropriate or may offer advantages in certain situations.

Analytic Tools and Foundational Principles

To facilitate comparisons of analytic tools, we first provide a list of terms used within each discipline to help reinforce ideas and provide reference material (see Table 2). Second, we use the data in Table 1 to contrast analytic tools.[5] We note up front that this comparison is simplified. In practice, researchers in both disciplines would refine the statistical models by adjusting for alternative error structures (e.g., heteroscedasticity, autocorrelation) or lagging variables to provide temporal separation (Greene, 2018; Wooldridge, 2010). These types of refinements are important, but they do not change the fundamental principles we illustrate. Third, we discuss how preference for different types of models motivates disciplinary differences in decomposing total variance into within and between components (e.g., estimating intraclass correlation coefficient, ICC, values). We conclude the section by discussing inferential errors.

Analytic Tools

Analyzing within-entity relationships. Examinations of within-entity relationships tend to dominate analyses of panel data. To test a hypothesis that "within-firm collective turnover is negatively related to within-firm ROI," a researcher can use either fixed-effect or random-effect models; however, the two approaches do not necessarily default to the same estimate. The fixed-effect model controls for all omitted variables associated with the higher-level entity (e.g., the firm) and time using regression with dummy-coded variables for firm and time. As such, fixed-effect models can be "expensive" with respect to degrees of freedom, particularly if the data contain a large number of higher-level entities (Greene, 2018; Kreft & de Leeuw, 1998). In our example, 7 degrees of freedom (two dummy-coded variables for firm and five for year) would be used to obtain an estimate of the within-entity relationship between turnover and ROI.

The fixed-effects model removes all unobserved heterogeneity due to firm and time (Bailey, 2017; Kreft & de Leeuw, 1998) and is a consistent estimator of within-firm relationships. Macroresearchers often consider the parameter estimates from this model to be the "gold standard" to which results from other analytic options are compared. Table 3 provides the fixed-effect results for our example data. Firm and time dummy codes are usually of no particular interest;[6] rather, the focus is on the estimated relationships between turnover and ROI.

Not to belabor the obvious, but when fixed effects were omitted and we estimated a simple OLS model regressing ROI on turnover, the resulting estimate was positive (0.247). In contrast, when unobserved heterogeneity due to firm and time is removed, the estimate in Table 3 is negative ($\beta = -0.278$, $p = .03$). This estimate from the fixed-effect model has a clear and precise theoretical interpretation: it represents within-firm change in ROI associated with a 1-point change in within-firm turnover. With respect to theory development, results based on fixed-effect models should leave no ambiguity that relative within-firm effects are being tested.

Table 2
Macro- and Microdefinitions and Macro- and Microdiscrepancies

Term	Definition	Domain Where Commonly Used	Corollaries or Other Examples
General terminology			
Endogeneity	A condition where the independent variable is related to other factors captured by a model's error term	Macro	Threats to validity (micro); common method bias (micro); omitted variable/third variable bias; nonindependence of observations
Panel data	Data that contain multiple responses from the same entity on the independent variable and dependent variable	Macro	Longitudinal, repeated measures, experience sampling methodology (micro)
Mixed-effects model	Model that partitions and estimates variance into within- and between-unit components	Micro	Random-effects model (macro), hierarchical linear modeling (micro), random coefficient model (micro)
Data-related terminology			
Level 1	Lowest level of interest in panel data; represents the multiple responses nested within a higher-level entity (e.g., firm or individual)	Micro	Yearly turnover and yearly ROI from firms are examples of Level 1 variables (firms would be Level 2).
Level 2	Higher level of interest in panel data; the higher-level entity that contains multiple nested Level 1 responses (e.g., firms or individuals)	Micro	Characteristics of the higher-level entity (firm industry; person gender)
Firm fixed effects	Controlling for unobserved heterogeneity between entities using dummy codes for each higher-level entity (e.g., firm)	Macro	Dummy variables for groups (micro); same estimates obtained using demeaning (macro) / group-mean centering (micro)
Time fixed effects	Controlling for unobserved temporal heterogeneity using dummy codes for time	Macro	Dummy variables for time (micro)
Analytical model–related terminology			
Incremental or emergent effects	Testing whether the between-entity relationship significantly differs from the within-entity relationship	Micro	A significant effect for a higher-level variable (e.g., average firm turnover) when the raw lower-level variable is included (yearly turnover) indicates an emergent or incremental effect
Cross-level interactions	Where a higher-level variable moderates the relationship between a Level 1 independent variable and a dependent variable	N/A	In growth models, Level 2 predictors of slope variability between time and the dependent variable represent a cross-level interaction
Temporal hypotheses	Hypotheses that incorporate an element of time (e.g., change; growth; pre- vs. postevent; trend)	N/A	N/A
Specific tests of data and models			
Hausman test	Examines consistency in coefficients between fixed- and random-effects specifications; nonsignificant test indicates random effects are preferred as a result of efficiency	Macro	A significant Hausman test indicates significant ICC values associated with higher-level entities
ICC	Total amount of variance in the dependent variable attributed to Level 2 (higher-level) differences	Micro	Adjusted *r*-square in a null fixed-effect model (no predictors)

Note: ROI = return on investment; ICC = intraclass correlation coefficient.

Table 3
Fixed Effects for Firm and Time From Example Panel Data

	Estimate	*SE*	*t* value	Pr(>|*t*|)
(Intercept)	10.57	1.27	8.30	.00
Turnover	−0.28	0.11	−2.54	.03
factor(Firm)2	8.41	2.30	3.66	.01
factor(Firm)3	−6.98	0.54	−13.00	.00
factor(Year)2013	0.44	0.16	2.70	.02
factor(Year)2014	0.77	0.23	3.42	.01
factor(Year)2015	0.84	0.30	2.76	.02
factor(Year)2016	0.91	0.39	2.34	.04
factor(Year)2017	0.83	0.49	1.69	.13

Random-effects models are a common alternative to fixed-effects models. One potential source of confusion is that the larger class of random-effects models are referred to using a variety of names, including "hierarchical linear model or HLM" (Raudenbush & Bryk, 2002), "random coefficient model or RCM," and "mixed-effects model" (Pinheiro & Bates, 2000). The term "mixed-effects model" is one of the more common, although we use "random effects" as this term is generally familiar to macro-oriented researchers. An additional source of confusion is that random-effects models (unlike OLS models) can be estimated using a variety of different algorithms. Microresearchers are familiar with an option to choose between full maximum likelihood and restricted maximum likelihood, but many other estimation methods exist (e.g., Amemiya, 1971; Nerlove, 1971; Swamy & Arora, 1972; Wallace & Hussain, 1969). Later, we show that some of these algorithms implemented in programs such as Stata and R can produce model estimates that vary dramatically, particularly in extreme examples such as the small data presented in Table 1.

A basic random-effects model includes a random intercept for the higher-level entity. To understand the model's logic, consider a null model where the dependent variable is predicted by only an intercept. The null model estimates a common intercept (i.e., overall mean) and also estimates the variability of the means associated with the higher-level entities. Rather than return parameter estimates for each higher-level entity, a random intercept model uses 1 degree of freedom to account for differences among higher-level entities (the variance estimate τ_{00}). The fact that the random-effects model uses only 1 degree of freedom to model firm differences, *for any number of higher-level entities*, is an important difference from the fixed-effects model. The other major difference is that the random-effect specification can include Level 2 characteristics in the model (e.g., characteristics of the firm) while simultaneously accounting for Level 2 variance.

