

Using Macro Archival Databases to Expand Theory in Micro Research

**N. Sharon Hill^a, Herman Aguinis^a, Josiah Drewry^a,
Sanjay Patnaik^b and Jennifer J. Griffin^c**

^a*The George Washington University*; ^b*The Brookings Institution*; ^c*Loyola University Chicago*

ABSTRACT Databases containing macro-level data are an underutilized methodological tool for expanding theory in micro research (i.e., individual and team) to the macro (i.e., organizational and higher) level of analysis. We describe how macro archival databases support different theoretical approaches for upwardly expanding micro research and summarize unanswered research questions across micro domains requiring upward expansion. We describe 31 macro archival databases as a resource for testing research questions that upwardly expand theorizing in micro domains and how databases enable methodological best practices (i.e., data collection over time, multiple measures of a construct, multilevel statistical controls, missing data and outlier management) that are often difficult to apply in typical micro research. Finally, we demonstrate the feasibility and benefits of using macro archival databases through an empirical illustration in the workplace diversity domain: positive effects of manager gender diversity and manager industry gender diversity on firm performance (i.e., ROA and ROE).

Keywords: diversity, macro archival databases, micro research, micro theory, multilevel

INTRODUCTION

An important yet underutilized opportunity for advancing theory in domains such as organizational behaviour and human resource management (OBHRM) is to expand theorizing in micro research to the macro level (i.e., to the organizational and higher levels of analysis) (Kozlowski, 2017; Porter and Schneider, 2014). Researchers in OBHRM and related fields such as organizational psychology and entrepreneurship focusing on

Address for reprints: N. Sharon Hill, Department of Management, School of Business, The George Washington University, 2201 G Street, NW, Washington, DC 20052, USA (nshill@gwu.edu).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

the entrepreneurs (rather than their ventures) usually focus on the individual and team levels of analysis – what is referred to as a ‘micro’ focus (Aguinis et al., 2011). On the other hand, a ‘macro’ focus refers to organizational-level phenomena and the context in which they occur (i.e., industries, countries). Thus, expanding micro research to the macro level involves developing theory that examines constructs and relations among constructs traditionally studied at the individual and team level at higher levels. This upward expansion of micro research can supplement, complement, and extend more traditional micro research by advancing understanding of how human capital functions influence organizational outcomes (Jiang et al., 2013; Kozlowski, 2017) and takes an important step toward ‘bridging different communities’ in order to ‘bring together different academic domains’ (Faems et al., 2016, p. 111).

Despite its recognized importance for advancing theory in micro research (Porter and Schneider, 2014), this upward theoretical expansion has been limited, resulting in repeated calls for this type of research across numerous micro research domains. For example, in the domain of workplace diversity, micro researchers have mostly examined antecedents and effects at the individual and team levels, resulting in calls to better understand the influence of different types of diversity within the workforce at the organizational level (Joshi et al., 2011; Roberson et al., 2017a). As another example, there have been calls to understand the strategic implications of employee health and well-being programs (Ployhart and Hale, 2014; Tetrick and Winslow, 2015).

A major challenge related to this desired upward expansion of theorizing in micro research is empirical testing, which requires data with macro-level variance on traditionally micro topics, including organizational-level aggregations of variables typically measured at the individual and team level. Specifically, these data are difficult to obtain using typical micro data collection techniques such as surveys, which have remained the most popular data collection approach in micro fields over many decades (Aguinis and Edwards, 2014; Aguinis et al., 2019; Bainbridge et al., 2017). Surveys and other labour-intensive data collection methods (e.g., interviews) have primarily been used to gather data from individuals and teams within a single organization or a few organizations to examine relations at the micro level (i.e., individual and team level) (Solarino and Aguinis, 2021). As discussed extensively in the multilevel literature (e.g., Aguinis et al., 2013a; Mathieu and Chen, 2011), an empirical examination of macro-level relations (i.e., organizational-level and higher) is not possible with data from a single organization, or even a few organizations (Mathieu et al., 2012), because they do not provide sufficient variance at the macro level. This limits empirical testing of upwardly expanded micro theories.

In the limited number of studies where surveys have been used to collect data at the organizational level, the organizational sample size has been relatively small. For example, Buyl et al. (2011) included 33 firms, Kunze et al. (2013) 147 firms, Liao and Chuang (2004) 52 restaurants, and Wu and Chaturvedi (2009) 23 firms. These techniques also make it virtually impossible to collect data at the industry and higher (e.g., country) levels of analysis. Related to this, in their study of the relation between workforce age diversity and firm performance, Kunze et al. discussed the limitation of the relatively small sample size and the important need for future research using much larger samples with companies of different sizes and from different cultural backgrounds to test theory. However, traditional labour-intensive micro data-gathering techniques make it difficult to collect

large multi-organization samples. For these reasons, the focus on traditional micro techniques such as surveys can pose a methodological barrier for examining micro research questions that extend to the macro level.

THE PRESENT STUDY

The goal of our article is to explain and then demonstrate with an empirical illustration how archival databases containing macro-level data (henceforth referred to as macro archival databases) are a useful source of data for studies aiming to test theory that expands micro research to the organizational and higher levels of analysis. Macro archival databases include organizational-, industry-, and country-level (e.g., societal, economic, legal and cultural characteristics) variables, which are often collected over time. These databases are used routinely in macro domains such as strategic management, organization and management theory, and venture-focused entrepreneurship. But micro researchers have yet to systematically take advantage of macro archival databases as a methodological tool to advance theory involving phenomena of interest in micro domains by examining constructs, and relations among constructs, traditionally examined at the individual and team level at higher levels. In this way, the article's goal is like that of studies in *Journal of Management Studies* that offered recommendations and best practices for methodological advancements that facilitate theoretical extensions in micro research (e.g., Aguinis and Edwards, 2014; Echambadi et al., 2006; Greco et al., 2018; Leitch and Palmer, 2010; Solarino and Aguinis, 2021). However, we differ from previous works that offer methodological improvements of traditional micro research techniques (e.g., Solarino and Aguinis, 2021) in that we focus on the application of macro-research techniques to advance micro research theorizing.

Our article demonstrates the benefits of macro archival databases for testing upwardly expanded theory in micro research. Relative to more traditional micro research techniques such as surveys and interviews, macro archival databases are an easier means of obtaining large samples of relevant macro-level variables as well as data across multiple organizations, industries, and even countries. Specifically, they contain variables related to topics traditionally of interest in micro research (e.g., diversity, turnover, compensation and rewards, employee health and safety), including organizational-level aggregations of variables typically measured at the individual and team level. Moreover, macro archival databases also include organizational-level outcome measures (e.g., firm performance) and variables related to characteristics of the macro context (i.e., the organization, industry, and country) that are useful for understanding the macro-level antecedents and consequences, as well as moderators, of upwardly expanded micro phenomena.

In addition, we show their important methodological benefits in the process of upward expansion resulting from the large sample sizes and broad range of variables they contain. These include the availability of variables collected over time for more robust tests of hypothesized causal relations; the opportunity to gather additional evidence about hypothesized effects by testing relations with multiple measures of the same underlying construct (e.g., firm performance); as well as large sample sizes and a broad set of variables at the organizational, industry, and country level that facilitate using macro-contextual

statistical controls to help rule out alternative explanations and implement best practices related to missing data and outlier management.

We emphasize that we are *not* suggesting that micro researchers abandon their current methodological approaches. Rather, we suggest they use macro archival databases, only when appropriate, in areas where there are opportunities to supplement, complement and extend micro theories by expanding them to the macro level. Moreover, the use of macro archival databases may not be useful in some micro research domains. For example, micro research domains where questions are best suited to empirical examination using an exclusive micro approach (e.g., research questions related to employee motivation, personality, individual abilities, leadership and team member exchange as well as job analysis) may not benefit from upward theoretical expansion. In addition, researchers may use micro techniques such as surveys and interviews to analyse a topic in more detail than is possible using macro archival databases. In summary, we view archival databases as a useful resource that supplements, but certainly does not replace, more traditional micro research methods.

Our article provides a consolidated, practical resource for researchers who are largely unfamiliar with using macro archival databases. It helps them identify relevant databases for testing different types of research questions that expand micro research. We discuss how macro archival databases can facilitate various theoretical approaches to upward expansion and provide evidence of the need for such theoretical advancement by highlighting calls across numerous research domains to address unanswered research questions requiring upward expansion. We also describe methodological best practices that help researchers maximize the advantages of macro archival databases. Finally, we illustrate with an empirical study on workplace diversity the feasibility and methodological benefits of using macro archival databases to test upwardly expanded theory in micro research.

THEORETICAL APPROACHES FOR UPWARD EXPANSION OF MICRO RESEARCH SUPPORTED BY MACRO ARCHIVAL DATABASES

Macro archival databases support several different theoretical approaches that have been used in the literature for upward theoretical expansion of micro research to help bridge the 'divide' between micro and macro research domains (Aguinis et al., 2011; Hitt et al., 2007; Maloney et al., 2016; Ployhart and Hale, 2014; Porter and Schneider, 2014). This divide exists because micro and macro research are rooted in different disciplines. Micro research is grounded in psychology with a focus on individual thoughts, feelings and actions as well as social interactions between individuals. Macro research is based in sociology, economics and understanding organizations and markets. These interdisciplinary differences cause scholars in micro and macro research to focus on different parts of the social and economic system in which management phenomena occur (Molloy et al., 2011; Roberts et al., 1978), with micro theories generally ignoring dynamics arising at higher (i.e., macro) levels of analysis. This helps to explain the heretofore limited use of macro archival databases to examine micro research topics.

Using macro archival databases for upward expansion of micro research supports strategies suggested in past research for bridging the micro-macro research divide

(Aguinis et al., 2011; Hitt et al., 2007; Molloy et al., 2011). These strategies include making scholars more aware of other disciplinary perspectives that are relevant to their field of inquiry, fostering collaboration between scholars who investigate organizational phenomena through different disciplinary lenses and promoting a multilevel approach by integrating theories from different levels or applying theories from one level to another to provide new insights.

One approach to expanding theory in micro research that is supported by using macro archival databases is based on multilevel theorizing (Aguinis and Molina-Azorin, 2015; Mathieu and Chen, 2011), in which characteristics of individuals and teams or collective constructs that emerge from the aggregation of these lower-level units are theorized to influence organizational-level outcomes – what is usually referred to as upward effects. For example, research in the field of strategic human resource management (SHRM) integrates perspectives from psychology and economics to shift the focus from studying human capital outcomes at the individual and team level to understanding how human capital can serve as a resource to improve organizational-level outcomes (Wright et al., 2014, 2018). The SHRM special issue in *Journal of Management* (Wright et al., 2014) provided excellent examples of studies that used macro archival databases to test research questions that align with this theoretical approach (Khanna et al., 2014; Liu, 2014; Liu et al., 2014; Tzabbar and Kehoe, 2014). Examples include studies that examine how the aggregate education and experience of a firm's board of directors influences firm performance (Khanna et al., 2014), and how the individual characteristics of a star scientist who exits a firm determine the effects of the star's turnover for the organization (Tzabbar and Kehoe, 2014). Another area in which this theoretical approach to expanding micro research applies is our earlier example of workforce diversity where researchers have called for a better understanding of the organizational-level effects of different types of workforce diversity (Joshi et al., 2011).

A second theoretical approach relates to understanding how the macro-level context (organizations, industries, societies, economic systems, and national cultures) influences micro phenomena (Aguinis et al., 2013a; Aguinis and Molina-Azorin, 2015; Mathieu and Chen, 2011) – what is usually referred to as downward effects. Related to this, time has also been recognized as an essential element of context in micro research (Johns, 2006) and macro archival databases can be a useful vehicle to examine effects over time. A special issue in *ILR Review* (Burton et al., 2019) provided examples of how researchers have applied this approach in the domain of entrepreneurship to understand who becomes an entrepreneur and entrepreneurship success. This research has typically focused on the influence of individual-level factors; however, studies in this special issue examined how the broader context in which entrepreneurship occurs – including labour market institutions (e.g., wage setting policies, employment security) and firm human resource management practices – influences entrepreneurship activity and outcomes beyond the influence of entrepreneurs' individual characteristics. Studies in this special issue also examined the broader context of organizations to understand differences in effects across different geographies, countries and over time. Turnover research provides another example of where this conceptual downward-effect approach to expanding micro research is relevant. Researchers have highlighted the need to examine how changes in organizational-level turnover over time (increasing or decreasing) impact organizational performance as

well as how the relation between turnover and its antecedents and outcomes at different levels vary across different types of contexts (e.g., organizations and cultures; Hancock et al., 2013; Hom et al., 2017).

Upward expansion of micro research is also relevant to micro foundations programs in macro domains such as strategy and entrepreneurship where an important goal is to develop and test theories about how individuals, their characteristics, and interactions among individuals contribute to understanding macro phenomena (Aguinis and Molina-Azorín, 2015; Felin et al., 2015; Ployhart and Hale, 2014). An example of an emerging area that has applied this approach is micro-foundations of corporate social responsibility (CSR), which has traditionally focused at the macro level (Aguinis and Glavas, 2019; Morgeson et al., 2013), rooted in disciplines such as strategy, finance and corporate governance. A special issue of *Personnel Psychology* showcased research using micro-oriented theories from OBHRM to study CSR-related issues (Morgeson et al., 2013), including studies using macro archival database data (e.g., Ormiston and Wong, 2013). Recent reviews in micro research have called for more research that applies theories (Ployhart and Hale, 2014) and multilevel modelling approaches from micro research (Aguinis and Molina-Azorín, 2015) to advance micro-foundations theorizing.

Research studies like the ones in the special issues we have described provided important insights related to expanding micro theory to the macro level. However, most of the research involving upward expansion in micro research remains undertheorized and primarily conceptual due to the lack of appropriate data (i.e., data with macro-level variance on traditionally micro topics) to test these theories (Eckardt et al., 2019; Ployhart and Hale, 2014; Porter and Schneider, 2014). Our article describes an important methodological approach to test extant theory and develop new theory that upwardly expands micro research.

USING MACRO ARCHIVAL DATABASES TO ADDRESS UNANSWERED RESEARCH QUESTIONS REQUIRING UPWARD EXPANSION OF MICRO RESEARCH

We conducted a review of literature reviews of micro research domains to demonstrate and provide examples of the opportunity that exists to expand micro research to the organizational and higher levels of analysis. We specifically targeted review articles because the future research directions included therein can be viewed as ‘wish lists’ of critical issues that need to be addressed to make theoretical progress. To select articles for our review of reviews, we searched in journals that publish (1) a significant number of articles in domains of interest in micro research and (2) qualitative reviews of these topics. Based on this, we reviewed articles in *Annual Review of Organizational Psychology and Organizational Behavior* (2014–20) and the March 2017 centennial issue of *Journal of Applied Psychology*. We also included review articles addressing OBHRM topics in *Journal of Management Studies*, *Journal of Management*, and *Academy of Management Annals* (2010–20).

Table I provides examples of research questions across domains of interest in micro research that highlight the need for upward expansion. For example, as shown in Table I, there have been calls to understand macro-level influences of phenomena that

Table I. Examples of unanswered research questions and knowledge gaps requiring upward expansion of theory in micro research

<i>Illustrative Research Questions and Knowledge Gaps</i>	<i>Relevant Sources</i>
<p>Turnover:</p> <ul style="list-style-type: none"> • What are the temporal effects of organizational-level turnover? How do changes in organizational-level turnover over time (increasing or decreasing) impact organizational performance? • How does the relation between organizational-level turnover and its antecedents and outcomes vary across different types of organizations, industries, cultures? 	Hancock et al. (2013); Hausknecht (2017); Hom et al. (2017)
<p>Training and Development</p> <ul style="list-style-type: none"> • How does organizations' focus on training and development or other learning initiatives impact their competitive advantage and innovation? 	Noe et al. (2014); Ployhart and Hale (2014)
<p>Compensation and Reward Systems:</p> <ul style="list-style-type: none"> • How does the composition of the workforce influence the relation between pay structures and organizational outcomes? 	Shaw (2014)
<p>Performance Appraisal and Management:</p> <ul style="list-style-type: none"> • How does the quality and nature of performance appraisal programs influence organizational-level performance? • How do organizational factors (e.g., size) and industry type shape the nature and quality of organizations' performance appraisal programmes? 	DeNisi and Smith (2014); DeNisi and Murphy (2017)
<p>Workplace Diversity:</p> <ul style="list-style-type: none"> • What are the organizational-level outcomes of different types of diversity in the overall workforce, and what macro contingency factors (e.g., industry characteristics, national culture) moderate their effects? • How does workforce diversity influence outcomes beyond organizational performance (e.g., service ratings, reputation indices, environmental performance)? 	Joshi et al. (2011); Roberson (2019); Roberson et al. (2017a)
<p>Employee Health and Wellbeing:</p> <ul style="list-style-type: none"> • How do different employee stress management and wellness programmes influence organizational outcomes, and what organizational-level and country-level factors moderate these effects? • What are the strategic implications of employee safety and health programmes? • What are the negative effects of HR systems on employee health and wellbeing, and how do these outcomes relate to firm performance? 	Jiang et al. (2013); Ployhart and Hale (2014); Tetrick and Winslow (2015)
<p>Creativity:</p> <ul style="list-style-type: none"> • What are organizational influences on collective employee creativity? • How does organizations' collective employee creativity vary across cultural contexts? 	Zhou and Hoever (2014)
<p>Leadership:</p> <ul style="list-style-type: none"> • What macro-contextual factors (i.e., societal, political, economic, and organizational) influence the representation of women leaders in organizations? 	Lyness and Grotto (2018)

are traditionally of interest in micro research on a range of organizational-level outcomes as well as macro-contextual factors that predict them and moderate their effects. These calls relate to micro phenomena in domains such as turnover (Hancock et al., 2013; Hausknecht, 2017; Hom et al., 2017), training and development (Noe et al., 2014; Ployhart and Hale, 2014), compensation and reward systems (Shaw, 2014), performance appraisal and management (DeNisi and Murphy, 2017; DeNisi and Smith, 2014), workplace diversity (Joshi et al., 2011; Roberson, 2019; Roberson et al., 2017a), employee health and well-being (Jiang et al., 2013; Ployhart and Hale, 2014; Tetrick and Winslow; 2015), creativity (Zhou and Hoever, 2014), and leadership (Lyness and Grotto, 2018).

In sum, our review of recently published reviews in micro research demonstrates the rather pervasive need across multiple domains to expand micro research to the macro level. But, being able to fulfil this 'wish list' of research questions requires useful data with macro-level variance on typical micro topics, which poses a methodological challenge for traditional data collection methods used in micro research. Accordingly, in the next section, we identify and describe 31 macro archival databases containing relevant variables that are particularly useful for addressing unanswered questions and knowledge gaps such as those listed in Table I.

Selecting Macro Archival Databases and Variables to Expand Micro Research

Databases, variables, and their roles. We created detailed summaries of 31 commonly used macro archival databases that contain data on organizations, industries, and countries (e.g., societal, economic, legal and cultural data). Given the goal of this article to provide a general resource, we focused on major databases available for free or through academic subscriptions that are relevant to a broad range of micro research domains. A member of our research team was previously a business and economics librarian and is very familiar with available databases relevant to business domains. We also consulted Strauss's *Handbook of Business Information: A Guide for Librarians, Students, and Researchers* (Moss and Ernsthausen, 2014), which lists databases relevant to various business domains and also identified other databases related to business and economics that are freely available online. We reviewed the information available in each database and selected those containing variables that are most relevant to addressing the types of research questions described in the previous section involving upward expansion in micro-research domains. This list of 31 databases is not exhaustive; other databases are available (potentially at a higher cost) that are narrower in scope (e.g., focused on a particular research domain, industry or country) and/or more specific to a particular area of inquiry. The databases in the online supplement provide a useful starting point for researchers who are unfamiliar with macro archival databases.