While a random-effects model is potentially more efficient, it is not by default consistent. Furthermore, the random-effects model assumes that unmeasured variables are uncorrelated with the included independent variables and does not automatically remove all variance due to the higher-level entity or to time. For instance, if we estimate a random-effects model using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015), we can specify random intercepts for both time and firm and obtain an estimate of turnover of −0.412 (quite different from the gold standard of −0.278). To better approximate our fixed-effects model, a random-effect model can include time fixed effects while keeping the random intercept for firm. In such a model, the estimate for turnover using Table 1 data is −0.235 versus the gold standard of −0.278. Practically speaking, researchers are then faced with determining whether the efficiency of the random-effect model is to be preferred over the fixed-effects model's consistency (Greene, 2018).

In the macro-oriented literature, parameter estimates from a random-effects model are routinely compared to parameter estimates from a fixed-effects model using the Hausman (1978) test. A significant Hausman test is a heuristic to favor the consistency of the fixed-effect model. In our example, a Hausman test provides no reason to reject the random-effects model.[7]

When random-effect models fail the Hausman test, there are three options that recover the "gold standard" estimate for both the parameter estimate and the standard error in a random-effect model. One alternative is to demean (i.e., group-mean center) the independent variable and include its group mean as an additional predictor along with the demeaned independent variable (Allison, 2005). This "hybrid" approach was highlighted by Certo et al. (2017). A second alternative is to include the group mean of the independent variable as a predictor along with the raw independent variable (Raudenbush, 2009). The third approach is to "demean" the independent variable and omit the group mean. When demeaning, it is advantageous to estimate the model using a random-effects program instead of OLS. As noted by Greene (2018: 188-189), standard OLS uses the wrong term in the denominator of the standard error by failing to subtract the number of groups that were used to demean the variables. Random-effect models, however, correctly calculate standard errors. Ultimately, the key point to note is that if the research focus is entirely on understanding within-firm relationships, *then either fixed-effects models or any of the three random-effect variants (hybrid, add group mean, or demean) provide identical results*.

Analyzing between-entity relationships. Recall that these relationships involve testing questions such as whether average firm turnover is related to average firm ROI. In balanced data where each higher-level entity has the same number of yearly observations, the most direct way to estimate between-entity relationships is to calculate means for each higher-level entity on the variables of interest (i.e., average across years) and estimate a regression equation. In our example, an OLS model regressing three values of firm average ROI on three values of firm average turnover returns a parameter estimate of 0.260 with a standard error estimate of 0.162. Greene (2018: 192) refers to this analysis as relying on the between-groups estimator.

If panel data are not balanced, there are advantages to estimating models that adjust for differences in the number of observations. The hybrid approach using a demeaned independent variable along with the group mean of the independent variable proposed by Allison (2005) and highlighted by Certo et al. (2017) returns a between-group estimate adjusted for differences in the number of within-group measures (Bliese et al., 2018; Raudenbush & Bryk, 2002). A random-effects model based on the data in Table 1 regressing ROI on demeaned turnover, firm average turnover, and year fixed effects returns a parameter estimate of 0.260 and a standard error of 0.162 for firm average turnover—both the parameter estimate and the standard error are identical to the OLS results based on three firm means because the data are balanced.

Table 4
Emergent or Incremental Test of Within-Firm and Between-Firm Turnover

	Value	*SE*	*df*	*t* value	*p* value
(Intercept)	3.06	3.03	9	1.01	.34
Turnover	−0.28	0.11	9	−2.54	.03
Firm average turnover	0.54	0.20	1	2.75	.22
factor(Year)2013	0.44	0.16	9	2.70	.02
factor(Year)2014	0.77	0.23	9	3.42	.01
factor(Year)2015	0.84	0.30	9	2.76	.02
factor(Year)2016	0.91	0.39	9	2.34	.04
factor(Year)2017	0.83	0.49	9	1.69	.13

Using Table 1 data to estimate the same hybrid model (regressing ROI on demeaned turnover, firm average turnover, and year fixed effects) in OLS instead of in a random-effects model provides an unbiased between-group estimate at 0.260; however, the standard error for the between-group estimate is severely downwardly biased at 0.042, resulting in a highly inflated *t* value (Raudenbush & Bryk, 2002). Thus, while the data in Table 1 underlying the hybrid approach can technically be used to estimate an OLS version of the hybrid model, doing so will lead to erroneous conclusions unless standard errors are corrected.

In the data structure in Table 1, within-firm ROI is regressed on average firm turnover, giving the appearance that average firm turnover "explains" within-firm variance. We emphasize that the correct interpretation of the hybrid model is that "average firm turnover explains average firm ROI" rather than "average firm turnover explains within-firm ROI" (see Bliese et al., 2018; LoPilato & Vanderberg, 2015; Preacher, Zyphur, & Zhang, 2010; Raudenbush & Bryk, 2002).

From a theoretical perspective, the challenge with interpreting the between-entity effect associated with the hybrid approach is that the between-entity value inherits the lower-level relationship. Recall that because of this inheritance, proposing that group means are related may not be particularly informative if we already expect within relationships to be present. In contrast, in our example, we can clearly see that the estimates differ across levels. The gold standard for the within relationship is −0.278, and the between relationship is of an opposite sign at 0.260. This difference represents an incremental or emergent relationship across levels.

Analyzing incremental or emergent relationships. A hypothesis involving an incremental or emergent relationship might state that "the between-firm relationship between firm average turnover and firm average ROI significantly differs from the within-firm relationship between turnover and ROI." In the incremental model, the Level 1 variable (turnover) is in raw form, so in Table 1 we would regress ROI (column 5) on turnover (column 4) and average turnover (column 6). Table 4 shows results from a random-effects model regressing ROI on raw within-firm turnover, average turnover, and year fixed effects (e.g., time dummy variables). The model returns the gold standard estimate of the within-firm effect (–0.277 rather than −0.278 due to minor rounding error); however, the estimate for the between effect is 0.537 with a standard error of 0.195. The value of 0.537 indicates that the between-entity effect involving relationships between firm means of 0.260 is 0.537 stronger than the within effect of −0.278 (within rounding error). Notice that even with only three Level 2 entities, the incremental test is statistically significant in this extreme example.

Other tools can be used to test incremental or emergent relationships. For instance, in balanced panel data, a researcher could estimate a model based on group means and formally test whether the observed value significantly differed from the value obtained based on a model with fixed effects for firm and time. In practice, though, the random-effect model provides a convenient way to test whether relationships differ across levels. Note again that the data structure underlying the incremental test makes it possible to estimate a simple OLS regression model, but doing so would produce significant bias in standard errors.

Analyzing cross-level interactions. Both the fixed-effect and the random-effect model can include cross-level interaction terms to test hypotheses such as "whether the within-firm turnover-ROI relationship is moderated by industry." In our example, if our Level 2 industry variable was "healthcare" (value of 1 or 0), the cross-level interaction would be captured by including an interaction variable that multiplied "healthcare" by each within-entity turnover value. To a microresearcher, the fixed-effect model would appear odd because it would lack a main effect for healthcare, as the main-effect variance is already captured by the firm fixed effects (McNeish & Kelley, 2019). Nonetheless, the hypothesized form would be familiar to all researchers by expressing that the strength of the relationship between turnover (Level 1) and ROI (Level 1) varies by the Level 2 industry attribute of healthcare.

A random-effects model that included time fixed effects and a random intercept for firm would provide the exact same estimate of the cross-level interaction term. A random-effect model, however, has two potential advantages. First, the random-effects model could include the Level 2 main effect testing whether average levels of ROI between firms was related to being a healthcare firm. Second, a common step recommended in estimating random-effects models is to determine whether the variance term associated with the slopes (τ_{11}) is significant prior to adding a cross-level predictor (Bliese & Ployhart, 2002; Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). A significant variance term suggests that the relationship between two lower-level variables varies more than would be expected by chance. In our running example, significant slope variance would suggest that the strength of the relationship between turnover and ROI is stronger in some firms than in others. The absence of significant variance does not preclude testing for a cross-level interaction (Snijders & Bosker, 1999), but the presence of a significant variance estimate supports modeling cross-level interactions. We note that removing unobserved heterogeneity for both firm and time is complicated in testing interactions, so we direct readers to Shaver (2019) for details on specifying interaction terms to remove unobserved heterogeneity.