The 31 database summaries are included in Table SI in the online supplement and contain the following information: (1) vendor and website; (2) overall description; (3) unique characteristics; (4) level of analysis of the data (e.g., individual, organization, industry, country); (5) geographic focus; (6) relevant categories of variables; and (7) subscription information. As an illustration of the type of information provided for each of the 31 databases, Table II shows an example of a summary description for the database Thomson Reuters ESG.

Table II. Illustrative database description: Thomson Reuters ESG (Environmental, Social, and Governance; formerly known as Asset4)

<i>Vendor and website</i>	<i>Available as part of Datastream and Eikon, data platforms sold by Thomson Reuters: http://financial.thomsonreuters.com/esg</i>
Overall description	This database aims to provide a comprehensive picture of a company's sustainability reporting and commitments, measuring firm spending, structure and actions in four 'pillars': environmental, social and governance, and economic performance. For 6000 companies, there are just over 750 variables, which are then aggregated into 400 key performance indicators (KPIs) that enable comparisons across companies, industries, and regions. For higher-level analysis, the KPIs are weighted and assigned to 18 categories, which are given scores, and then combined into the 4 pillars, which are also given scores that go into an overall company-level score
Unique characteristics	<ul style="list-style-type: none"> • Assigns quantitative scores to many qualitative variables, as assessed by in-house experts. Also reports quantitative data taken from company filings and other public sources. • KPIs are weighted based on relevance of the category to the industry and region, percentages of companies reporting, measurability, ability to confirm data independently, and other factors • Includes Yes-No indicator variables for company practices (e.g., animal testing, existence of policies and procedures, controversies in the media)
Level of analysis	Data are at the firm level, from 2002-present
Geographic focus	Global. Covers firms in the S&P 500, STOXX 600, NASDAQ 100, Russell 1000, FTSE 100, ASX 300, MSCI Europe, MSCI World, MSCI Emerging Market indices
Relevant variable categories	<ul style="list-style-type: none"> • Environmental variables include: resource use and reduction, emissions, energy use, electricity purchased, environmental impact controversies • Social variables include: employment quality, employee training and development, employee well-being, health and safety, diversity, human rights, community engagement, product responsibility • Corporate Governance variables include: board structure and functions, compensation policies, shareholders' rights • Economic variables include: financial performance, employee productivity, client loyalty, shareholder loyalty, accounting controversies
Subscription information	<ul style="list-style-type: none"> • Available as a subscription through Thomson Reuters Eikon. The Core package contains scores the top-level pillars and categories. The Premium package has more granular data, including the 750+ variables and scores for KPIs • Data are also available in Thomson Reuters Datastream (as 'Data Types' in the Equities module)

Note: See Table SI in the online supplement for similarly detailed descriptions of the databases listed in Tables III and IV.

As an additional resource, Table III shows different types of variables available across the 31 databases that are particularly useful for expanding theory in micro research to the macro level. These variables reflect topics traditionally of interest in micro research, including organizational-level aggregations of variables typically measured at the individual and team level. Table III also includes information on performance and other

Table III. Types of variables in macro archival databases useful for expanding micro research to the macro level

Databases	Performance management (e.g., hours of training per employee, total training spending, mentoring & development programs)	Employee health and well-being (e.g., health & safety policies and training, work/life balance, family friendly policies)	Employee compensation (e.g., average salary, bonus plans, fringe benefits, CEO pay gap)	Workforce diversity & inclusion (e.g., % women, minority, disabled employees in workforce and in management equal opportunity policies)	Business ethics & legal compliance (e.g., number of ethical & legal, controversies, product responsibility policies, product recalls)	Safety (e.g., accident rates, time lost, waste); and customer service (e.g., customer satisfaction, quality data)	Financial performance (e.g., revenue, profit, growth, return on capital); and productivity (e.g., efficiency, costs, output per worker)
Amadeus	✓						✓
Bloomberg Professional	✓	✓	✓	✓	✓	✓	✓
BoardEx Academic			✓				
Capital IQ			✓				✓
Compustat							✓
Datastream							✓
Execucomp			✓				✓
Factiva			✓			✓	✓
Foundation Center			✓		✓		✓
LexisNexis Academic			✓		✓		✓
Orbis							✓
Osiris					✓		✓
Roper Center Polling Data					✓		
Thomson ONE							✓

Table III. Continued

Databases	Performance management (e.g., hours of training per employee, total training spendings, mentoring & development programs)	Employee health and well-being (e.g., health & safety policies and training, work/life balance, family friendly policies)	Employee compensation (e.g., average salary, bonus plans, fringe benefits, CEO pay gap)	Workforce diversity & inclusion (e.g., % women, minority, disabled employees in workforce and in management equal opportunity policies)	Business ethics & legal compliance (e.g., number of ethical & legal, controversies, product responsibility policies, product recalls)	Safety (e.g., accident rates, time lost, waste); and customer service (e.g., customer satisfaction, quality data)	Financial performance (e.g., revenue, profits, growth, return on capital); and productivity (e.g., efficiency, costs, output per worker)
Thomson Reuters ESG (Environmental, Social, and Governance; formerly known as Asset4)	✓	✓	✓	✓	✓	✓	✓
Wharton Research Data Services (WRDS)	✓	✓	✓	✓	✓		✓
Worldscope							✓
Zephyr							✓

Note: ✓ denotes availability of variables in the specific database. See Table SI in the online supplement for detailed descriptions of each of the databases.

Table IV. Types of macro-contextual variables in macro archival databases useful for expanding micro research to the macro level

<i>Databases</i>	<i>Organizational characteristics (e.g., size, location, product offerings, supply chain relationships)</i>	<i>Senior leadership team (top mgt. team, & directors) characteristics (e.g., career history, demographic & functional diversity)</i>	<i>Industry characteristics (e.g., size, growth, innovation, alliances, competition, regulation)</i>	<i>Country characteristics (e.g., governance, politics, population, taxes, trade, culture)</i>
Amadeus	✓	✓	✓	
Bloomberg Professional	✓	✓	✓	✓
BoardEx Academic		✓		
Capital IQ	✓		✓	✓
Carbon Disclosure Project	✓			
Census Bureau			✓	✓
Compustat	✓			
Database of Political Institutions				✓
Datastream	✓		✓	✓
Economist Intelligence Unit (EIU)			✓	✓
Execucomp		✓		
Factiva	✓	✓	✓	✓
Foundation Center		✓		
Gallup Analytics				
Global Leadership & Organizational Behavior Effectiveness (GLOBE)				✓
				✓

Table IV. Continued

Databases	Organizational characteristics (e.g., size, location, product offerings, supply chain relationships)	Senior leadership team (top mgt. team, & directors) characteristics (e.g., career history, demographic & functional diversity)	Industry characteristics (e.g., size, growth, innovation, alliances, competition, regulation)	Country characteristics (e.g., governance, politics, population, taxes, trade, culture)
Hoover's	✓		✓	
ILOSTAT (and other International Labour Organization Data)				✓
LexisNexis Academic	✓	✓	✓	✓
Organisation for Economic Cooperation and Development (OECD) iLibrary			✓	✓
Open Secrets	✓		✓	✓
Orbis	✓	✓	✓	✓
Osiris	✓	✓	✓	✓
Polity IV				✓
Roper Center Polling Data				✓
Thomson Reuters ESG (Environmental, Social, and Governance; formerly known as Asset4)	✓	✓		✓
Thomson ONE	✓		✓	✓
Wharton Research Data Services (WRDS)	✓	✓	✓	✓
World Bank Data			✓	✓

(Continues)

Table IV. Continued

Databases	Organizational characteristics (e.g., size, location, product offerings, supply chain relationships)	Senior leadership team (top mgt. team, & directors) characteristics (e.g., career history, demographic & functional diversity)	Industry characteristics (e.g., size, growth, innovation, alliances, competition, regulation)	Country characteristics (e.g., governance, politics, population, taxes, trade, culture)
Worldscope	✓	✓		
Worldwide Governance Indicators (WGI)				✓
Zephyr	✓	✓	✓	

Note: ✓ denotes availability of variables in the specific database. See Table SI in the online supplement for detailed descriptions of each of the databases.

outcomes at the organizational level that these variables might predict. For example, a researcher seeking to understand the organizational-level influence of workforce diversity (e.g., percent of women or minorities in the entire workforce) might use Table III to identify Thomson Reuters ESG as a database that contains relevant organizational-level diversity and performance variables. The researcher could then consult our article's online supplement (Table SI) for additional information about this particular database.

As yet an additional resource, Table IV shows different types of macro-contextual variables in the databases useful for expanding micro research to the macro level. These variables include characteristics of organizations (e.g., size, location, and product offerings) and their senior leadership team (e.g., career history, demographics), industries (e.g., competition, regulation, and growth) and countries (e.g., societal, economic, legal and cultural) that can be used for answering the types of research questions shown in Table I. First, these variables might predict macro-level constructs related to traditionally micro phenomena. For example, a researcher might be interested in examining how characteristics of organizations covary with the types of performance management programs organizations implement. Second, these macro-contextual variables can also be conceptualized as moderators. For example, a researcher might be interested in investigating how factors at the organizational, industry, or country level moderate the relation between organizational-level turnover and its outcomes. Finally, some of the macro-contextual variables in Table IV can also be conceptualized as macro-level statistical control variables to help rule out alternative explanations for hypothesized effects.