Analyzing growth models (temporal hypotheses). In growth models, trajectories over time are of substantive interest and time is included as a 1 degree of freedom vector or several vectors to capture nonlinearity and/or discontinuities (e.g., Singer & Willett, 2003). In other words, time becomes a special form of a within-entity predictor that (in balanced data) takes on the same values for each higher-level entity. Recall that a growth model hypothesis might state that "ROI significantly increases over time and that rates of change over time differ across firms."

When time is treated as a vector rather than as T-1 dummy codes, the models can still contain unobserved heterogeneity due to time. For instance, if we specify an OLS model regressing ROI on turnover, firm fixed effects, and time as a vector, the within estimate for turnover is −0.233 rather than −0.278. In the trade-off, however, we now observe that ROI is increasing by 0.20 each year (nonsignificant in our small data set). Growth models are most frequently discussed from a random-effect modeling framework (e.g., Raudenbush & Bryk, 2002; Singer & Willett, 2003), but because time is simply another form of a within-entity predictor, a growth model could be specified and tested in an OLS model with firm fixed effects. That said, random-effect models benefit from the ability to include main effects associated with the higher-level entity and to formally test whether time-related trajectories significantly vary across higher-level entities. That is, a random-effects growth model is a form of a cross-level model that can include an estimate of the slope variance (τ_{11}) for time across higher-level entities.

Variance Decomposition

Our summary shows that both fixed-effect and random-effect models are often viable alternatives for analyzing panel data. One key difference, however, is that researchers using random-effect models tend to rely more on variance decomposition as a step to inform subsequent analyses. More specifically, microresearchers using random-effects models often estimate the total amount of variance in the outcome that can be attributed to the higher-level entities in the form of an ICC (Bliese, 2000; Hox, 2002; Raudenbush & Bryk, 2002). The ICC is estimated using a null model—a model without any predictors that partitions variance into within- and between-group components.

In the example data, the ICC for the null model is .97, suggesting that 97% of ROI's total variance can be attributed to firm-level differences. It would be helpful to share that only 3% of the data's total variance was within firm to help readers put findings into context, particularly if the analytic approach explained only within-firm variance. In panel data where time is not of substantive interest, an ICC conditional on time is informative. Time would be included as a fixed effect along with a random intercept for higher-level entities. In our example, the conditional ICC with a random effect for firm and a fixed effect for time is .998.

In a fixed-effects model, the null model ICC can be approximated by examining the adjusted r-square for a model containing only fixed effects for higher-level entities (Bliese, 2000). In our example, the adjusted r-square is .95 (marginally lower than .97). The adjusted r-square from a fixed-effects model with both year and firm fixed effects is .997, which is similar to the conditional ICC value of .998 from the random-effects model.[8]

Summary of Foundational Principles of Analytic Tools

While the fixed-effect specification returns a "gold standard" estimate, its use comes at some cost with respect to formulating and testing theory. The main limitation is that models cannot include main effects for higher-level variables. As such, a fixed-effects model also precludes testing incremental or emergent models. It is not clear how often emergent effects exist because they are rarely tested; however, Certo et al. (2017) noted that 66% of studies they reviewed reported a significant Hausman (1978) test. The significant test statistic indicates one or more parameter estimates differ between a fixed- and random-effect model. These results suggest a high likelihood of unexplored emergent effects.

The fixed-effect specification may also limit research on time-related effects. Including time fixed effects precludes testing whether attributes of certain time periods (e.g., economic growth years) are related to the dependent variable. In addition, including time fixed effects limits researchers' ability to frame theoretical questions about change trajectories. For instance, random-effects models are arguably better suited for examining exogenous shocks and modeling attributes from higher-level entities that predict different trajectory responses (e.g., Bliese & Lang, 2016; Kim & Ployhart, 2014; Singer & Willett, 2003). We emphasize that theory is the main determinant in choosing how to model time within panel data, but if time characteristics are of interest, time fixed effects are likely overly restrictive.

Inferential Errors

The illustrative data in Table 1 show how panel data can produce a wide variety of different results. The models provide opportunities to test different theoretical hypotheses, but care must be taken to avoid inferential errors. Even simple forms of multilevel data without a repeated measures component (e.g., individuals nested within larger collectives) are known to produce results prone to inferential errors (see Thorndike, 1939).

When relationships differ across levels, researchers can inadvertently commit either the ecological or the atomistic fallacy (Diez-Roux, 1998). The ecological fallacy occurs when aggregate-level results are used to make inferences about lower-level relationships. In our example involving turnover and ROI, an ecological fallacy would involve using the estimated between-entity result to make a within-firm inference. Recall that the estimate based on firm means was 0.260, whereas the within-firm estimate was −0.278. In concrete terms, the fallacy here would involve encouraging a firm to increase turnover (a within change) to increase ROI based on the 0.260 finding that firms with high turnover (on average) have high ROI (on average).

The atomistic fallacy is less frequently discussed but involves making a higher-level inference from lower-level results. In our example, we commit the atomistic fallacy if we use the within-firm estimate of −0.278 to conclude that firms who historically have high turnover (averaged over many years) are likely to have lower average ROIs. The fallacy reflects that we are attempting to explain between-firm patterns from within-firm relationships. At the between level, a firm's average turnover likely reflects Level 2 factors such as industry (which are also related to ROI), so using within-firm results to make inferences represents a serious fallacy.

It is tempting to try to identify the "right" correlation, but from an inferential perspective, the "right" correlation is determined by the level of inference. One way to help avoid either fallacy is to frame hypotheses in language that clearly informs the reader about the level of inference. In our review, we highlight hypotheses that are exemplars in both the micro- and macro-oriented literature. In the following section, we review management research employing panel designs and discuss differences between macro- and microresearch with how data are treated theoretically and methodologically.

Literature Review: Methodological Approaches and Their Implications

We conducted a review of management research published in 2017 and identified articles that used panel data in the following leading journals: *Academy of Management Journal,*

Table 5
Review of Panel Data Articles (2017)

	Macro	Micro	Combined
Studies reviewed by journal			
Academy of Management Journal	23	14	37
Administrative Science Quarterly	8	0	8
Journal of Applied Psychology	0	12	12
Journal of Management	15	5	20
Organization Science	12	0	12
Strategic Management Journal	57	0	57
Total	115	31	146[a]
Sample sizes[b]			
Higher-level entities			
Mean	7,689	844	6,105
Median	522	150	409
Time periods			
Mean	27	19	25
Median	13	8	12
Study design (%)			
Within-entity relationships	43	81	70
Between-entity relationships	57	84	76
Emergent relationships	0	0	0
Cross-level interactions	42	39	41
Growth model (temporal hypotheses)	29	52	34
Time fixed effects	70	6	56

[a]We identified 146 studies within 142 articles (4 articles with multiple studies and samples).
[b]While all studies reported the number of observations used to test hypotheses, we were unable to determine the precise number of higher-level entities (e.g., firms) or time periods for several studies because they were unclear regarding the number of entities or time periods. Thus, our sample size calculations for higher-level entities and time periods are based on 134 and 144 studies, respectively.

Administrative Science Quarterly, Journal of Applied Psychology, Journal of Management, Organization Science, and *Strategic Management Journal*. In total, 142 articles were identified, representing approximately 31% of empirical research. Macro-oriented journals had a higher proportion of articles using panel data than did micro-oriented journals (e.g., 54% of *Strategic Management Journal* empirical articles used panel data compared to only 15% in the *Journal of Applied Psychology*).