In sum, researchers can use Tables III and IV to efficiently select potentially relevant databases and variables and then consult the detailed descriptions of those databases in the online supplement to identify the ones with the specific variables of interest. The variables required to answer questions and address knowledge gaps such as those in Table I may all reside in one database; or, however, a researcher may need to merge data from multiple databases. We will demonstrate these procedures in our empirical illustration later in our article. It is also possible to combine macro archival database data with data collected using more traditional micro data collection techniques. For example, in the SHRM domain, Liu et al.'s (2014) study of how a firm's capital structure influences its investment in strategic human capital combined survey data from general managers with macro archival database data. In addition, Ormiston and Wong (2013) examined how CEOs' psychological processes influence corporate social irresponsibility using macro archival data combined with data about CEOs gathered using a qualitative methodology from publicly available data from the business press and books.

Benefits of Using Macro Archival Databases for Implementing Methodological Best Practices in Micro Research

Due to the large sample sizes and broad range of variables they contain, macro archival databases can provide methodological benefits in the process of upward expansion of micro research. These benefits relate to data collection (e.g., research design and measurement, control variables, and missing data management) and data preparation (e.g., outlier management) (Aguinis et al., 2021), which may be difficult to achieve using traditional micro research techniques. These methodological improvement needs have been

discussed in past publications (e.g., Aguinis et al., 2021); therefore, we focus our discussion below on how macro archival databases can help researchers realize these benefits. We also discuss these methodological advantages in more detail in the context of applying them to our empirical demonstration.

Research design and measurement. First, macro archival databases contain data collected over several time periods, which can facilitate more effective testing of causal effects and the specific role of time (Aguinis and Edwards, 2014). Second, the broad set of variables in macro archival databases allows for the use of multiple measures of the same underlying construct (e.g., firm performance measured as both return on assets and return on equity) for richer tests of hypothesized relations between constructs (i.e., tests across different operational definitions).

Control variables. Theoretically relevant controls can strengthen causal inferences by allowing researchers to rule out alternative explanations for the hypothesized effects (Bernierth and Aguinis, 2016), leading to more robust results that enhance the ability to replicate, extend, and generalize a study's findings (Nielsen and Raswant, 2018). Specifically, macro-contextual variables in databases (see Table IV) can serve as theoretically justified controls in studies involving upward expansion of micro research and provide a more practically feasible means of including statistical controls compared to traditional micro techniques such as surveys, which would have to be distributed to a large number of organizations, industries, and countries. The availability of variables collected over time also provides the opportunity to control for the complex feedback loops that can exist in upwardly expanded research. For example, when studying how organizations' investment in education and training initiatives impacts their performance, it is important to remember that organizational performance is also likely to impact investments in training and development. Macro archival databases allow researchers to test the impact of training and development investment at Time 1 on organizational performance at Time 2 and examine how this performance influences Time 3 training and development, controlling for Time 1 investment.

Missing data management. The large samples in databases allow for the implementation of best practices regarding missing data management. First, it is feasible to use a simpler approach to missing data such as listwise or pairwise deletion because the large sample sizes result in a relatively smaller percentage of missing data. This is a benefit given that more complex missing data techniques such as multiple imputation (MI) and maximum likelihood (ML) may not provide noticeable improvements over simpler techniques when the percentage of missing data is small and the sample size remains large enough to avoid insufficient statistical power (Cheema, 2014; Newman, 2014). Second, macro archival databases can also improve results using more complex techniques such as MI and ML. These techniques rely on large-sample assumptions, and therefore results are less likely to contain small-N bias (Schafer and Graham, 2002). Third, the large number of variables in macro archival databases allows researchers to diagnose patterns of missing data more effectively and test assumptions. This makes it easier to implement best-practice recommendations in micro research that depend on the pattern of missingness:

completely at random, at random, or not at random (Newman, 2014). For example, due to the more comprehensive information in the database, a researcher could test whether data are missing at random by examining whether missing variables relate to other variables in the database (Newman, 2014; Schafer and Graham, 2002). Finally, it is possible to replace missing data in one database with data from another database containing the same variable.

Outlier management. The presence of outliers – data points that are markedly different from others – can affect substantive conclusions in non-trivial ways. Best-practice recommendations for outlier management include making informed decisions by first classifying these observations as either error, interesting, or influential outliers (Aguinis et al., 2013b). In micro research, it is difficult to retroactively determine the classification of an outlier once the data have already been collected (e.g., using surveys). However, macro archival databases allow researchers, guided by theory, to examine the larger set of variables included in the database to search for patterns that help to determine the cause (e.g., an extreme value on a financial statement might, on closer inspection, turn out to be accurate due to a particular event such as an acquisition, divestment, or sale of stock). Researchers can also check for error outliers by systematically comparing extreme values with those for the same variable in other databases; and they can use database data to identify ‘interesting’ outliers, by comparing the outlier observation to others across the same and different contexts (e.g., organization, industry, country) to see if it is typical or stands out in interesting ways. Another approach to outlier management, which is more commonly used in macro research, is winsorizing. This approach involves transforming extreme values to a specified percentile of the data. While this approach reduces the influence of some observations, it retains all of them, which can be an advantage over approaches that require outright elimination of extreme values.

Next, we describe an original empirical study illustrating the feasibility and usefulness of macro archival databases for upward theoretical expansion in a domain that is typically of interest in micro research: workplace diversity. Although the empirical illustration focuses on workplace diversity, we could have chosen any of the domains summarized in Table I.

EMPIRICAL ILLUSTRATION: USING MACRO ARCHIVAL DATABASES FOR UPWARD EXPANSION IN WORKPLACE DIVERSITY RESEARCH

In OBHRM, workplace diversity research has primarily focused on understanding the effects of individual demographic dissimilarity to the workgroup on individual outcomes or, alternatively, the effects of workgroup diversity on workgroup outcomes (e.g., Joshi et al., 2011; Roberson et al., 2017b). Motivated by our review of reviews in micro research, this empirical illustration addresses Joshi et al.’s (2011, p. 531) conclusion that ‘relatively little attention has been devoted to the firm-level consequences of firm-level diversity’. In particular, research on the organizational-level effects of manager gender diversity remains rare (Roberson et al., 2017a). Further, three existing studies related to organizations’ manager gender diversity had small samples (between 177 and 243 organizations), which were limited to a single country (Dwyer et al., 2003; Schwab et al., 2016; Shrader

et al., 1997) or a single industry (banks and insurance companies: Dwyer et al., 2003; Schwab et al., 2016). Our illustration study involves upward expansion of workplace diversity research to the organizational and higher levels of analysis using a large sample of firms in multiple industries and countries.

For illustration purposes, we focus on using macro archival database data to examine the relation between firms' manager gender diversity (i.e., the percent of all managers who are women) and firms' performance. We also demonstrate the methodological benefits of macro archival databases described in the previous section related to research design and measurement, control variables, missing data and outlier management.

In micro research, there is a rich history of research aimed at understanding the effects of diversity in workgroups and teams (Joshi et al., 2011). The value-in-diversity (Cox et al., 1991) and information-processing (Williams and O'Reilly, 1998) frameworks are dominant theoretical perspectives predicting positive effects of diversity on collective outcomes (e.g., Buyl et al., 2011; Miller and del Carmen Triana, 2009; Schwab et al., 2016). According to these perspectives, diversity can expand a group's knowledge, viewpoints, and informational resources, resulting in enhanced problem-solving, decision making, and creativity. These conceptual frameworks explain why manager gender diversity can have a positive influence on firm performance (Schwab et al., 2016; Shrader et al., 1997). Male and female managers have different experiences and therefore different knowledge and perspectives; hence, manager gender diversity can broaden a firm's information base for decision making and problem solving to enhance firm performance (van Knippenberg et al., 2004). In addition, men and women have been shown to differ in their approaches to evaluating alternatives, leading to a more thorough evaluation process that enhances decision making (Crow et al., 1991; Park, 1996). Based on these considerations, we hypothesize:

Hypothesis 1: Manager gender diversity at the firm level of analysis will positively relate to firm performance.

Methodological Choices and Best Practices with Benefits for Theory Testing

In our empirical illustration, we made several methodological choices based on best practices to demonstrate the benefits provided by macro archival databases discussed earlier.

Research design and measurement. To demonstrate the usefulness of macro archival databases for gathering evidence regarding causal relations, and consistent with prior firm-level diversity studies (e.g., Andrevski et al., 2014; Shrader et al., 1997), we also assessed the relation between manager gender diversity and firms' performance in the following year.

Hypothesis 2: Manager gender diversity at the firm level of analysis will positively relate to firm performance in the following year.

In addition, we took advantage of the large number of variables in macro archival databases and used two different measures of firm performance: return on assets (ROA)

and return on equity (ROE). These are useful for comparing the financial performance of firms of different sizes, across industries and countries. They are also common financial outcome measures in corporate governance research (Dalton and Aguinis, 2013) as well as in studies of how women's leadership affects firms' performance (Hoobler et al., 2018).

Control variables. The few existing studies that examined the relation between manager gender diversity and firms' performance (Dwyer et al., 2003; Schwab et al., 2016; Shrader et al., 1997) were limited to a single industry and/or country. Thus, because of research design and measurement limitations, they were unable to consider macro-contextual statistical controls related to industry and country characteristics. In our illustration, we took advantage of macro archival database data to include firm-, industry-, and country-level controls based on best practices (Bernerth and Aguinis, 2016).

At the firm level of analysis, we controlled for firm size because it is related to several firm-level outcomes including performance (Josefy et al., 2015) and is also linked to gender diversity (e.g., Metz, 2009). At the industry level, we controlled for industry manager gender diversity (i.e., the mean manager gender diversity for all firms in the same industry) because firms in industries with higher levels of gender diversity are likely to have more gender diverse management. Also, such firms may benefit from the overall broader information base for creativity, decision making and problem solving in an industry, leading to higher levels of firm performance. At the country level of analysis, we controlled for two characteristics that may relate to the diversity predictors in the model. The first is gender egalitarianism, a societal culture dimension reported in the GLOBE database (see Table SI in the online supplement for a description of this database) that reflects 'the degree to which a collective promotes gender equality' (House et al., 2004, p. 30). Firms in more gender egalitarian cultures are likely to have more women managers (Hoobler et al., 2018). The second characteristic is countries' labour force participation ratio (i.e., the ratio of women's vs. men's rate of participation in the labour force), which can affect gender diversity in management through female labour supply.

Missing data management. We used listwise deletion to demonstrate that simpler missing data techniques are feasible without a noticeable detrimental effect on statistical power. The low percent of missingness in our empirical illustration (less than 10 per cent) was facilitated not only by the large sample but also by filling in missing values on the outcome variables using data from another database. Second, we implemented best-practice recommendations in micro research to assess the pattern of missingness: completely at random, at random, or not at random (Newman, 2014).

Outlier management. We implemented best-practice recommendations for outlier management (Aguinis et al., 2013b) by using two different approaches. First, we checked for error outliers by taking advantage of the large set of information across databases to compare outlier observations across multiple databases to see if there was a match. Second, we winsorized the outcome variables ROA and ROE using the 1st and 99th percentiles.

Multilevel data structure. We used data for firms across multiple years and we also examined control variables at different levels of analysis. Thus, our dataset consists of two different three-level nested structures, and we refer to these as the (a) ‘industry-nested model’ and (b) ‘country-nested model’. First, the industry-nested model consists of firm-years (level 1), nested in firms (level 2), nested in industries (level 3). Second, the country-nested model consists of firm-years (level 1), nested in firms (level 2), nested in countries (level 3). Examining both models allowed us to empirically investigate the influence of different types of nesting structures as well as statistical controls at different levels. Clearly, three-level models are not practically feasible in most micro research, but very much so when using macro archival databases.

Method

In this section, we describe how we selected databases and relevant variables, merged data from multiple databases, and calculated new measures.

Databases and sample. We searched for databases that include information on manager gender diversity and firm performance as well as relevant control variables at different levels of analysis. Examining the summaries of the 31 databases in the online supplement suggested that Thomson Reuters ESG, Datastream, Worldscope, and ILOSTAT include the data to test our hypotheses over a period of 15 years from 2002–16 and multiple measures of firm performance. Thomson Reuters ESG reports annual data for each fiscal year in which the company is publicly listed. We kept all firm-year observations (7,058) that contain data from Thomson Reuters ESG for our predictor variable manager gender diversity, our firm-level control variable firm size, and a code identifying the firm’s industry from 2002 through 2016 to be able to use our more specific industry-level manager gender diversity variable. To ensure we only compared data corresponding to full and not partial fiscal years, we excluded 96 observations where firms were acquired, taken private, or otherwise delisted from stock exchanges before their fiscal-year end. Also, as is recommended in the finance and accounting literatures (Dhaliwal et al., 2015), we excluded banks and related lenders (1,015 observations) because the majority of their assets are comprised of loans. Next, we merged in the variables serving as the aforementioned macro-contextual controls, dropping 287 firm-year observations due to missing data on one or more of these variables. Then, we merged in three variables from the Worldscope database: total assets, total shareholders’ equity, and operating profit to create our outcome measures, return on assets (ROA) and return on equity (ROE), dropping 17 observations with negative values of either total assets or shareholders’ equity.

The final sample to test Hypothesis 1 using ROA_{t0} (i.e., t_0 denotes current year) was an unbalanced panel of 5,644 firm-year observations for 1,354 firms from 40 countries from the years 2002 through 2016. Non-US firms accounted for 87 per cent of firm-year observations across six continents (Africa, 2 per cent of firm-year observations; Asia, 16 per cent; Australia/Oceania, 7 per cent; Europe, 51 per cent; North America, 16 per cent; and, South America, 5 per cent). Data on shareholders’ equity were not available for all firm-years leaving us with 62 fewer observations for ROE_{t0} . Therefore, in testing Hypothesis 1 using ROE_{t0} , $N = 5,582$ observations for 1,341 firms. In testing Hypothesis

2, $N = 5,591$ firm-year observations for 1,341 firms for ROA_{t+1} (i.e., $t+1$ denotes following year) and slightly fewer $N = 5,512$ observations for 1,329 firms when using ROE_{t+1} .

Measures

Predictor variable. For manager gender diversity, we used the variable Women Managers (percentage of all managers in a firm that are women) from Thomson Reuters ESG.

Outcome variables. We conducted all analyses with two different operationalizations of firm performance assessed in both current-year (ROA_{t0} and ROE_{t0}) and following-year (ROA_{t+1} and ROE_{t+1}), as stated in Hypothesis 1 and 2, respectively. Thomson Reuters ESG reported firm performance only for some companies and only through the year 2013. To take full advantage of the variables available across multiple databases, we downloaded data for operating profit, total assets, and total shareholders' equity from Worldscope and calculated outcome measures for the years 2002 through 2016. Following Ben-Oz and Greve (2015), we used operating profit to calculate ROA and ROE. Operating profit reflects the core activities of the firm and excludes the effects of income taxes and financing practices that determine interest expenses, both of which vary widely across countries and may be affected by other factors unique to individual firms. Total Assets consist of all assets a firm owns and uses to generate revenue, including tangible items like inventory, real property, plants and equipment, and intangible items like patents and trademarks. Total Shareholders' Equity is a measure of total investment in the firm by shareholders, primarily composed of book value (amount paid by investors) of both common and preferred shares, and the firm's retained earnings, which are after-tax profits that have not yet been distributed to shareholders. We constructed ROA_{t0} as current-year operating profit divided by current-year total assets and ROE_{t0} as current-year operating profit divided by current-year total shareholders' equity. We used the same formulae to construct ROA_{t+1} and ROE_{t+1} using Operating Profit, Total Assets and Total Shareholders' Equity for the following year.

Control variables. For the firm-level control firm size, we used the variable Number of Employees from Thomson Reuters ESG. Firm size is highly skewed to the right and log-transformation decreases the influence of extreme values. As recommended by Becker et al. (2019), we conducted our analysis using both the transformed and untransformed firm size and found similar results. Therefore, for ease of interpretability, we report results using the untransformed scores only. For the industry-level control industry manager gender diversity, we used Manager Gender Diversity from Thomson Reuters ESG and the Industrial Group Number (an industry classification code) from Datastream. We used all available firm-year observations of manager gender diversity to generate this variable rather than using our (smaller) final sample, ensuring estimates as close as possible to population-level industry-average values. Our sample contains firms classified into 106 different industry types. For country-level controls we used the variable Gender Egalitarianism from GLOBE (see Table SI in the online supplement for a description of this database). Also, we used the variable Labor Force Participation Ratio from the ILOSTAT database, defined as the percentage of women 15 years old or older in the

formal labour force, divided by the same measure for men, for each country-year in our sample.

Because our sample contained data for each firm across multiple years, some of the control variables varied across the years within the sample. For example, the firm-level control firm size varied across the years for each firm in the database. This is to be expected given natural patterns of hiring and employee exits from year to year. However, fluctuations in firm size over time are likely to take time to have significant effects on the overall dynamics of the firm (George and Jones, 2000; Mitchell and James, 2001). Therefore, for the firm-level control variable firm size, we used the median value of firm size across the years for each firm in the database. Similarly, we used the median value across industry-years for the industry-level manager gender diversity and the median value across country-years for the country-level control variable, Labor Force Participation Ratio.

Missing data management. In addition to implementing listwise deletion, we investigated empirically whether data were truly missing at random. We did so by examining whether the availability of data on manager gender diversity is related to values for other variables.

Outlier management. The first outlier management approach involved classifying observations as either error, interesting, or influential outliers (Aguinis et al., 2013b). To do this, we examined the bottom 1 per cent and top 1 per cent of the distributions of both ROA and ROE (i.e., 115 and 113 observations, respectively). To check for error outliers, we first compared the value for each observation from Worldscope to data for the same variable available in the Osiris and Compustat databases (see the online supplement for database descriptions). In most cases, we had data from all three databases. If the value from Worldscope matched at least one of the other two data sources, we concluded it was not an error outlier. We considered values from Osiris and Compustat to match if they were the same sign (positive or negative) and within one standard deviation of our Worldscope data. The second outlier management approach involved winsorizing ROA and ROE at the 1 per cent and 99 per cent levels.

Data analysis approach. We conducted analyses by fitting the two separate three-level nested models as discussed earlier: (a) industry-nested model, and (b) country-nested model. In what follows, we first discuss results for the analyses using winsorized data. We report detailed results for the industry-nested model and then discuss important differences in the results compared to the country-nested model. We then discuss results based on unwinsorized data and a comparison based on inclusion or exclusion of control variables at different levels.

Results

Table V provides descriptive statistics and correlations for the winsorized dataset. We tested our hypotheses following best practices for multilevel modelling as recommended in the micro literature (Aguinis et al., 2013a). We first fitted a null model to analyse the variance components and compute intraclass correlations (ICCs) for each of the two nested models at level 2 and level 3. As shown in Table VI (Step 1 for each firm

Table V. Descriptive statistics and correlations for empirical illustration involving upward expansion of workplace diversity research

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9
Control Variables												
1 Firm size	5,643	46,568	10,631	–								
2 Industry manager gender diversity	5,643	0.24	0.10	0.07	–							
3 Gender egalitarianism	5,643	4.81	0.35	0.01	0.09	–						
4 Labour force participation ratio	5,643	0.78	0.09	0.03	0.10	0.43	–					
Predictor Variable												
5 Manager gender diversity	5,643	0.24	0.14	0.06	0.69	0.17	0.27	–				
Outcome Variables												
6 Firm performance (Return on Assets ₀)	5,643	0.08	0.08	0.01	0.18	0.06	–0.02	0.17	–			
7 Firm performance (Return on Assets _{t+1})	5,590	0.08	0.09	0.01	0.19	0.06	–0.01	0.17	0.88	–		
8 Firm performance (Return on Equity ₀)	5,582	0.24	0.25	0.06	0.19	0.11	–0.00	0.20	0.69	0.61	–	
9 Firm performance (Return on Equity _{t+1})	5,511	0.23	0.25	0.06	0.18	0.11	–0.00	0.20	0.60	0.69	0.86	–

Notes: *N*, number of firm-year observations; *M*, mean; *SD*, standard deviation; ₀, current-year; _{t+1}, following year. Correlations are based on (a) firm-year (level 1) values of all variables and (b) variables winsorized at the 1% and 99% levels. All correlations greater than 0.03, *p* < 0.05. All correlations greater than 0.04, *p* < 0.01.