For each article, we recorded details such as the specific analytic approaches, wording of hypotheses, and numbers of higher-level entities and longitudinal time periods. Table 5 outlines the number of articles we reviewed by journal, as well as descriptive statistics between domains. We summarize modeling framework and use of analytic tools by discipline.

Macro-Oriented Research

In 2017, we identified 115 macro-oriented articles that used panel data. The mean number of time periods was 27 with a median of 13. The mean number of Level 2 entities reported (typically firms) was 7,689 with a median of 522. We classified articles as between or within based on the methodology employed, given that hypotheses were often ambiguous.

Within-entity relationships. As anticipated, macro-oriented research frequently used the fixed-effect model, which resulted in testing within-firm relationships. Firm fixed effects[9] were used in 44% of macroarticles (vs. 10% using random effects).[10] The Hausman test was used in 29% of these articles. Furthermore, these studies routinely included fixed effects for time (82%), resulting in 37% of articles that included both firm and time fixed effects. For a study with the median number of higher-level entities (522) and time periods (13), 533 dummy codes would be needed to fit the full fixed-effect model. This suggests that macro-oriented research might benefit from using variants of the random-effect models (e.g., demeaning) with random intercepts for firms and fixed effects for time to obtain the same gold standard estimate.

In our review, we identified several macro-oriented articles that did particularly well in discussing the results of the empirical analysis in a way that was aligned to the use of fixed-effects models. For example, Williams, Chen, and Agarwal (2017) examine the effect of new top management team member experience and outsider status on firm growth. Williams et al. noted that because they used a fixed-effect model, the effects were the result of "changes within firms rather than variance between firms" (1400). Many studies, however, were ambiguous as to whether they were theorizing, hypothesizing, or testing within-firm relationships.

As noted earlier, macro-oriented research commonly used time fixed effects associated with the unit of observation (e.g., year dummies with firm-year observations) when there were no time-related hypotheses (76% of articles used time fixed effects). Many of the remaining studies attempted to control for time in other ways, such as controlling for year when observations were monthly or seasonal (e.g., Tan & Tan, 2017) or using a dummy code to indicate a time period with an exogenous event (e.g., Xu, Tihanyi, & Hitt, 2017).

Between-entity relationships. In total, 57% of macro-oriented studies tested between-entity relationships. As mentioned previously, random effects were not frequently used (10%). The few that did were often very clear about the reasons for using random effects instead of fixed effects, such as when using a time-invariant predictor (e.g., Barthélemy, 2017; Lee, Hwang, & Chen, 2017). Alternatively, many studies used cluster robust standard errors, which account for within-entity heterogeneity due to repeated observations over time but do not eliminate between-entity heterogeneity (McNeish & Kelley, 2019), to test for between-entity relationships.

Incremental or emergent relationships. None of the macro-oriented studies we reviewed directly tested for emergent or incremental relationships using methods familiar to micro-oriented scholars. Overall, 42% of the 24 articles that used the Hausman test found significant test statistics. The existence of a significant Hausman test illustrates a lack of consistency in parameter estimates between fixed- and random-effects models; this suggests that the nature of the model's relationships within entity differs from the between-firm relationship. Thus, these articles may have missed emergent effects and interesting theoretical contributions.

For example, Haynes, Campbell, and Hitt (2017) explored whether CEO greed is negatively related to shareholder return. The article states that a significant Hausman test supported the fixed-effects model as the most appropriate modeling approach. When interpreting the article's findings, the use of CEO-firm fixed effects means that higher levels of a given CEO's greed are associated with lower shareholder wealth during the given CEO's tenure at the firm. Thus, a recommendation might be that firms should take measures that decrease or limit their CEO's greed. The use of a CEO-firm fixed-effect model, however, does not allow us to draw inferences regarding whether different levels of greed between CEOs negatively influence shareholder wealth across firms. Our prior suggestions regarding the use of a random-effect model to explore emergent effects would allow an empirical examination of whether the between-entity (CEO-firm groupings) relationship is similar to that of the within-entity relationship.

Cross-level interactions. We found 42% of macrostudies proposed and tested cross-level interactions (35% of these studies in fixed-effects models). For example, Chin and Semadeni (2017) hypothesize that CEO and compensation committee liberalism interact to predict pay equality within top management teams. CEO liberalism was measured using political contributions prior to CEO appointment and did not vary within firm over time but did vary between firms. Additionally, several studies used time-invariant Level 2 moderators (e.g., industry). For example, Deb, David, and O'Brien (2017) examined whether the within-firm relationship between cash and performance was moderated by industry characteristics (i.e., R&D intensity).

Growth model (temporal hypotheses). Our review found that few macro-oriented hypotheses explicitly involve the time component. Overall, 29% of macroarticles hypothesized about time; however, time was rarely the article's primary focus. Thus, time-related hypotheses likely represent a smaller proportion of total hypotheses. Macrostudies rarely used growth models (see Williams et al., 2017, for an exception). Instead, some studies used dummy codes to examine changes before and after an event (e.g., Eberhart, Eesley, & Eisenhardt, 2017; Singh, Mahmood, & Natarajan, 2017). Appropriately, these articles used dummy variables to indicate pre- and postevent years rather than time fixed effects. Our review also found that survival analysis and hazard models were relatively common (e.g., Greve & Man Zhang, 2017; Pontikes & Barnett, 2017). As noted, these approaches are outside the scope of our discussion.

Micro-Oriented Research

We identified 31 studies across 27 micro-oriented articles that used some form of panel data. The mean number of time periods was 19 with a median of 8. The mean number of Level 2 entities (often persons but sometimes teams) was 844 with a median of 150. To examine the substantive questions pursued with panel data, 19 of the 31 studies (56%) used experience sampling methodology (ESM). ESM research seeks to understand response variation over relatively short periods of time: individuals are often assessed over multiple days or weeks and provide measures of independent variables and dependent variables daily. Within an ESM design, participants usually complete a large survey at the beginning of a study to assess time-invariant (Level 2) variables such as personality traits (e.g., conscientiousness).

Micro-oriented research reported ICC values in 45% of the studies. In ESM designs, the ICC is often presented as the percent of variance within the individual (1-ICC; Matta, Scott, Colquitt, Koopman, & Passantino, 2017, Table 2). In ESM designs, the 1-ICC value indicates whether the dependent variable of interest has a meaningful level of change over time.

Within-entity relationships. Within-entity relationships were examined in 81% of the studies. Micro-oriented researchers predominately use variants of random-effects models when examining within-entity relationships (72% of articles). We consider multilevel structural equation models (SEMs) to be a form of random-effect models because multilevel SEM can return intercept and slope variance estimates and include Level 2 predictors. Only 2 studies (6% of the 31 studies) used fixed-effect specifications, and the remainder used other approaches.

Researchers group-mean centered (demeaned) within-person variables of interest in 17 of the 22 (77%) random-effects models identified. The percentage of researchers using group-mean centered variables is probably somewhat higher than reported because several cases in the literature review were ambiguous, and group-mean centering is less common when the predictor is a variant of a growth model, which was the case in three articles using random effects.

In contrast, only 2 of 25 (4%) micro-oriented articles included fixed effects for each panel time period, while 2 studies estimated nonlinear time trajectories likely approximating time fixed effects. As a consequence of not accounting for unobserved heterogeneity associated with time, the random-effects model estimates reported in the micro-oriented literature are not entirely on par with the "gold standard" fixed-effects estimates typically reported in the macro-oriented literature. In sum, *microresearchers using group-mean centering are being as effective in dealing with endogeneity due to pooled or nested data as are macroresearchers using fixed-effects models but are less consistent in addressing endogeneity due to time.*

Several authors who used group-mean centering also offered exemplary hypotheses that convey a focus on within-person variance. Ilies, Liu, Liu, and Zheng phrased their Hypothesis 1 as "daily work engagement will be positively associated, within individuals, with daily family satisfaction such that employees will be more satisfied with their family lives on days when they are more, as compared to less, engaged at work" (2017: 959). Some studies were also clear that focusing on within-person relationships offered a novel theoretical contribution (e.g., Woolum, Foulk, Lanaj, & Erez, 2017). Not all published micro-oriented studies displayed similar clarity.