Table VI. Multilevel analysis results for regressing firm performance (ROA and ROE) on manager gender diversity: Industry-nested model using winsorized data

Level and Predictors	Current-Year ROA ^a ₁₀		Current-Year ROE ^b ₁₀		Following-Year ROA ^a _{t+1}		Following-Year ROE ^b _{t+1}	
	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)
Level-1 predictor (firm-year):								
Intercept	0.084*** (0.004)	0.052*** (0.010)	0.236*** (0.012)	0.119*** (0.029)	0.081*** (0.004)	0.050*** (0.010)	0.233*** (0.012)	0.120*** (0.029)
Manager gender diversity		-0.007 (0.012)		-0.008 (0.043)		-0.000 (0.012)		-0.021 (0.044)
Level-2 control (firm):								
Firm Size		0.000 (0.000)		0.000 (0.000)		0.000 (0.000)		0.000 (0.000)
Level-2 predictor (firm):								
Average manager gender diversity		0.053*** (0.015)		0.205*** (0.054)		0.055*** (0.015)		0.221*** (0.056)
Level-3 control (industry):								
Industry manager gender diversity		0.078* (0.040)		0.269* (0.120)		0.074 (0.040)		0.233* (0.121)
Variance components:								

Table VI. Continued

Level and Predictors	Current-Year ROA ₁₀ ^a		Current-Year ROE ₁₀ ^b		Following-Year ROA _{t+1} ^a		Following-Year ROE _{t+1} ^b	
	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)	Null (Step 1)	Random Intercept/ Fixed Slope (Step 2)
Within-year (Level 1), σ_2	0.001	0.001	0.017	0.017	0.001	0.001	0.017	0.018
Between-firm (Level 2), $\tau_{\pi 00}$	0.004	0.003	0.043	0.042	0.003	0.003	0.045	0.044
Between- industry (Level 3), $\tau_{\beta 00}$	0.001	0.001	0.009	0.007	0.001	0.001	0.009	0.084
Additional Information:								
Level 2 ICC ^c	0.570***		0.618***		0.576***		0.627***	
Level 3 ICC ^c	0.193***		0.136***		0.183***		0.127***	
-2 log likelihood	-17,875	-17,899	-3,761	-3,783	-17,840	-17,863	-3,528	-3,560
Number of estimated parameters	4	8	4	8	4	8	4	8

^aNote: All models are tested on data from the years 2002–16. Results show unstandardized coefficients and standard errors (in parenthesis).

^bROA, return on assets.

^cROE, return on equity.

^dICC, intraclass correlation.

*p < 0.05; ***p < 0.001.

performance dependent variable), the ICCs showed significant variance at level 2 (firm level) and level 3 (industry level). The results for the industry-level model (see Table SII in the online supplement), also show significant level-2 and level-3 variances. This confirmed the need for multilevel analysis with variables at three levels for both the industry- and county-nested models.

Next, implementing a random intercept and fixed slope model, we entered manager gender diversity predicting firm performance at level 1 (firm-year level). At level 2 (firm level), we entered the firm-level control (firm size) and the average manager gender diversity across firm years. Finally, we also entered controls at level 3 based on the model to be tested: industry manager gender diversity for the industry-nested model and gender egalitarianism and labour force participation ratio for the country-nested model. Regarding centring, we followed recommendations in the multilevel literature for modelling longitudinal multilevel data (Hamaker and Muthén, 2020). Specifically, we group-mean centred the level-1 (firm-year) manager gender diversity predictor to estimate the within-firm slope and added the average of the level-1 predictor to the level-2 (firm) equation to estimate the between-firm slope.

Outlier management. We found 19 observations for ROA_{t0} and 35 for ROE_{t0} that did not match, so we classified them as error outliers and removed them. The same was true for 21 observations for ROA_{t+1} and 36 for ROE_{t+1} . We also identified two observations for ROA_{t0} and three for ROE_{t0} (as well as seven observations for ROA_{t+1} and six for ROE_{t+1}) where data were missing from both Osiris and Compustat and removed them. After addressing error outliers, we still had several observations that may be ‘interesting outliers’ due to their extreme values on one or more variables.

As a second approach to outlier management, we winsorized the outcome variables. The results in Table VI for the industry-nested model show a significant level-2 between-firm relation between manager gender diversity and all current- and following-year firm performance variables: ROA_{t0} ($\gamma = 0.053$, $p < 0.001$); ROE_{t0} ($\gamma = 0.205$, $p < 0.001$); ROA_{t+1} ($\gamma = 0.055$, $p < 0.001$); and ROE_{t+1} ($\gamma = 0.221$, $p < 0.001$). In sum, for the industry-nested model, the predicted relation between manager gender diversity and firms’ performance was fully supported for both current-year performance (Hypothesis 1) and following-year performance (Hypothesis 2). For the country-nested model (see Table SII in the online supplement), Hypotheses 1 and 2 also received empirical support: ROA_{t0} ($\gamma = 0.079$, $p < 0.001$); ROE_{t0} ($\gamma = 0.281$, $p < 0.001$); ROA_{t+1} ($\gamma = 0.081$, $p < 0.001$); and ROE_{t+1} ($\gamma = 0.281$, $p < 0.001$). The relation between manager gender diversity and firms’ performance based on unwinsorized outcome variables with removal of error outliers was also significant for all current- and following-year firm performance variables for both the industry-nested and country-nested models (see Tables SIII and SIV in the online supplement for detailed results based on unwinsorized outcome variables).

Results based on inclusion or exclusion of control variables at different levels. Table V shows that all of the control variables at different levels were significantly correlated with manager gender diversity and/or at least one of the firm performance variables, offering empirical justification for their inclusion as statistical controls. To understand the impact of the control variables on our study results, we followed best-practice recommendations

(Bernerth and Aguinis, 2016) and tested the hypotheses with and without inclusion of each control variable to assess the resulting change in the relation between manager gender diversity and firms' performance. All coefficients for this relation in both the industry-nested and the country-nested models were positive and statistically significant for all firm performance variables (ROA_{t0} , ROE_{t0} , ROA_{t+1} , ROE_{t+1}) with and without controls. These coefficients changed in size by an average of 7.5 per cent when controls were excluded, but this change is not sufficiently large to change our substantive conclusions about Hypotheses 1 and 2. In sum, substantive results and conclusions remained unchanged by including or excluding control variables. We report results with controls given that they are theoretically justified, related to the study variables (per Table V), and therefore yield more precise coefficient estimates for the hypothesized relations.

Discussion and Implications

The results fully supported the hypothesized positive relation between manager gender diversity and firms' performance in both current- and following-year. This addresses Joshi et al.'s (2011) call for more research to understand the firm-level consequences of manager gender diversity and extends existing research in this area by examining this relationship using a sample of firms in multiple industries and countries. We found support for the study hypotheses with a model that included industry-level statistical controls and one that included country-level controls. We discuss these results in this section with a focus on the methodological choices and procedures related to the benefits that macro archival databases can provide.

Research design and measurement. We were able to test the hypothesized relations using both current-year and following-year performance. Finding support for following-year performance (ROA_{t+1} , ROE_{t+1}) in addition to current-year performance (ROA_{t0} , ROE_{t0}) provides evidence regarding a putative causal relation between manager gender diversity and firms' performance. This was made possible by the use of macro archival databases, which allow researchers, guided by theory, to select variables across an appropriate timeframe based on their study's hypotheses. Another advantage was being able to use multiple measures of the same underlying construct (i.e., firm performance). We expected ROA and ROE to covary strongly because they both represent accounting measures of firm performance, but we found the correlation between these two measures for both current and following year was 0.69, which means they only share 47.61 per cent of their variance. The results for ROA and ROE were consistent in spite of their less-than-50 per cent variance overlap, which strengthens support for the firm performance-enhancing effects of manager gender diversity.

Control variables. We tested relations of interest using multiple theoretically justified macro-contextual statistical controls at the firm, industry, and country levels. The limited number of existing studies related to manager gender diversity included samples constrained to a single industry and/or country, which precluded consideration of controls at levels higher than the firm. The results of our empirical illustration justified inclusion of the industry- and country-level controls because they were significantly correlated with the

predictor and/or outcome variables. In addition, we found significant relations between manager gender diversity and all the firm performance variables for both the industry- and country-nested models with statistical control variables included. Another important finding enabled by using macro archival database data was the significant relations between industry manager gender diversity and three of the firm performance measures (ROA_{t0} , ROE_{t0} , ROE_{t+1}) for the industry-nested model based on winsorized data. This result supports theorizing that firms may benefit from higher levels of gender diversity in their industry and points to the importance of considering diversity at the industry-level in future studies of workforce gender diversity effects. This type of relation involving the firm and industry levels of analysis is difficult to estimate if a study relies on surveys or interviews.

Macro contextual variables such as those we currently consider as control variables in our study might also be used to extend theorizing in our empirical illustration by hypothesizing boundary conditions of the predicted positive relation between workforce diversity and firm performance. For example, researchers might examine the relationship between firm age, the national cultural belief of past-oriented vs. future-oriented (i.e., importance of tradition), and national cultural dimension of gender egalitarianism on the pace of change in gender diversity in organizations.

Missing data management. We used a larger percentage of the data available (i.e., reduced the level of missingness) for our predictor variables by compensating for missing data in Thomson Reuters ESG with outcome measures constructed from data available in a separate database (i.e., Worldscope). The low percentage of missingness in our sample and large sample size allowed for the use of listwise deletion. We also found two patterns of non-randomness: firm size and country location. Specifically, data on manager gender diversity are more likely to be reported by larger firms and by firms in developed countries, which suggests there might be a limitation on the generalizability of our findings to firms with these characteristics. This might have remained unnoticed if we had not leveraged the broad set of information (variables) available in macro archival databases to diagnose patterns of missing data. As with all data collection techniques, there may be limitations on the generalizability of findings using macro archival databases and researchers should clearly specify the scope of their sample and the implications for the generalization of their study results. Diagnosing patterns of missing data can provide a more complete understanding of a study's generalizability and, in turn, assist in theorizing about underlying, under-examined relationships.

Outlier management. We demonstrated how macro archival databases enable multiple approaches to outlier management, which led to new insights, facilitated using more than one approach to outlier management. First, we identified error outliers and eliminated them. But, following best practices in outlier management, we also uncovered additional extreme observations that may be 'interesting outliers'. One example is the observation for Avon Products from 2014. Avon had the fifth-highest value on manager gender diversity (70 per cent) in that year of our sample, and it also had the 10th-highest value of ROE that year (177 per cent). This is an interesting outlier and worthy of further investigation, which was made possible by using macro archival databases to examine

upwardly expanded micro research questions. In addition, we found that the mean ROE for the personal products industry is 71 per cent, roughly three times the mean for all firms. Knowing this, Avon may not appear to be such an extreme outlier within this context. These results also suggest the personal products industry may be an interesting sector for future research on boundary conditions that link manager gender diversity with firms' performance.

We also implemented winsorizing to reduce the influence of extreme outliers. The relations between manager gender diversity and both measures of firm performance in current and following year were significant for both the industry- and country-nested models using winsorized and unwinsorized firm performance variables. Thus, we were able to demonstrate support for the study hypotheses using different approaches to outlier management, which was made possible due to the broad set of information and large sample sizes macro archival databases contain. As another point of interest, the significant relations between the industry manager gender diversity control variable and three of the firm performance variables (ROA_{t0} , ROE_{t0} , ROE_{t+1}) based on winsorized data were not significant using unwinsorized data. This further supports the notion that extreme values for certain industries might be influencing the results, suggesting a need for future research focused on the role of industry in understanding manager gender diversity effects.

In sum, our empirical illustration demonstrates the feasibility and methodological benefits of using macro archival databases for upward theoretical expansion in the micro domain of workplace diversity using samples with organizations in different industries and countries.

CONCLUSION

Expanding micro research to the macro level (i.e., to the organizational and higher levels of analysis), where appropriate, can be a useful approach for extending theory in micro research; yet, it presents a persistent methodological challenge. Meeting this challenge requires formulating relevant research questions, hypotheses and propositions; but a necessary subsequent step involves testing the veracity of the hypotheses and propositions empirically. Macro archival databases are an underutilized methodological resource that can supplement more traditional micro research techniques when upward expansion of micro research is the goal. They allow for empirical testing, which is a documented need based on literature reviews in many micro domains including turnover, training and development, compensation and reward systems, performance appraisal and management, workplace diversity, employee health and wellbeing, and creativity (see Table I). As a practical resource to facilitate future research, we provided an online supplement that includes a detailed summary of 31 commonly used macro archival databases containing organizational-level aggregations of variables typically measured at the individual and team level as well as additional variables describing organizations, industries, and countries (e.g., societal, economic, legal and cultural data) that are useful for testing research questions involving upward expansion of micro research. In addition, we described how the large number of variables and sample sizes in macro archival databases provide

methodological benefits related to research design and measurement, control variables as well as missing data and outlier management that lead to useful insights for existing theory as well as future research directions. We hope our article and online supplement will serve as a practical resource and catalyst for micro researchers to use macro archival databases to make theory advancements by meeting the persistent calls to expand micro research to the macro level.

ACKNOWLEDGEMENTS

We thank General Editor Donald Siegel, an anonymous *Journal of Management Studies* (JMS) Editor, and three JMS anonymous reviewers for highly constructive feedback. A previous version of this manuscript was presented at the (virtual) annual meetings of the Society for Industrial and Organizational Psychology, April 2021.

REFERENCES

- Aguinis, H., Boyd, B. K., Pierce, C. A. and Short, J. C. (2011). 'Walking new avenues in management research methods and theories: Bridging micro and macro domains'. *Journal of Management*, **37**, 395–403.
- Aguinis, H. and Edwards, J. R. (2014). 'Methodological wishes for the next decade and how to make wishes come true'. *Journal of Management Studies*, **51**, 143–74.
- Aguinis, H. and Glavas, A. (2019). 'On corporate social responsibility, sensemaking, and the search for meaningfulness through work'. *Journal of Management*, **45**, 1057–86.
- Aguinis, H., Gottfredson, R. K. and Culpepper, S. A. (2013a). 'Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling'. *Journal of Management*, **39**, 1490–528.
- Aguinis, H., Gottfredson, R. K. and Joo, H. (2013b). 'Best-practice recommendations for defining, identifying, and handling outliers'. *Organizational Research Methods*, **16**, 270–301.
- Aguinis, H., Hill, N. S. and Bailey, J. (2021). 'Best practices in data collection and preparation: Recommendations for reviewers, editors, and authors'. *Organizational Research Methods*, **24**, 678–93. <https://doi.org/10.1177/1094428119836485>
- Aguinis, H. and Molina-Azorin, J. F. (2015). 'Using multilevel modeling and mixed methods to make theoretical progress in microfoundations for strategy research'. *Strategic Organization*, **13**, 353–64.
- Aguinis, H., Ramani, R. S. and Villamor, I. (2019). 'The first 20 years of *Organizational Research Methods*: Trajectory, impact, and predictions for the future'. *Organizational Research Methods*, **22**, 463–89.
- Andreuski, G., Richard, O. C., Shaw, J. D. and Ferrier, W. J. (2014). 'Racial diversity and firm performance: The mediating role of competitive intensity'. *Journal of Management*, **40**, 820–44.
- Bainbridge, H. T. J., Sanders, K., Cogin, J. A. and Lin, C.-H. (2017). 'The pervasiveness and trajectory of methodological choices: A 20-year review of human resource management research'. *Human Resource Management*, **56**, 887–913.
- Becker, T. E., Robertson, M. M. and Vandenberg, R. J. (2019). 'Nonlinear transformations in organizational research: Possible problems and potential solutions'. *Organizational Research Methods*, **22**, 831–66.
- Ben-Oz, C. and Greve, H. R. (2015). 'Short- and long-term performance feedback and absorptive capacity'. *Journal of Management*, **41**, 1827–53.
- Bernerth, J. and Aguinis, H. (2016). 'A critical review and best-practice recommendations for control variable usage'. *Personnel Psychology*, **69**, 229–83.
- Burton, M. D., Fairlie, R. W. and Siegel, D. (2019). 'Introduction to a special issue on entrepreneurship and employment: Connecting labor market institutions, corporate demography, and human resource management practices'. *Industrial & Labor Relations Review*, **72**, 1050–64.
- Buyl, T., Boone, C., Hendriks, W. and Matthyssens, P. (2011). 'Top management team functional diversity and firm performance: The moderating role of CEO characteristics'. *Journal of Management Studies*, **48**, 151–77.
- Cheema, J. (2014). 'Some general guidelines for choosing missing data handling methods in educational research'. *Journal of Modern Applied Statistical Methods*, **13**, 53–75.
- Cox, T. H., Lobel, S. and McLeod, P. (1991). 'Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task'. *Academy of Management Journal*, **34**, 827–47.