Between-entity relationships. Between-entity relationships were examined in 42% of articles, often involving a person-level variable related to multiple measures of an outcome. For instance, Ramarajan, Rothbard, and Wilk proposed that "Identity conflict is negatively related to performance in tasks that involve interpersonal interactions" (2017: 2211). Identity conflict was assessed via survey and linked to 4 months of sales performance data. As a Level 2 variable, identify conflict was effectively related to the average sales performance over 4 months. It is relatively common for ESM designs to include person-level predictors as main effects.

Incremental or emergent relationships. None of the micro-oriented literature we reviewed formally specified emergent effects across levels. Spieler, Scheibe, Stamov-Roßnagel, and Kappas (2017), however, ran supplemental analyses testing whether average and day-level effects were consistent in their results. In addition, the correlation tables from other studies (Beck, Scholer, & Hughes, 2017; Y. Liu, Song, Koopmann, Wang, Chang, & Shi, 2017) appeared to show evidence of emergent relationships because several correlations at the between-person level differed strongly in absolute terms from the corresponding within-person relationships.

Table 5 from Beck et al. (2017), in particular, showed a large number of incremental or emergent effects. For instance, at the within level on a trial-by-trial basis, participants' ratings of frustration were unrelated to task performance ($r = -.04$), but participant average frustration was related to average task performance ($r = -.25$). Similarly, at the within level, participants' ratings of enthusiasm were unrelated to task performance ($r = -.06$), but participant average enthusiasm was positively related to average task performance ($r = .19$). Beck et al. did not formally test these differences because their model was a variant of the hybrid model with the Level 1 variable group-mean centered. In addition, hypotheses were somewhat vague about levels (Hypothesis 3b: "Following the disturbance, frustration will be positively related to goal commitment, effort, and task performance"; Beck et al., 2017: 1111). Finally, Beck et al. used evidence at either level as support for hypotheses, for example, saying that "frustration was significantly negatively related to goal commitment (at both levels of analysis) and task performance (at the between-person level of analysis only)" (1118). We are not trying to be critical of Beck et al., but we do believe that differences across levels such as those they found provide rich opportunities for theory development. We also believe that hypotheses should clearly identify whether within or between effects will be used to provide evidence confirming or disconfirming hypotheses.

Cross-level interactions. We found 11 articles that examined cross-level interactions. Five tested for slope variance and provided estimates via tables, text, or both. One article (W. Liu, Song, Li, & Liao, 2017) stopped testing for a hypothesized cross-level interaction as a result of a nonsignificant slope variance (even though they could have continued the test). In general, micro-oriented articles were clearer when specifying that a moderating variable was a Level 2 variable.

Growth model (temporal hypotheses). Fourteen of the 27 articles (52%) proposed and tested some form of a temporal hypothesis. Eight of the 14 studies employed time as a vector (1 degree of freedom) or as several time-related vectors. For instance, Stewart, Astrove, Reeves, Crawford, and Solimeo (2017) modeled time as discontinuous trajectories. Other studies used change in the predictor to predict change in the outcome (e.g., Taylor, Bedeian, Cole, & Zhang, 2017).

Implications and Recommendations

As integration has begun between macro- and microdomains, reconciling methodological differences remains one of the more difficult and persistent challenges (Wright et al., 2014). In this review, we address such challenges by focusing on how both domains approach panel data. Regardless of domain, panel data routinely have two distinct forms of unobserved heterogeneity—that due to higher-level entities and that due to time. Failing to account for these two sources of variance can produce biased statistical results. Both macro- and micro-oriented researchers understand the importance of addressing these issues (though perhaps less so with time in microresearch) but do so using different terms and statistical tools.

Table 6
Summary Implications and Recommendations

Random-effects models can provide the same within-group estimate as fixed-effects models when the following conditions hold:
Demean (group-mean center) the independent variable.
Include Level 2 means of the independent variable along with raw independent variable values.
Demean the independent variable and include the Level 2 mean of the independent variable.
Consider predictors at higher levels and emergence:
Test whether within-firm and between-firm relationships *differ* by including Level 2 means of independent variable and raw independent variable values.
Test whether between-firm relationships are *significantly different from zero* by including Level 2 means of independent variable and demeaned independent variables (i.e., group-mean center).
Use and interpretation of variance terms:
Always report ICCs or a variant as a descriptive statistic when using panel data.
Calculate ICC to determine variance explained within and between firms and support theory building regarding the level at which effects occur. When the ICC is close to 1, there may be little value to interpreting within-group effects.
When testing cross-level interactions, perform tests of slope variance. Even when slope variance is insignificant, it is still appropriate to test for cross-level interactions.
Model time appropriately:
When time is not of theoretical interest, control for time with fixed effects to eliminate all unobserved temporal heterogeneity.
When time is of theoretical interest, model time substantively (e.g., trends) and do not employ fixed effects for time periods.
Expand theoretical approaches to incorporate time as a meaningful predictor in future research to understand how predictors of interest change over time.
Handle Hausman tests with care:
A significant Hausman test does not require a fixed-effects specification but does require modifying the predictor of interest to account for between-firm variation.
As an omnibus test, incorporating multiple predictors of no theoretical interest may bias the Hausman test statistic.
Identify and understand the default algorithms used in random-effects packages and how they might affect Hausman test statistics. Different algorithms can return coefficients that vary widely.
Carefully word hypotheses to link to analytical choices:
When using fixed effects or group-mean centered variables, word hypotheses to indicate the tests are within the higher-level entity.
When testing higher-level (Level 2) variables, word hypotheses to indicate that the predictor of interest (average) is related to the average of the dependent variable.

Note: ICC = intraclass correlation coefficient.

Reconciling methodological differences is important because statistical choices frame theoretical questions. When scholars in one domain are unfamiliar with terms and methods in the other, the research has the appearance of testing different theoretical questions in a potentially inappropriate manner. A scholar trained in human capital theory and familiar with articles using fixed effects will likely be skeptical of a manuscript testing human capital theory with group-mean centered random-effect models. Likewise, a scholar trained in ESM and familiar with random effects may be skeptical of a manuscript employing dummies for unit and time.

Our overall goal has been to facilitate cross-disciplinary communication by helping both macro- and micro-oriented researchers understand each other. We conclude by summarizing and discussing six themes identified in Table 6: (1) random-effects and fixed-effects models can provide the same within-group estimate; (2) random-effects models provide the option to include predictors at higher levels and test theory involving incremental or emergent effects; (3) macro- and micro-oriented researchers can likely make better theoretical use of variance estimates; (4) both macro- and microresearch might benefit from considering time as a substantive variable but at a minimum, need to control for its effects in all cases; (5) care should be taken when interpreting the default values with the Hausman test; and (6) both macro- and microresearch would benefit from aligning hypothesis wording to analyses.

Fixed Effects and Random Effects: Not So Different Really

While terminology used in macro- and micro-oriented research differs, both areas see the need to use statistical methods that account unobserved heterogeneity in panel data. Overwhelmingly, we found researchers in both fields use methods that account for the variance associated with higher-level entities. Macro-oriented articles tended to be more systematic than micro-oriented research with respect to accounting for variance associated with time.

A main difference in approaches centers on the decision to use fixed-effects or random-effects models. Fixed-effects models remove all endogeneity associated with the higher-level entity and time, so their appeal is understandable. Unfortunately, the approach comes with the cost of potentially removing systematic variance that may be substantively and theoretically important (Certo et al., 2017). Random-effects models do not by default return the "gold standard," but they can with modest effort by including a random intercept for the higher-level entity and fixed effects for time and either (a) demeaning the independent variable, (b) including the higher-level entity mean of the independent variable of interest and the raw variable, or (c) both demeaning and including the higher-level entity mean in the model (Bell & Jones, 2015; Certo et al., 2017; Raudenbush, 2009).[11] Our takeaway is that both fixed- and random-effects models are equally viable options for modeling within-entity relationships as long as researchers using random-effects models specify the appropriate model.

At a practical level, also note that random-effects models may offer some advantages for prediction and forecasting. Raudenbush and Bryk (2002) discuss how empirical Bayes estimates help provide firm-specific values. For instance, if we estimated a growth model and wanted to obtain a value that represented firm-specific growth, we would favor an empirical Bayes estimate of the growth over an estimate based on OLS regression. The empirical Bayes estimate would "normalize" extreme OLS estimates of slopes (e.g., Chen, Ployhart, Thomas, Anderson, & Bliese, 2011).

Predictors at Higher Levels and Emergence

The inability to include characteristics of the higher-level entity in a fixed-effects model may limit theory development and testing. Part of understanding a phenomenon more broadly involves asking questions that involve contextual or higher-level influences. The ability to model incremental or emergent effects of what is ostensibly the same variable (e.g., turnover each year and firm average turnover) in the same model opens up considerable research potential. Formally proposing conceptual reasons and testing whether within-entity relationships between two variables differ in magnitude from between-entity relationships allows researchers to develop a more enriched understanding of phenomena. Correlation tables contrasting effects across levels in the microliterature and significant Hausman tests in macroresearch suggest that emergent effects are relatively common. Models that include variants of the same variable at different levels can also help researchers avoid inferential errors such as the atomistic and ecological fallacies. Observing an emergent effect such as that for firm average turnover in Table 4 serves as a reminder that within-firm effects do not generalize between firms and vice versa.

In macro-oriented research, between-firm relationships may be just as relevant as within-firm relationships. One of strategic management's fundamental questions is why some firms outperform others (Mahoney & Qian, 2013). For example, researchers have identified the risk-return paradox of strategic management (e.g., Fiegenbaum & Thomas, 1988), whereby greater risk yields poorer performance. In testing this relationship, we could articulate theoretical mechanisms that operate both within firm and between firms. Within firms, we may explore how levels of firm risk affect firm performance over time. In an emergent or incremental model, we might be interested in testing whether greater average risk is associated with greater average performance, consistent with the notion that firms who take on higher average levels of risk have different performance from those with lesser average risk. It is entirely possible that the between-firm relationship between average risk and performance differs from the within-firm relationship. However, if fixed-effect models are used as a default, between-firm inferences cannot be examined and contrasted to within-firm relationships. Thus, research using panel data in strategy may not fully help to answer some of the field's important questions. Indeed, researchers who study between-firm theories may implicitly commit the atomistic fallacy by believing results based on fixed effects support theory involving between-firm relationships.

More Use and Interpretation of Variance Terms

Both macro- and micro-oriented researchers could likely make better use of variance terms associated with panel data even before substantive predictors are included. Micro-oriented researchers often reported ICC or 1-ICC values associated with how much variance in the dependent variable was associated with higher-level entities. In panel data, a conditional ICC that partitions variance into within and between components in a model that includes fixed effects for time would be informative, particularly if time was not of substantive interest.

In the fixed-effects model, researchers can estimate adjusted r-square values in the baseline model including firm and year fixed effects as an analog to the *conditional* ICC from the random-effects model. As noted by Certo et al. (2017), ICC values have notable implications for theory development and drawing inferences. Values convey important information about the main correlates of the phenomena. ICC values near zero suggest that most of the variance occurs within the higher-level entity and that the outcome of interest will be related to within-firm, time-varying covariates. In contrast, an extremely high ICC suggests that the outcome is primarily a function of between-firm differences and that there may be little value in investigating within-firm change. In short, the ICC provides important theoretical information about the main correlates of the phenomena of interest.

Providing information about the ICC can also help align analyses and inferences. A researcher reporting an ICC of .997 may find a significant within-group effect as we did in our models with fixed effects and year dummies. Part of a balanced discussion that spurs theory development would need to mention that the vast majority of the variance exists between firms. In sum, ICC estimates provide an important baseline for understanding statistical results and should be commonly reported along with other descriptive statistics.

Variance terms are also important in understanding cross-level interactions and slopes in growth models. The ability to examine slope variance of a predictor enhances theory testing. For instance, consider the case where a researcher fails to find a significant predictor for why the strength of the relationship between within-firm turnover and within-firm ROI differs. If significant slope variance is present, a researcher would presumably encourage future research to find predictors of the variance. In contrast, if tests of slope variance are not conducted, then there may be little reason to attempt to identify alternative predictors.

Model Time (Appropriately)

Including time fixed effects removes all temporal unobserved heterogeneity; however, doing so relegates time to the status of an uninteresting control variable. That is, use of time dummies guarantees that attempts to model time substantively (e.g., examining trends, change, or growth) will fail because all temporal variation is removed by the dummy variables. There are clearly situations where including time fixed effects is appropriate and strengthens research conclusions. For example, studies published in the micro-oriented literature will be able to draw stronger inferences if they remove temporal variation, particularly if the independent variable and dependent variable are measured concurrently (representing a cross-sectional design repeated across numerous days). For instance, Sonnentag, Pundt, and Venz (2017) hypothesized that daily health motives and daily snacking would be related, but inferences about the strength of the relationship between daily health motives and daily snacking would have been stronger if the research eliminated unobserved heterogeneity associated with time.

While we clearly see opportunities for micro-oriented researchers to improve how they deal with endogeneity due to time, we also caution researchers against simply defaulting to include time fixed effects. Thinking of time simply as a source of endogeneity limits opportunities to incorporate change into our theoretical and statistical models. Considering time more broadly can continue to advance both macro- and microresearch.

Handle Hausman (Tests) With Care

We share the perspective of Certo et al. (2017) that a significant Hausman test need not preclude using a random-effects model. Instead, the test result informs users of the need to modify the random-effects model. We also encourage users to exercise caution when applying and interpreting the Hausman test in two respects. First, the Hausman test is often run on models with many predictors and represents an omnibus test of whether the coefficient estimates for all predictors are consistent across the fixed- and random-effect specifications. It may be more informative to test single variables (see Schunck, 2013). That is, the Hausman test might be more meaningful when it includes only the predictors of interest. Second, the choice of random-effect algorithms in the plm package in R (Croissant & Millo, 2008) can provide substantially different random-effect estimates, at least in our extreme data. Our main point is simply to encourage researchers to understand the defaults in their statistical programs and to make sure that they are not using the Hausman test in ways that inhibit theory development and testing.

Carefully Word Hypotheses to Link to Analytical Choices

In our review, we highlighted instances in both the macro- and micro-oriented literature where authors worded hypotheses in ways that were particularly precise and clear. One implication for our decision to highlight exemplary hypotheses is that many hypotheses and, thus, related discussion of results, in both the macro-oriented and the micro-oriented literatures were vague with respect to level. In research relying on fixed-effects models, researchers can specify that the statistical tests determine whether within-entity independent variables are related to within-entity dependent variables (e.g., "higher levels of turnover within firms will be negatively related to within-firm ROI"). Researchers using random-effects models with group-mean centered time-varying independent variables can use the same wording. Researchers who incorporate attributes of the higher-level entity, however, should specify that the models are predicting averages of the dependent variable (e.g., "healthcare firms have higher average levels of ROI than other types of firms" or "employee conscientious is positively related to average performance").

Researchers who estimate models including variants of the same variable (yearly turnover and average firm turnover) need to carefully consider whether they are proposing a test of an emergent relationship (e.g., "the relationship between average firm turnover and average firm ROI is *stronger* than the relationship between within-firm turnover and within-firm ROI") or simply testing whether the aggregated independent variable is related to average levels of the dependent variable (e.g., "average turnover in a firm is related to average firm ROI") and use the appropriate model (Hofmann & Gavin, 1998).

Conclusion

Panel data contain rich sources of information that can be leveraged to understand relationships within and between entities and over time. Our review attempts to facilitate communication between macro- and micro-oriented researchers by detailing the methodological frameworks and analytic tools used by both disciplines when approaching panel data. While we identified many differences in terminology and analytic approaches, we also identified fundamental similarities and ways in which both disciplines can borrow best practices to advance theory that provides a more cohesive understanding of organizational phenomena.

ORCID iDs

Paul D. Bliese https://orcid.org/0000-0002-5384-8879
Robert E. Ployhart https://orcid.org/0000-0003-1815-9215

Notes

1. In simulations, Semadeni et al. (2014) found that even when the independent variable is correlated at a low level with the error term (e.g., $r = .1$), coefficient estimates are inflated significantly.

2. In some instances, discrete time event history models (e.g., survival analysis) are applied to panel data. We direct readers to Singer and Willett (2003) for details on these models. As noted in our review, discrete time event history models are common in macroresearch; however, they would not be applied to the example data in Table 1.

3. Some variables, such as firm size, may change in minor ways on an annual basis but may still be best conceptualized as an invariant (unchanging) attribute of the firm for both theoretical and estimation reasons, particularly if firms in the sample tend to vary substantially on this dimension.

4. In practice, group-mean correlations are downwardly biased in this extreme case with only three groups. Across 10,000 pseudogroup iterations, we observed an average group-level correlation of .697.

5. To assist readers, we have provided R and STATA code, including the data in Table 1, to run all models discussed in the text, with detailed comments, in the online supplemental material.

6. In practice, researchers rarely provide parameter estimates for individual firms or years in tables, particularly if the data contain a large number of firms and many years. Indeed, some statistical packages used by macroresearchers do not return estimates of the firm fixed effects as routine output (e.g., xtreg, in STATA). In Table 4, the referent is the first group (Firm 1 and Year 2012) but arbitrary. Models can also be fit without a common intercept, which will not change the estimate of the within effect.

7. It is worth emphasizing that different random-effect algorithms can provide different estimates, so users should be cautious about relying on defaults from statistical programs. The default STATA algorithm using xtreg and the default option in R (plm library) both return random-effects estimates of −0.1094 instead of the mixed-effect estimate of −0.235 listed above. With an estimate of −0.1094, the Hausman test statistic of 24.47 is significant ($p < .001$), implying the need to use a fixed-effects model. Nondefault options in R return estimates much more aligned with our mixed-effect estimate of −0.235, while alternative estimators are not readily available in STATA.

8. As noted in our online appendix, when using panel data estimation techniques such as STATA's xtreg with fixed effects, the coefficient for rho (variance explained) is .97 in the null model and .998 for the conditional model.

9. This refers to articles that use fixed effects for the Level 2 unit of analysis (i.e., firm fixed effects when the unit of observation is a firm-year). Industry or year fixed effects would not be counted as using fixed effects to test within-entity relationships.

10. We also observed generalized estimating equations, generalized methods of moments, regression with clustered standard errors, and Cox proportional hazard models.

11. Advances in random-effects models, which allow fully crossed random effects for both higher-level entity and time (e.g., lmer in R), will return the "gold-standard" within-firm estimate if means for higher-level entities and means for each time period are also included.

References

Allison, P. D. 2005. *Fixed effects regression methods for longitudinal data using SAS*. Cary, NC: SAS Institute.

Amemiya, T. 1971. The estimation of the variances in a variance-components model. *International Economic Review*, 12: 1-13.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2010. On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21: 1086-1120.

Bailey, M. A. 2017. *Real econometrics: The right tools to answer important questions*. New York: Oxford University Press.

Barthélemy, J. 2017. The impact of technical consultants on the quality of their clients' products: Evidence from the Bordeaux wine industry. *Strategic Management Journal*, 38: 1174-1190.

Bates, D., Mächler, M., Bolker, B., & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67: 1-48.

Beck, J. W., Scholer, A. A., & Hughes, J. 2017. Divergent effects of distance versus velocity disturbances on emotional experiences during goal pursuit. *Journal of Applied Psychology*, 102: 1109-1123.

Bell, A., & Jones, K. 2015. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3: 133-153.

Bliese, P. D. 2000. Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*: 3-90. San Francisco: Jossey-Bass.

Bliese, P. D., & Hanges, P. J. 2004. Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. *Organizational Research Methods*, 7: 400-417.

Bliese, P. D., & Lang, J. W. 2016. Understanding relative and absolute change in discontinuous growth models: Coding alternatives and implications for hypothesis testing. *Organizational Research Methods*, 19: 562-592.

Bliese, P. D., Maltarich, M. A., & Hendricks, J. L. 2018. Back to basics with mixed-effects models: Nine take-away points. *Journal of Business and Psychology*, 33: 1-23.

Bliese, P. D., & Ployhart, R. E. 2002. Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, 5: 362-387.

Bodner, T. E., & Bliese, P. D. 2018. Detecting and differentiating the direction of change and intervention effects in randomized trials. *Journal of Applied Psychology*, 103: 37-53.

Certo, S. T., Withers, M. C., & Semadeni, M. 2017. A tale of two effects: Using longitudinal data to compare within- and between-firm effects. *Strategic Management Journal*, 38: 1536-1556.

Chan, D. 1998. Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, 83: 234-246.

Chatterji, A. K., Findley, M., Jensen, N. M., Meier, S., & Nielson, D. 2016. Field experiments in strategy research. *Strategic Management Journal*, 37: 116-132.

Chen, G., Ployhart, R. E., Thomas, H. C., Anderson, N., & Bliese, P. D. 2011. The power of momentum: A new model of dynamic relationships between job satisfaction change and turnover intentions. *Academy of Management Journal*, 54: 159-181.

Chin, M. K., & Semadeni, M. 2017. CEO political ideologies and pay egalitarianism within top management teams. *Strategic Management Journal*, 38: 1608-1625.

Croissant, Y., & Millo, G. 2008. Panel data econometrics with R: The plm package. *Journal of Statistical Software*, 27: 1-43.

Deb, P., David, P., & O'Brien, J. 2017. When is cash good or bad for firm performance? *Strategic Management Journal*, 38: 436-454.

Diez-Roux, A. V. 1998. Bringing context back into epidemiology: Variables and fallacies in multilevel analysis. *American Journal of Public Health*, 88: 216-222.

Eberhart, R. N., Eesley, C. E., & Eisenhardt, K. M. 2017. Failure is an option: Institutional change, entrepreneurial risk, and new firm growth. *Organization Science*, 28: 93-112.

Eden, D. 2017. Field experiments in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, 4: 91-122.

Fiegenbaum, A., & Thomas, H. 1988. Attitudes toward risk and the risk–return paradox: Prospect theory explanations. *Academy of Management Journal*, 31: 85-106.

Greene, W. H. 2018. *Econometric analysis* (8th ed.). Upper Saddle River, NJ: Prentice Hall.

Greve, H. R., & Man Zhang, C. 2017. Institutional logics and power sources: Merger and acquisition decisions. *Academy of Management Journal*, 60: 671-694.

Hamilton, B. H., & Nickerson, J. A. 2003. Correcting for endogeneity in strategic management research. *Strategic Organization*, 1: 51-78.

Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, 46: 1251-1271.

Haynes, K., Campbell, J. T., & Hitt, M. A. 2017. When more is not enough: Executive greed and its influence on shareholder wealth. *Journal of Management*, 43: 555-584.

Highhouse, S. 2009. Designing experiments that generalize. *Organizational Research Methods*, 12: 554-566.

Hofmann, D. A., & Gavin, M. B. 1998. Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, 24: 623-641.

Hox, J. J. 2002. *Multilevel analysis techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hutzschenreuter, T., & Horstkotte, J. 2013. Performance effects of top management team demographic faultlines in the process of product diversification. *Strategic Management Journal*, 34: 704-726.

Ilies, R., Liu, X. Y., Liu, Y., & Zheng, X. 2017. Why do employees have better family lives when they are highly engaged at work? *Journal of Applied Psychology*, 102: 956-970.

Kennedy, P. 2008. *A guide to modern econometrics*. Oxford, England: Blackwell.

Kenny, D. A. 1979. *Correlation and causality*. New York: Wiley.

Kim, Y., & Ployhart, R. E. 2014. The effects of staffing and training on firm productivity and profit growth before, during, and after the Great Recession. *Journal of Applied Psychology*, 99: 361-389.

Kozlowski, S. W. J., & Klein, K. J. 2000. *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. San Francisco: Jossey-Bass.

Kreft, I., & de Leeuw, J. 1998. *Introducing multilevel modeling*. London: Sage.
Lee, J. M., Hwang, B. H., & Chen, H. 2017. Are founder CEOs more overconfident than professional CEOs? Evidence from S&P 1500 companies. *Strategic Management Journal*, 38: 751-769.
Lincoln, J. R., & Zeitz, G. 1980. Organizational properties from aggregate data: Separating individual and structural effects. *American Sociological Review*, 45: 391-408.
Liu, W., Song, Z., Li, X., & Liao, Z. 2017. Why and when leaders' affective states influence employee upward voice. *Academy of Management Journal*, 60: 238-263.
Liu, Y., Song, Y., Koopmann, J., Wang, M., Chang, C. H. D., & Shi, J. 2017. Eating your feelings? Testing a model of employees' work-related stressors, sleep quality, and unhealthy eating. *Journal of Applied Psychology*, 102: 1237-1258.
LoPilato, A. C., & Vandenberg, R. J. 2015. The not so direct cross-level direct effect. In C. E. Lance & R. J. Vandenberg (Eds.), *More statistical and methodological myths and urban legends*: 292-310. New York: Taylor & Francis.
Mahoney, J. T., & Qian, L. 2013. Market frictions as building blocks of an organizational economics approach to strategic management. *Strategic Management Journal*, 34: 1019-1041.
Matta, F. K., Scott, B. A., Colquitt, J. A., Koopman, J., & Passantino, L. G. 2017. Is consistently unfair better than sporadically fair? An investigation of justice variability and stress. *Academy of Management Journal*, 60: 743-770.
McNeish, D., & Kelley, K. 2019. Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, 24: 20-35.
Nerlove, M. 1971. A note on error components models. *Econometrica: Journal of the Econometric Society*, 39: 383-396.
Nyberg, A., & Ployhart, R. E. 2013. Context-emergent turnover (CET) theory: A theory of collective turnover. *Academy of Management Review*, 38: 109-131.
Pinheiro, J. C., & Bates, D. M. 2000. Linear mixed-effects models: Basic concepts and examples. In D. M. Bates & J. C. Pinheiro (Eds.), *Mixed-effects models in S and S-Plus*: 3-56. New York: Springer Verlag.
Pontikes, E. G., & Barnett, W. P. 2017. The non-consensus entrepreneur: Organizational responses to vital events. *Administrative Science Quarterly*, 62: 140-178.
Preacher, K. J., Zyphur, M. J., & Zhang, Z. 2010. A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, 15: 209-233.
Ramarajan, L., Rothbard, N. P., & Wilk, S. L. 2017. Discordant vs. harmonious selves: The effects of identity conflict and enhancement on sales performance in employee–customer interactions. *Academy of Management Journal*, 60: 2208-2238.
Raudenbush, S. W. 2009. Adaptive centering with random effects: An alternative to the fixed effects model for studying time-varying treatments in school settings. *Education*, 4: 468-491.
Raudenbush, S. W., & Bryk, A. S. 2002. *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
Rubin, D. B. 2008. For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, 2: 808-840.
Schunck, R. 2013. Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *Stata Journal*, 13: 65-76.
Semadeni, M., Withers, M. C., & Certo, S. T. 2014. The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, 35: 1070-1079.
Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
Shaver, J. M. 1998. Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival? *Management Science*, 44: 571-585.
Shaver, J. M. 2019. Interpreting interactions in linear fixed-effect regression models: When fixed-effect estimates are no longer within-effects. *Strategy Science*, 4: 25-40.
Singer, J. D., & Willett, J. B. 2003. *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
Singh, K., Mahmood, I. P., & Natarajan, S. 2017. Capital market development and firm restructuring during an economic shock. *Organization Science*, 28: 552-573.

Snijders, T., & Bosker, R. 1999. *Multilevel analysis: An introduction to basic and applied multilevel analysis*. London: Sage.

Sonnentag, S., Pundt, A., & Venz, L. 2017. Distal and proximal predictors of snacking at work: A daily-survey study. *Journal of Applied Psychology*, 102: 151-162.

Spieler, I., Scheibe, S., Stamov-Roßnagel, C., & Kappas, A. 2017. Help or hindrance? Day-level relationships between flextime use, work-nonwork boundaries, and affective well-being. *Journal of Applied Psychology*, 102: 67-87.

Stewart, G. L., Astrove, S. L., Reeves, C. J., Crawford, E. R., & Solimeo, S. L. 2017. Those with the most find it hardest to share: Exploring leader resistance to the implementation of team-based empowerment. *Academy of Management Journal*, 60: 2266-2293.

Swamy, P. A. V. B., & Arora, S. S. 1972. The exact finite sample properties of the estimators of coefficients in the error components regression models. *Econometrica: Journal of the Econometric Society*, 40: 261-275.

Tan, D., & Tan, J. 2017. Far from the tree? Do private entrepreneurs agglomerate around public sector incumbents during economic transition? *Organization Science*, 28: 113-132.

Taylor, S. G., Bedeian, A. G., Cole, M. S., & Zhang, Z. 2017. Developing and testing a dynamic model of workplace incivility change. *Journal of Management*, 43: 645-670.

Thorndike, E. L. 1939. On the fallacy of imputing the correlations found for groups to the individuals or smaller groups composing them. *The American Journal of Psychology*, 52: 122-124.

Verbeek, M. 2008. *A guide to modern econometrics*. Chichester, England: John Wiley & Sons.

Wallace, T. D., & Hussain, A. 1969. The use of error components models in combining cross section with time series data. *Econometrica: Journal of the Econometric Society*, 37: 55-72.

Williams, C., Chen, P. L., & Agarwal, R. 2017. Rookies and seasoned recruits: How experience in different levels, firms, and industries shapes strategic renewal in top management. *Strategic Management Journal*, 38: 1391-1415.

Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Woolum, A., Foulk, T., Lanaj, K., & Erez, A. 2017. Rude color glasses: The contaminating effects of witnessed morning rudeness on perceptions and behaviors throughout the workday. *Journal of Applied Psychology*, 102: 1658-1672.

Wright, P. M., Coff, R., & Moliterno, T. P. 2014. Strategic human capital: Crossing the great divide. *Journal of Management*, 40: 353-370.

Xu, K., Tihanyi, L., & Hitt, M. A. 2017. Firm resources, governmental power, and privatization. *Journal of Management*, 43: 998-1024.