- Crow, S. M., Fok, L. Y., Hartman, S. J. and Payne, D. M. (1991). 'Gender and values: What is the impact on decision making?'. *Sex Roles*, **25**, 255–66.
- Dalton, D. R. and Aguinis, H. (2013). 'Measurement malaise in strategic management studies: The case of corporate governance research'. *Organizational Research Methods*, **16**, 88–99.
- DeNisi, A. S. and Murphy, K. R. (2017). 'Performance appraisal and performance management: 100 years of progress?'. *Journal of Applied Psychology*, **102**, 421–33.
- DeNisi, A. S. and Smith, C. E. (2014). 'Performance appraisal, performance management, and firm-level performance: A review, a proposed model, and new directions for future research'. *Academy of Management Annals*, **8**, 127–79.
- Dhaliwal, D. S., Lamoreaux, P. T., Lennox, C. S. and Mauler, L. M. (2015). 'Management influence on auditor selection and subsequent impairments of auditor independence during the post-SOX period'. *Contemporary Accounting Research*, **32**, 575–607.
- Dwyer, S., Richard, O. C. and Chadwick, K. (2003). 'Gender diversity in management and firm performance: The influence of growth orientation and organizational culture'. *Journal of Business Research*, **56**, 1009–19.
- Echambadi, R., Campbell, B. and Agarwal, R. (2006). 'Encouraging best practice in quantitative management research: An incomplete list of opportunities'. *Journal of Management Studies*, **43**, 1801–20.
- Eckardt, R., Crocker, A., Ahn, Y., Floyd, S. W., Boyd, B. K., Hodgkinson, G. P., Kozlowski, S. W., Moliterno, T. P. and Starbuck, W. H. (2019). 'Reflections on the micro–macro divide: Ideas from the trenches and moving forward'. *Strategic Organization*, **17**, 385–402.
- Faems, D., Filatotchev, I., Harley, B. and Siegel, D. (2016). 'The distinctive identity of the Journal of Management Studies'. *Journal of Management Studies*, **53**, 1109–12.
- Felin, T., Foss, N. J. and Ployhart, R. E. (2015). 'The microfoundations movement in strategy and organization theory'. *The Academy of Management Annals*, **9**, 575–632.
- George, J. M. and Jones, G. R. (2000). 'The role of time in theory and theory building'. *Journal of Management*, **26**, 657–84.
- Greco, L. M., O'Boyle, E. H., Cockburn, B. S. and Yuan, Z. (2018). 'Meta-analysis of coefficient alpha: A reliability generalization study'. *Journal of Management Studies*, **55**, 583–618.
- Hamaker, E. L. and Muthén, B. (2020). 'The fixed versus random effects debate and how it relates to centering in multilevel modeling'. *Psychological Methods*, **25**, 365–79.
- Hancock, J. I., Allen, D. G., Bosco, F. A., McDaniel, K. R. and Pierce, C. A. (2013). 'Meta-analytic review of employee turnover as a predictor of firm performance'. *Journal of Management*, **39**, 573–603.
- Hausknecht, J. P. (2017). 'Collective turnover'. *Annual Review of Organizational Psychology and Organizational Behavior*, **4**, 527–44.
- Hitt, M. A., Beamish, P. W., Jackson, S. E. and Mathieu, J. E. (2007). 'Building theoretical and empirical bridges across levels: Multilevel research in management'. *Academy of Management Journal*, **50**, 1385–99.
- Hom, P. W., Lee, T. W., Shaw, J. D. and Hausknecht, J. P. (2017). 'One hundred years of employee turnover theory and research'. *Journal of Applied Psychology*, **102**, 530–45.
- Hoobler, J. M., Masterson, C. R., Nkomo, S. M. and Michel, E. J. (2018). 'The business case for women leaders: Meta-analysis, research critique, and path forward'. *Journal of Management*, **44**, 2473–99.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W. and Gupta, V. (2004). *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies*. New York: Sage Publishing.
- Jiang, K., Takeuchi, R. and Lepak, D. P. (2013). 'Where do we go from here? New perspectives on the black box in strategic human resource management research'. *Journal of Management Studies*, **50**, 1448–80.
- Johns, G. (2006). 'The essential impact of context on organizational behavior'. *Academy of Management Review*, **31**, 386–408.
- Josefy, M., Kuban, S., Ireland, R. D. and Hitt, M. A. (2015). 'All things great and small: Organizational size, boundaries of the firm, and a changing environment'. *Academy of Management Annals*, **9**, 715–802.
- Joshi, A., Liao, H. and Roh, H. (2011). 'Bridging domains in workplace demography research: A review and reconceptualization'. *Journal of Management*, **37**, 521–52.
- Khanna, P., Jones, C. D. and Boivie, S. (2014). 'Director human capital, information processing demands, and board effectiveness'. *Journal of Management*, **40**, 557–85.
- Kozlowski, S. (2017). 'Reflections on the *Journal of Applied Psychology* for 2009 to 2014: Infrastructure, operations, innovations, impact, evolution, and desirable directions'. *Journal of Applied Psychology*, **102**, 580–88.
- Kunze, F., Boehm, S. and Bruch, H. (2013). 'Organizational performance consequences of age diversity: Inspecting the role of diversity-friendly HR policies and top managers' negative age stereotypes'. *Journal of Management Studies*, **50**, 413–42.

- Leitch, S. and Palmer, I. (2010). 'Analysing texts in context: Current practices and new protocols for critical discourse analysis in organization studies'. *Journal of Management Studies*, **47**, 1194–212.
- Liao, H. and Chuang, A. (2004). 'A multilevel investigation of factors influencing employee service performance and customer outcomes'. *Academy of Management Journal*, **47**, 41–58.
- Liu, K. (2014). 'Human capital, social collaboration, and patent renewal within U.S. pharmaceutical firms'. *Journal of Management*, **40**, 616–36.
- Liu, X., van Jaarsveld, D. D., Batt, R. and Frost, A. C. (2014). 'The influence of capital structure on strategic human capital: Evidence from U.S. and Canadian firms'. *Journal of Management*, **40**, 422–48.
- Lyness, K. S. and Grotto, A. R. (2018). 'Women and leadership in the United States: Are we closing the gender gap?'. *Annual Review of Organizational Psychology and Organizational Behavior*, **5**, 227–65.
- Maloney, M. M., Bresman, H., Zellmer-Bruhn, M. E. and Beaver, G. R. (2016). 'Contextualization and context theorizing in teams research: A look back and a path forward'. *Academy of Management Annals*, **10**, 891–942.
- Mathieu, J. E., Aguinis, H., Culpepper, S. A. and Chen, G. (2012). 'Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling'. *Journal of Applied Psychology*, **97**, 951–66.
- Mathieu, J. E. and Chen, G. (2011). 'The etiology of the multilevel paradigm in management research'. *Journal of Management*, **37**, 610–41.
- Metz, I. (2009). 'Organisational factors, social factors, and women's advancement'. *Applied Psychology*, **58**, 193–213.
- Miller, T. and del Carmen Triana, M. (2009). 'Demographic diversity in the boardroom: Mediators of the board diversity–firm performance relationship'. *Journal of Management Studies*, **46**, 755–86.
- Mitchell, T. R. and James, L. R. (2001). 'Building better theory: Time and the specification of when things happen'. *Academy of Management Review*, **26**, 530–47.
- Molloy, J. C., Ployhart, R. E. and Wright, P. M. (2011). 'The myth of “the” micro-macro divide: Bridging system-level and disciplinary divides'. *Journal of Management*, **37**, 581–609.
- Morgeson, F. P., Aguinis, H., Waldman, D. A. and Siegel, D. S. (2013). 'Extending corporate social responsibility research to the human resource management and organizational behavior domains: A look to the future'. *Personnel Psychology*, **66**, 805–24.
- Moss, R. W. and Ernsthausen, D. G. (2014). *Strauss's Handbook of Business Information: A Guide for Librarians, Students, and Researchers*, 3rd edition. Santa Barbara, CA: Libraries Unlimited.
- Newman, D. A. (2014). 'Missing data: Five practical guidelines'. *Organizational Research Methods*, **17**, 372–411.
- Nielsen, B. B. and Raswant, A. (2018). 'The selection, use, and reporting of control variables in international business research: A review and recommendations'. *Journal of World Business*, **53**, 958–68.
- Noe, R. A., Clarke, D. M. and Klein, H. J. (2014). 'Learning in the twenty-first-century workplace'. *Annual Review of Organizational Psychology and Organizational Behavior*, **1**, 245–75.
- Ormiston, M. E. and Wong, E. M. (2013). 'License to ill: The effects of corporate social responsibility and CEO moral identity on corporate social irresponsibility'. *Personnel Psychology*, **66**, 861–93.
- Park, D. (1996). 'Gender role, decision style and leadership style'. *Women in Management Review*, **11**, 13–17.
- Ployhart, R. E. and Hale, D. Jr. (2014). 'The fascinating psychological microfoundations of strategy and competitive advantage'. *Annual Review of Organizational Psychology and Organizational Behavior*, **1**, 145–72.
- Porter, W. P. and Schneider, B. (2014). 'What was, what is and what may be in OP/OB'. *Annual Review of Organizational Psychology and Organizational Behavior*, **1**, 1–21.
- Roberson, Q. M. (2019). 'Diversity in the workplace: A review, synthesis, and future research agenda'. *Annual Review of Organizational Psychology and Organizational Behavior*, **6**, 69–88.
- Roberson, Q. M., Holmes, O. IV and Perry, J. L. (2017a). 'Transforming research on diversity and firm performance: A dynamic capabilities perspective'. *Academy of Management Annals*, **11**, 189–216.
- Roberson, Q. M., Ryan, A. and Ragins, B. (2017b). 'The evolution and future of diversity at work'. *Journal of Applied Psychology*, **102**, 483–99.
- Roberts, K. H., Hulin, C. L. and Rousseau, D. M. (1978). *Developing an Interdisciplinary Science of Organizations*, 1st edition. San Francisco, CA: Jossey-Bass.
- Schafer, J. L. and Graham, J. W. (2002). 'Missing data: Our view of the state of the art'. *Psychological Methods*, **7**, 147–77.
- Schwab, A., Werbel, J. D., Hofmann, H. and Henriques, P. L. (2016). 'Managerial gender diversity and firm performance: An integration of different theoretical perspectives'. *Group & Organization Management*, **41**, 5–31.
- Shaw, J. D. (2014). 'Pay dispersion'. *Annual Review of Organizational Psychology and Organizational Behavior*, **1**, 261–93.

- Shrader, C. B., Blackburn, V. B. and Iles, P. (1997). 'Women in management and firm financial performance: An exploratory study'. *Journal of Managerial Issues*, **9**, 355–72.
- Solarino, A. M. and Aguinis, H. (2021). 'Challenges and best-practice recommendations for designing and conducting interviews with elite informants'. *Journal of Management Studies*, **58**, 649–72.
- Tetrick, L. E. and Winslow, C. J. (2015). 'Workplace stress management interventions and health promotion'. *Annual Review of Organizational Psychology and Organizational Behavior*, **2**, 583–603.
- Tzabbar, D. and Kehoe, R. R. (2014). 'Can opportunity emerge from disarray? An examination of exploration and exploitation following star scientist turnover'. *Journal of Management*, **40**, 449–82.
- van Knippenberg, D., De Dreu, C. K. and Homan, A. C. (2004). 'Work group diversity and group performance: An integrative model and research agenda'. *Journal of Applied Psychology*, **89**, 1008–22.
- Williams, K. Y. and O'Reilly, C. A. (1998). 'Demography and diversity in organizations: A review of 40 years of research'. *Research in Organizational Behavior*, **20**, 77–140.
- Wright, P. M., Coff, R. and Moliterno, T. P. (2014). 'Strategic human capital: Crossing the great divide'. *Journal of Management*, **40**, 353–70.
- Wright, P. M., Nyberg, A. J. and Ployhart, R. E. (2018). 'A research revolution in SHRM: New challenges and new research directions'. *Research in Personnel and Human Resources Management*, **36**, 141–61.
- Wu, P.-C. and Chaturvedi, S. (2009). 'The role of procedural justice and power distance in the relationship between high performance work systems and employee attitudes: A multilevel perspective'. *Journal of Management*, **35**, 1228–47.
- Zhou, J. and Hoever, I. J. (2014). 'Research on workplace creativity: A review and redirection'. *Annual Review of Organizational Psychology and Organizational Behavior*, **1**, 333–59.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web site: