

## Secondary Publication



Tubik, Daniel; Herberger, Tim Alexander

### **How endogeneity problems are addressed in analyzing the relationship between diversity in top management teams and company financial performance : a systematic literature review**

Date of secondary publication: 11.02.2026

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-113098x

#### **Primary publication**

Tubik, Daniel; Herberger, Tim Alexander (2026): How endogeneity problems are addressed in analyzing the relationship between diversity in top management teams and company financial performance : a systematic literature review, in: Management review quarterly, Cham: Springer, Vol. 76, Nr. 1, pp. 127–156, doi: 10.1007/s11301-024-00476-3.

#### **Legal Notice**

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>



# How endogeneity problems are addressed in analyzing the relationship between diversity in top management teams and company financial performance—a systematic literature review

Daniel Tubik<sup>1</sup> · Tim Alexander Herberger<sup>1,2,3,4</sup>

Received: 8 February 2024 / Accepted: 1 November 2024 / Published online: 16 December 2024  
© The Author(s) 2024

## Abstract

Endogeneity is a primary concern for empirical research that examines the relationship between firm leadership and company performance. This is especially true for demographic heterogeneity in boardrooms and top management teams, where reverse causalities are prevalent. Based on a systematic literature review, we contribute to the current state of empirical literature by highlighting the effects of mitigating endogeneity methodically, showing how studies differ in various types of demographic heterogeneity variables and endogeneity checks, and how all these influence performance outcomes in research. We thus offer a new perspective on the numerous empirical findings already available. Our findings show that most studies acknowledge and check for endogeneity concerns but do so by employing different methodical approaches. We further show that the utilization of various techniques affects the variability in study outcomes regarding relationship directions. Papers employing more sophisticated methods (e.g., System GMM) have a higher share of positive significant findings than our overall sample. Still, the general trend that an overwhelming share of papers finds at least some significant positive relationship between leadership demographic heterogeneity and financial performance holds firm all the techniques employed. We consider research created from 1990 to 2024 for our sample of 163 publications.

**Keywords** Diversity · Endogeneity · Financial performance · GMM · Systematic literature review · Top management teams

## Zusammenfassung

Endogenität stellt ein zentrales Problem für die empirische Forschung dar. Dies gilt besonders für die demografische Heterogenität in Vorständen und Top-Management-Teams, bei denen umgekehrte Kausalitäten in Bezug auf die Unternehmensperformance und damit verbundenen Kennzahlen häufig auftreten können. Basierend

auf einer systematischen Literaturanalyse leisten wir einen Beitrag zur aktuellen empirischen Forschung, indem wir das Instrumentarium gegen die negativen Auswirkungen von Endogenität untersuchen. Wir analysieren, wie Studien in diesem Themenfeld sich in Bezug auf verschiedene Arten von Variablen zur demografischen Heterogenität und Endogenitätsprüfung unterscheiden und wie all dies die Ergebnisse zur Unternehmensleistung beeinflussen kann. Damit eröffnen wir interessante Einblicke auf die bereits zahlreich vorhandenen empirischen Befunde in diesem Forschungsfeld. Unsere Ergebnisse zeigen, dass die meisten Studien auf etwaige Endogenitätsprobleme prüfen und gegebenenfalls nachsteuern, jedoch unterschiedliche methodische Ansätze hierfür verwendet werden. Zudem zeigen wir, dass der Einsatz verschiedener Techniken die Variabilität der Studienergebnisse in Bezug auf die Richtung der Beziehungen beeinflussen kann. Studien, die fortschrittlichere Methoden wie beispielsweise System-GMM einsetzen, weisen einen höheren Anteil signifikanter positiver Ergebnisse im Vergleich zum gesamten untersuchten Literaturmaterial. Dennoch bleibt die Grundaussage aus diesem Forschungsbereich bestehen, dass ein sehr großer Anteil der Studien mindestens einen signifikant positiven Zusammenhang zwischen der demografischen Heterogenität in der Führungsebene und der finanziellen Unternehmensleistung feststellt – unabhängig von der methodischen Vorgehensweise. In unserer Stichprobe von 163 Peer-Review-Publikationen haben wir Forschungsarbeiten aus den Jahren 1990 bis 2024 berücksichtigt.

**JEL Classification** C18 · G15 · G30 · M14

## 1 Introduction

Empirical research regarding the nexus of company leadership demographic heterogeneity and firm financial performance has enormously expanded in recent years, with often inconclusive outcomes, broadly contradictory findings, and multiple challenges regarding the interpretation of results.

Besides the usual problems of empirical research designs—e.g., varying regional legislative setups regarding company leadership structures, differing performance variables used during testing, vast methodical differences, the cacophony of diversity measurement techniques utilized, a broad variety of statistical methodologies employed—a major critique of the field concentrates on the endogeneity problem of such observational research. Ahern and Dittmar (2011) e.g. warn that the endogenous nature of board diversity-related research hinders the understanding of the subject.

Endogeneity is especially relevant in the leadership-performance nexus, as nearly all the characteristics around a board or top management team and a firm's performance can be endogenous, producing joint endogeneities (Hermalin and Weisbach 2003). The exogeneity of the independent variables (diversity factors) must be secured in order to provide consistent empirical results and implications. Failing to acknowledge endogeneity issues means risking reverse causality and/or omitted variables when employing empirical models in the field, ultimately

leading to misguided conclusions for practitioners, e.g., hiring related decision-making for company leadership teams or altering existing policy frameworks from a regulatory side. Consequently, the problem of endogeneity must also be addressed in the empirical analyses of scientific papers (e.g., related techniques and measurements).

We, therefore, aim to show how aware the research field in question is of endogeneity problems, to what extent methods mitigating endogeneity concerns are employed in empirical studies to date, and how this influences research outcomes. We will show what methodological setups are trending, their drawbacks, and whether the general outcome structure of our studies handling endogeneity differs from papers that do not account for this issue. We will look at whether the problem is seen as more prevalent for certain heterogeneity dimensions (e.g., gender diversity in company leadership and decision-making teams) and whether this affects study outcomes.

Empirical research relating to the diversity of leadership teams has sprawled since the first introduction of gender quota regulations in Norway (2006). Our analysis explicitly includes further dimensions of leadership diversity (age, ethnicity, nationality, etc.) to provide a full picture of endogeneity perceptions in empirical leadership diversity literature. This way, we can also highlight possible differences in empirical designs between papers by focusing on various dimensions of top decision-maker heterogeneity. We do not limit ourselves to boardrooms and include top management team focused empirical papers into our sample of papers. This is, in many cases, a straightforward extension of boardrooms by executives and/or managers not having a seat on the board.

To best of our knowledge there are no studies to date that systematically examine to what extent endogeneity is considered in the empirical demographic heterogeneity and performance nexus literature, including the above-mentioned details regarding performance outcomes and methodical techniques employed. Hence, our work contributes to a better understanding of how much endogeneity concerns influence outcomes in the trending empirical field of demographic heterogeneity literature.

We find that most studies to date include techniques mitigating endogeneity concerns and that the share of papers doing so is growing over time. Most research from our sample found at least some positive significant relationship between leadership demographic heterogeneity and firm financial performance. The overall subsample of papers employing endogeneity mitigating techniques did not materially differ regarding this statistic. However, certain demographic heterogeneity dimensions and methods to check for endogeneity showed altering shares of outcomes. Specifically, more sophisticated methods to check for endogeneity (e.g., system GMM regressions) were utilized by papers that found more positive and significant results.

The paper is structured as follows: Sect. 2 presents the relevant theoretical background for the systematic literature review regarding heterogeneity frameworks and endogeneity mitigating techniques. Section 3 presents the applied methodology. Section 4 deals with the presentation of the results and their discussion. A conclusion of the results is presented in Sect. 5.

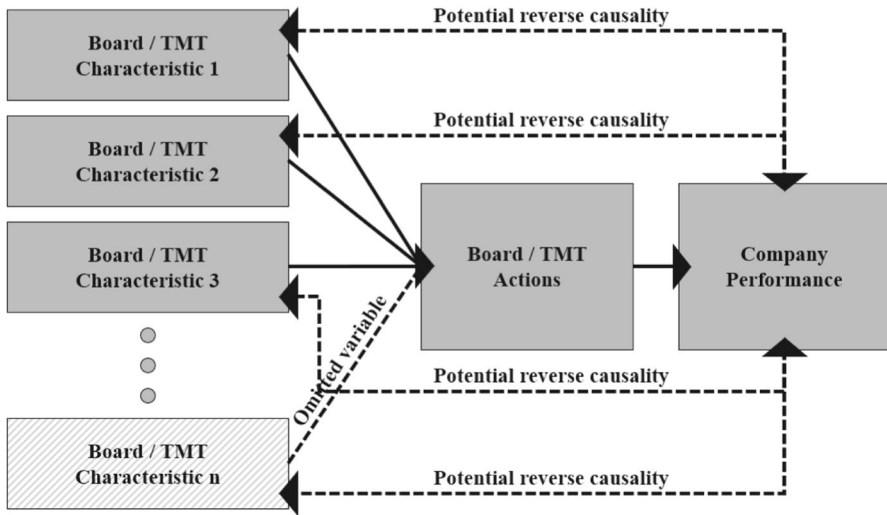
## 2 Theoretical background

Having a deeper view of the general endogeneity problem, we can introduce a general theoretical framework considering the effects of leadership heterogeneity on financial performance outcomes.

A major source of ambiguities regarding the effects of leadership heterogeneity on firm performance is endogeneity concerns. We will show in Chapter 4 that an overwhelming share of empirical papers acknowledge the endogeneity problems arising. The relationship between firm performance and board characteristics is traditionally focused on the effects of the proportion of outside directors on the board and the size of the board, with both factors well covered by empirical research (Hermalin and Weisbach 2003). Research has already been done on these two easily accessible and observable characteristics, and endogeneity has been studied. Hermalin and Weisbach (1998), e.g. find that an unsatisfactory company performance can cause board structures to get more independent as a reaction, highlighting that certain characteristics cause the board to change before it then can, in turn, change company performance through its proactive decision-making. Overall, the background of endogeneity is less prominently covered by theoretical research than by empirical tests of the subject matter (Hermalin and Weisbach 2003).

Endogeneity can be categorized into four major subtypes (Woolridge, 2010; Hill et al., 2021). Some of them are more relevant to our line of research than others (e.g., we do not dedicate significant space towards the measurement error-based endogeneity problem in the following). The categorization we follow considers omitted variables, simultaneity, measurement errors, and selection bias as major sources of endogeneity (Woolridge, 2010). Omitted variables (unobservable heterogeneity) now mean that there can be unobservable variables that might be correlated with, in our case, both company financial performance and leadership heterogeneity. However, they are not included as variables in empirical testing. Simultaneity (or reverse causality) refers to when clear causality is blurred because the two variables in question (leadership heterogeneity and financial performance) affect the other. In our general example, this would, for instance, mean that while heterogeneous boards generate superior financial performance, firms with superior financial performance also cause heterogeneous company leadership. Measurement errors can cause endogeneity if, e.g., the independent variable in question is measured with an error, and this error is correlated with the dependent variable. A fourth major category, according to Woolridge (2010), is selection bias, which arises when either sample selection is not randomized or when the selection of treatment is not randomly assigned (Hill et al., 2021). As shown in Fig. 1, empirical research in our sample focuses on the first two categories mentioned here: omitted variables and simultaneity.

Theories concerning the possible effects of leadership heterogeneity on various types of performance are broad and well-documented in existing empirical literature. The most prominent theory utilized in our sample is agency theory (Carter et al. 2003), given that it is a framework for studying the board of directors' monitoring functions. Agency theory was first described by Berle and Means (1932) and focused on monitoring the functions of a company's board. Fama and Jensen



**Fig. 1** Endogeneity problems related to company leadership and performance. (based on Hermalin and Weisbach 2003 and Hill et al. 2021). This Figure shows potential forms of endogeneity especially relevant for empirical studies in our sample. One form highlighted here is reverse causality between the dependent and independent variables. Another endogeneity issue depicted by the Figure above is unobservable heterogeneity, where potentially relevant variables are not included in a regression analysis. TMT means Top Management Team.

(1983) and Jensen and Meckling (1976) explicitly pointed the theory toward the insider–outsider differences in boardrooms and their effects on company performance. As, e.g., Adams and Ferreira (2009) explain, gender-diverse boards allocate further resources to monitoring, showing how demographic heterogeneity interplays with agency theory. Carter et al. (2003) further describe that this might occur due to directors with varying cultural backgrounds asking more questions than culturally more comparable and equivalent colleagues, also having the possible negative effects of marginalization and exclusion of such different group members sometimes hinder a strategic efficiency decision-making process.

Another major theory describing the relationship between leadership heterogeneity and performance is Pfeffer and Salancik's (1978) resource dependency theory, in which firms are seen as open systems whose success depends on their relationships with their external environment (Lückerath-Rovers 2013). Hillman et al. (2009) explain that a heterogenous leadership group can better mitigate uncertainties since a diverse group better covers a firm's diverse external environment. Essentially, the heterogenous leadership team has more power to acquire resources and overcome interdependencies.

Upper echelons theory (Hambrick and Mason 1984) focuses on the individual experiences, values, and personalities of decision-makers and states that these factors matter significantly for the situation-based individual decision-making of company leaders. In the relevant empirical diversity literature, demographic

heterogeneity is seen as something positive, enhancing the collective experience of a group, enriching the decision depth, and influencing firm performance.

Tokenism (Kanter and Bluedorn, 1977) is the idea that minority members are only included in a decision-making body to communicate openness to outsiders, and at best, they have limited effects on decision-making and, hence, company performance. Konrad's critical mass theory (Konrad et al. 2008) goes further and tries to establish a threshold for the share or quantity of minority members of a group, from which the minority members have an effect on decisions.

Our sample of 163 studies also employed several further frameworks for hypothesis development, including social identity theory (Tajfel and Turner 1986), social psychology theory (Westphal and Milton 2000; Milliken and Martins 1996), stakeholder theory (Freeman 1984), and institutional theory (Selznick 1957), most of which lead towards a positive assumed relationship between performance and leadership heterogeneity (Kagzi and Guha 2018).

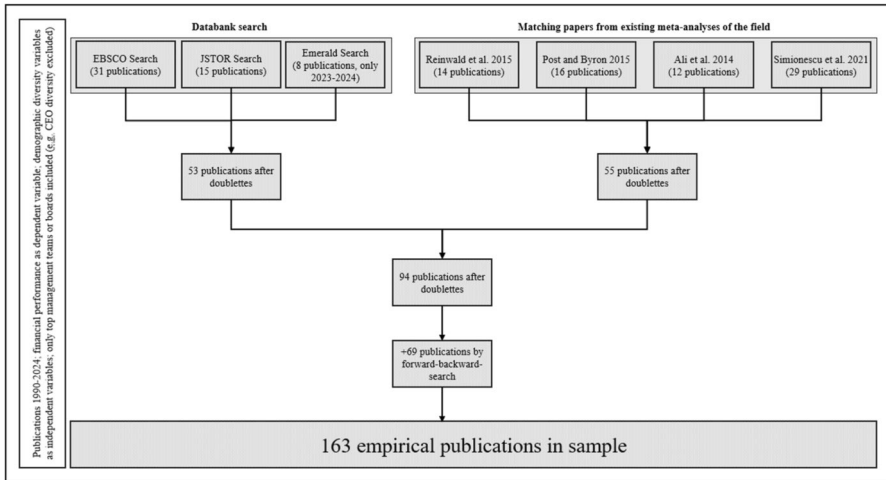
### 3 Data and methodology

The following research process is based on the 15-step systematic quantitative literature review (SQLR) developed by Pickering and Byrne (2014). This process focuses on systematic paper selection and data evaluation to compare research in a quantitative and expert-oriented way. Using this method, we minimized omissions and systematically created a broad and representative sample of the empirical landscape, keeping our analysis consistent with past research in the field and raising the applicability of our results for forthcoming studies.

Due to the methodology selected, we can show the extent of endogeneity effects on empirical leadership heterogeneity research, what methods are prevalent in mitigating these effects, and how or if attempts to handle the endogeneity problem influence performance outcomes.

This literature review differs from previous works on heterogeneity in top management teams: It does not solely concentrate on any one diversity dimension (as e.g. Velte's structured review from 2017, which only examines papers regarding gender diversity), is more quantitative than other research (e.g. compared to Baker et al. (2020), who only perform a bibliometric analysis), and focuses on the topic of the endogeneity problem, which has no precedent we are aware of. Bayly et al. (2023) dive deeper into the issue of endogeneity, systematically listing arising problems, possible solutions, and outcomes in a selection of papers. However, given the initially smaller sample size of 62 papers, of which only 14 looked at financial performance and simultaneously addressed endogeneity in the eyes of the authors adequately, the results found by the authors are not quite representative. Bayly et al. also only focused on gender diversity in their study, while we aim more broadly at the entire scope of demographic diversity when looking into the endogeneity issue.

Our primary search, as shown in Fig. 2, started with the EBSCO database based on a broad keyword search of related topics (with the keywords board diversity performance, firm leadership diversity performance, top management team diversity, heterogeneity performance, etc.), with publications starting in 1990 up until and



**Fig. 2** Sample selection process for empirical research papers considering the relationship between demographic diversity in leadership teams and financial performance, 1990–2024. This Figure shows the sample selection process for the entire sample of 163 studies for the systematic literature review, showcasing the core database and existing metaanalyses’ as major sources of the empirical diversity sample. The initial process is based on databank searches, but extend the thusly found scientific papers by a forward–backward-search. Only papers matching our initial criteria were selected for the above process (financial performance as dependent variables, demographic diversity dimensions as independent variables, diversity measurement based on top management teams or boards, publications 1990–2024, CEO or other single office holder diversity excluded)

including 2024. We extended this pool with a JSTOR search for these keywords with an identical timeframe and an Emerald Insight search for the last two years of the search period to further deepen the sample. We aimed to include as many personality-based diversity and heterogeneity dimensions of company leadership as possible in our sample, given that various factors of socioeconomic or demographic background could have a potential effect on performance based on initial diversity theories (Jensen and Meckling 1976; Pfeffer and Salancik 1978; Hambrick and Mason 1984). To hold our review as comprehensive as possible, to counter a possible publication bias (high-quality journals focusing on studies with significant connections found, mostly from prestigious institutions), to have a possibly diverse and broad set of geographies, markets, contexts covered, and methodologies represented, we explicitly did not limit our search to only high-quality journals. Instead, we included individual dissertations, studies from smaller publications, large public institutions, etc. After elimination, we identified 53 major and relevant empirical research publications from the mentioned data banks. This was then significantly extended by adding relevant research papers from existing meta-analyses’ of the empirical diversity field (Post and Byron 2015; Reinwald et al. 2015; Simionescu et al. 2021; Ali et al. 2014). Initially, we found 71 papers throughout these meta-analyses, which was reduced to 55 after eliminating double counts between these publications. Of these 55 individual papers, 14 were already found due to our above-described initial data bank search, leaving us with 41 additional empirical research papers found through

the mentioned meta-papers. Overall, with these two searches, we arrived at 94 studies. In the next step, we continued to include relevant, cited studies from this initial batch of research papers to get a comprehensive picture of current scientific empirical and theoretical literature (forward–backward-search), significantly expanding our sample to the count of 163 papers.

For each empirical research study in our sample, we recorded the data points in Table 1 to get a comprehensive picture of current developments and trends in empirical diversity literature.

We re-recorded data points for 20 randomly selected papers within our sample after six weeks of finalizing our initial dataset, yielding a 73% match of formerly recorded data (in 9 main data points recorded for each study) with our second batch of compilation. This is far from a perfect match but shows a generally stringent recording practice enough for this qualitative and quantitative evaluation of the leadership heterogeneity field.

We then screened for two factors: papers that recognized the endogeneity issue in some form and, separately papers that employed a method considered a mitigating tool for the endogeneity problem in statistical analysis. The two groups only differ in a handful of papers. The details regarding the outcomes of these screenings and their subsequent evaluation regarding the relationship between diversity in top management teams and financial performance will be discussed in Chapter 4.

Our search process did not stop at the level of boardroom heterogeneity but included studies searching for the relevance of heterogeneity in top management teams. This term usually involves more people than only the directors since senior executives sometimes have no seat on the board, but they should be included in heterogeneity tests of decision-making.

**Table 1** Data recorded for papers in 163 sample studies

Datapoint recorded	Description
Demographic heterogeneity dimensions covered (main independent variables of empirical research)	Concentrating on broader categories identified, recording individual metrics for each diversity dimension covered (e.g. gender, race/ethnicity)
Endogeneity mitigating method type	Type of endogeneity mitigating technique used in regression analysis by study (e.g. lagged variables, 2SLS, fixed effects)
Theories	Theories covered to explain diversity-effects on performance variables (e.g. agency theory)
Heterogeneity-performance relationship outcomes	Individual study findings according to relationships between heterogeneity parameters and performance variables, recorded on both the study and the heterogeneity dimension level regarding outcomes
Publishing paper	The research paper an article was published in

This Table describes all data points recorded for each of the 163 research papers of the sample, respectively. The data is comprehensive and categorized further than represented here. The heterogeneity-performance relationship is recorded on a paper basis and also on a demographic heterogeneity dimension basis

Our sample only included research that tested for the relationship between financial performance and top management-level heterogeneity. We did not include the numerous ESG-performance-centered empirical papers in our study. Furthermore, we only added empirical papers to our sample; papers with only a theoretical line of argumentation were excluded.

For each study in our sample, we accounted for various methodical frameworks used in empirical literature to handle the various forms of endogeneity. The following will highlight the broader and most representative checks and remedies for endogeneity used in our sample studies.

We accounted for studies employing some form of a lagged variable in performed regression analyses. We recorded not only if a research design included this methodological approach, but also included exactly what type of lagged variable was employed. Traditionally, papers employ lagged independent variables, since the effects of firm leadership decision-making on financial performance should be accounted for with a time lag. Our sample papers usually employ a lag of one year, with occasional studies also utilizing a two-year lag period in testing (e.g., Miller and Del Carmen Triana, 2009) and some GMM tests using multiple lags (e.g., Miletkov et al. 2012). System GMM models always include the lagged version of the dependent variable as an independent variable in the regression model as an instrument (e.g. Talavera et al. 2018). These simple lagged values of variables are easy to implement into regression models since they do not require additional data.

Although lagged (dependent) variables help solve the problem of unobserved heterogeneity, as, e.g., Cannella et al. (2008) point out, simultaneously, these variables can be highly correlated with disturbance terms. To mitigate this issue, authors can employ an Instrumental Variable (IV) approach for the lagged variable in question, essentially substituting the variable in question with an IV. As Rose et al. (2013) e.g. show, finding adequate instrumental variables for performance measures is an overall problem in research, making the lagged value approach in combination with instrumental variables complex to employ successfully. But, as Dezső and Ross (2012) state, using, e.g., only the lagged value of a dependent variable as an instrument in a panel data regression—typically employed by empirical papers of our field in question—may create problems with autocorrelation. All in all, lagged variables are an intuitive tool to resolve endogeneity issues, but ultimately, they cannot remedy all their sources independently.

We also accounted for the employment of fixed effects in regressions within the research in our sample. As Chijoke-Mgbame et al. (2020) note, fixed effects in regression models can help to account for unobservable firm-specific factors. Although this method handles unobservable heterogeneity (Bennouri et al. 2018), it only works in cases where past performance is uncorrelated to independent variables in question (e.g., heterogeneity in our case), or as Mastella et al. (2021) describe it, is caused by time-constant omitted variables. But as Bennouri et al. (2018) explain, this is not given, e.g., for the dimension of gender diversity, given that past performance can influence where female candidates for firm leadership roles decide to be employed. This dynamic endogeneity and other reverse causalities can cause fixed effect estimates to be biased. Additionally, as Masulis et al. (2012) point out, firm fixed effects regressions generally can have problems with

detecting the influence of only slightly changing independent variables over time (e.g. board characteristics, which is our focus).

Unite et al. (2019) point out that random effects can be employed to control for omitted variables besides fixed effects. Dang et al. (2020) use correlated random effects to control not only for time-invariant unobserved heterogeneity, as does fixed effects, but also to measure time-invariant explanatory variables. The authors combine random effects with a unique control function approach (essentially instrumental variables), “which addresses the potential endogeneity issue due to unobserved time-varying factors. [...] The exogenous variation induced by excluded instrumental variables provides a separate variation in the residuals obtained from a reduced form, which serve as the control functions.” Overall, this sophisticated model of Dang et al. (2020) makes the testing substantially more robust to heteroskedasticity and is therefore an interesting contribution to the econometric approaches otherwise used in empirical literature.

Two-stage least squares instrumental variable (2SLS-IV or simply 2SLS) models, as described by Arioglu (2020), are often employed to handle endogeneity and general robustness purposes. The technique uses two OLS models, first regressing the endogenous variable(s) on the (exogenous) instrumental variables and then running the second regression with the results of the first calculation as the independent variables instead of the presumed endogenous original variables. The method is based on two linear relationships and is intuitive and comparatively simple to employ. A major shortcoming of 2SLS is that it is generally suitable to handle the issue of reverse causality but can’t mitigate another source of endogeneity: the effects of past performance on leadership heterogeneity (Unite et al. 2019). Also, the challenge with 2SLS models, as with every IV application, is finding suitable instruments, as already pointed out above (Rose et al. 2013). Miletkov et al. (2012) describe that papers often combine 2SLS models with fixed effects to handle further heterogeneity, which our sample supports.

Instrumental variables (IV) are generally widespread in our sample, given that many methods described here are essentially IV-based approaches. Since the driver behind employing IVs can differ (e.g., 2SLS), we separately recorded this methodical approach’s utilization (or not-utilization) for every study in our sample. Specifically, some studies used instrumental variables added as independent variables to a simple regression and did not employ a more sophisticated approach (e.g., 2SLS) while doing so (e.g., Niikura and Seko 2020). As Miletkov et al. (2012) describe the challenges with IVs, these variables should have no effect on a firm’s operating (financial) performance but should affect the heterogeneity variables in question, and this way indirectly influence the dependent variable of a regression (in our case financial performance). The description makes it clear why finding adequate IVs is so important and challenging at the same time, leading to rather exotic choices for some papers. Within our sample, examples for IVs included the average length of education of the spouses of the other CEOs in the firm (Smith et al. 2006), the gender heterogeneity of the population in a city where a company’s HQ is located (Akram et al. 2020), US teen birth rates (Oldford et al. 2021), or age diversity of the local population at the province level (Talavera et al. 2018).

As an answer to the issue of finding adequate instrumental variables, literature promotes employing dynamic generalized method of moments (GMM) models to mitigate endogeneity concerns (Chijoke-Mgbame et al. 2020). GMM explicitly addresses endogeneity concerns and is, in its dynamic form, usually either a difference GMM or a system GMM model (Sani 2021).<sup>1</sup> Roodman (2009, p. 86) argues that the models are well suited for studies with “(1) ‘small T, large N’ panels, meaning few time periods and many individuals; (2) a linear functional relationship; (3) one left-hand-side variable that is dynamic, depending on its own past realizations; (4) independent variables that are not strictly exogenous, meaning they are correlated with past and possibly current realizations of the error; (5) fixed individual effects; and (6) heteroskedasticity and autocorrelation within individuals but not across them.” These mentioned bullet points describe the usual empirical studies in our overall sample well, where data is only tested for a low number of periods (years), with a higher number of companies (usually an index), a dependent variable (financial performance) being dynamic in nature (influenced by its past values), with fixed individual effects (company characteristics), etc. The Arellano and Bond (1991) based difference GMM technique transforms the independent variables (as Roodman points out by differencing) and then runs a GMM model. A further evolution of this GMM is the system GMM model developed by Arellano and Bover (1995) and Blundell and Bond (1998). The system GMM model assumes that “first differences of instrument variables are uncorrelated with the fixed effects” (Roodman 2009, p. 86) and permits further instrumental variables, ultimately improving regression outcomes. In practice, the models use the dependent variable’s prior (lagged) values as instruments in the regression equation. A favorable characteristic of system GMM further is, that it accounts for the issue of only slowly changing independent variables, which is naturally often the case for leadership heterogeneity in practice (Nguyen et al. 2015). As described, these models are sophisticated and can capture multiple problems arising from endogeneity, as they handle both reverse causation and time-invariant unobserved heterogeneity. But they are also complicated to comprehend and hence can lead to false readings or applications when employed. As Sani (2021) points out, the simpler difference GMM models also have efficiency issues when used for small T large N panels (which was one of the main reasons for developing the system GMM technique). Additionally, the GMM tests have the same problem as every method utilizing IVs, namely that the models are only as good as the IVs selected (Nguyen et al. 2015).

Empirical research increasingly also employs quasi-experimental tools, e.g. difference-in-difference (diff-in-diff or DD) and propensity score matching (PSM) techniques (Hesarzadeh 2020). The DD method focuses on the source of the external endogenous variable (specifically, its variability) (Gruszczyski, 2020). Unlike traditional event studies, Hesarzadeh (2020) points out that DD attempts to find an intervention or treatment, usually a passage of the law. For studies in our sample examining demographic heterogeneity in firm leadership, this is commonly the

---

<sup>1</sup> System GMM models can in some cases also be non-dynamic. Simple linear GMM models, of which OLS and 2SLS are special forms (Roodman 2009), are not employed significantly in our sample and hence are not described in detail further here.

implementation of quota regulation, as performed in multiple jurisdictions since Norway first introduced a gender quota law in the 2000s (Meyerinck et al. 2018). Papers employing DD compare the outcomes in a group affected by regulation vs. those in a control group not affected by regulation. The divergence between the two groups' results is attributable to the introduction of regulation or a change in hiring techniques. DD is especially suitable for designs where instrumental variables are hard to identify, which is, as described above, one major issue with the subject examined here. Another method, propensity score matching (PSM), is an ideal treatment for sample selection bias (Bennouri et al. 2018) and focuses on finding similar companies that only differ in the one variable in question for the testing. PSM tries to account for "the covariates that predict receiving the treatment" (Hesarzadeh 2020, p. 36) to reach higher-quality results. Gruszczyski (2020), when examining the causal effects of appointing female directors, e.g. put firms employing exactly one female director after the departure of a male in a treatment group, and a similar group of companies employing a male director after the departure of a male in the control group, matched the two groups, and then looked for significant differences to find causal relationships. The two techniques mentioned here also have their own endogeneity problems, as the interventions or treatments used for DD might show signs of endogeneity themselves (Hesarzadeh 2020).

Another method that was not explicitly employed in our sample to address endogeneity is the utilization of Bayesian methods in empirical research. As van de Schoot et al. (2014) highlight, this approach is becoming more used in social and behavioral science research thanks to advanced computational algorithms, which make its employment more convenient. Bayesian methods are a statistical approach that is different from most studies' usual null hypothesis testing. The technique, other than the null hypothesis approach, explicitly incorporates past findings into the model, by e. g. including an explicit prior distribution of the parameter in question from established studies or meta-analyses instead of using simply a normal distribution assumption (referred to as informative priors). This could be practically achieved by using instrumental variables, just as in the "traditional" IV approach, but explicitly incorporating prior distributions of the factors of a model in question as the instrumental variable. The issue with employing Bayesian methods for the demographic diversity performance nexus is that given methods are computationally complex and intensive to employ, besides the usual challenge of autocorrelation with instrumental variables. In the case of the diversity-performance relationship, it is also hard to connect past distribution patterns with better modeling outcomes, as is the case for inflation expectation modeling (Cross et al. 2024) or development studies in psychology. Simply put, in a rapidly changing environment, with the increasing prevalence of regulatory action in the field and a growing rate of demographic diversity in most firm leadership teams, employing past distribution factors seldom creates a superior model but raises complexity. Nevertheless, Bayesian methods may provide additional findings in future research for certain more stable diversity factors, e.g. age heterogeneity distributions or nationality diversity in boardrooms.

Regarding methods in general international business research, Reeb et al. (2012) describe how, in an ideal setting, randomized experimental research

designs would be the best-suited technique to mitigate endogeneity concerns, but also how this is seldom feasible in studying the corporate world in practice. Despite rarely having the chance for such experiments in real life, the abovementioned methods, like event studies and novel synthetic experimental tools, e.g., difference-in-difference and propensity score matching, are already employed in empirical demographic diversity studies. Another novel method for dealing with endogeneity in business research listed by Reeb et al. (2012) is regression discontinuity design (Lee and Lemieux 2010), focusing on the edge of treatment groups, selected based on certain criteria (e.g. specific employees in a work setting). Regression discontinuity design currently employs checks for those individuals or firms in a sample that just made it into a control group and those who did not get in based on selected thresholds.

As an example of a related field of research, empirical marketing literature also uses more novel techniques, e.g. the Gaussian copula approach, modeling the correlation between the endogenous variable and error term (Park and Gupta 2012), or the employment of latent instrumental variables, essentially splitting an endogenous variable into an exogenous and an endogenous part (Papies et al., 2017), besides generally using the usual instrumental and control variable settings (Hult et al. 2018). In the field dealing with national cultural aspects and their connections to financial performance, instrumental variables are often employed (Nash and Patel 2019). In political science work, instrumental variables are utilized commonly (Bellemare et al. 2017). Overall, other fields of empirical finance literature employ similar methods for endogeneity-checking as demographic diversity studies within our sample, with a heavy emphasis in general econometrics on the employment of adequate control variables and instrumental variable approaches (Hult et al. 2018), besides the occasional novel research design.

## 4 Results and discussion

Based on our extensive sample and as shown in Table 2, we can state that most papers examining the connection between top management team heterogeneity and the financial performance of firms employ methods that account for endogeneity, with age heterogeneity being the factor employing such methods in the fewest cases, with still 62% of studies doing so. This is still considerably different, e.g., 85% of papers examine functional heterogeneity checks for the issue.

Over a third of studies in our sample employ tests for more than one heterogeneity dimension (e.g. Halcro et al. (2021) test for gender, age, and nationality heterogeneity). This suggests that authors might have chosen endogeneity mitigation for one diversity dimension in some cases, only because they had already incorporated it into other dimensions in the same study. Nevertheless, it is clear from our sample that the employment of methods to alleviate endogeneity concerns is not highly dependent on demographic heterogeneity dimensions covered by a study, as most papers do in every case.

As shown in Figs. 3 and 4, based on our sample over time, the share of papers that recognize the endogeneity problem and the share of papers utilizing a method to

**Table 2** Diversity dimensions represented by papers in sample

Diversity dimension	# of papers in large sample	# of papers from dimension method utilizing to check for endogeneity issues	Share of total of those that utilize method to mitigate endogeneity (%)
Gender	139	102	73
Nationality	36	25	69
Age	34	21	62
Education	31	22	71
Ethnicity	20	15	75
Functional	13	11	85
Aggregate	8	7	88
Tenure	13	11	85
Religion	1	1	100

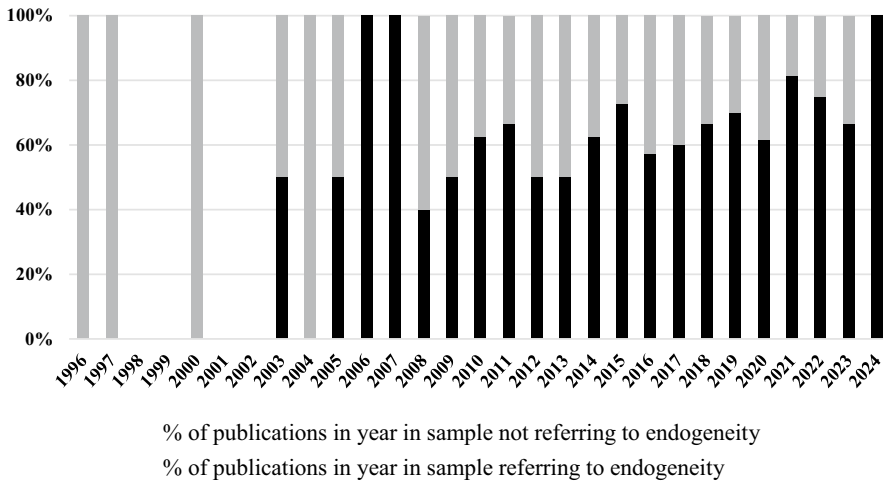
This Table describes the heterogeneity dimensions represented in our study sample of 163 papers and what share of respective dimension utilizes a technique to check for endogeneity concerns

alleviate endogeneity concerns have steadily increased. This reveals the increasing consciousness of authors in empirical leadership heterogeneity literature regarding the problem of endogeneity in the field.

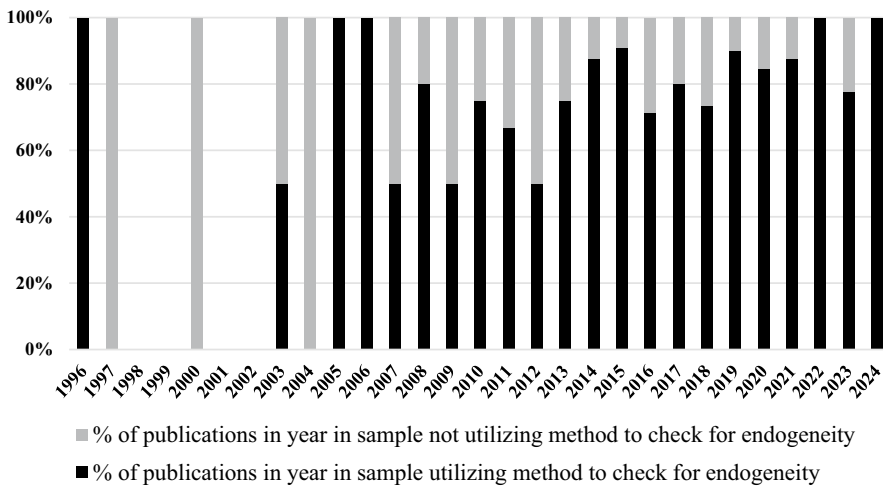
As shown in Table 3, a significant share of papers that check for endogeneity concerns do so by employing more than one method – although this also depends on how these methods are categorized since categories displayed here might overlap.<sup>2</sup> Nevertheless, a lesson learned here is that if authors recognize the problem of endogeneity with the general model, they often pursue multiple solutions to alleviate given concerns, e.g., by utilizing fixed effects in the regression besides even a simple form of lagged variables (e.g., as performed by Chapple and Humphrey 2014).

As shown in Table 4, in our overall sample of 163 studies (including papers not mitigating endogeneity concerns), 75% found a significant and positive connection between leadership heterogeneity and firm financial performance, 30% found a significant and negative connection, while 53% of papers in some instances found no significant connection. A handful of papers also found some form of U-shaped significant relationship between the factors (e.g., Ali et al. 2014; Pucheta-Martínez et al. 2018), not included in the table here. Overall, studies that accounted for endogeneity in some methodical way had a similar finding distribution to the entire sample, which included studies not checking for endogeneity in any methodical way.

<sup>2</sup> E.g. System GMM models always utilize a form of lagged variable, meaning a study employing this method would automatically be accounted for in both, the section System GMM and the section for lagged variables, too. Similarly, as 2SLS models are essentially a special form of linear GMM models, these too include Instrumental Variables, also automatically checking both (2SLS and Instrumental Variable) boxes for a given study. Employing multiple remedies for endogeneity often happens not only by integrating various methodical approaches into one model, but by calculating multiple variations of a general model with different approaches (e.g. for robustness check purposes, or simply to enhance validity of testing).



**Fig. 3** Mentions of endogeneity concerns over time. This figure shows the number of studies mentioning the endogeneity problem regarding the relationship between leadership heterogeneity and firm financial performance for publications in any given year within our sample of 163 papers published from 1996 to 2024



**Fig. 4** Utilization of methods accounting for endogeneity. This Figure shows the share of studies employing at least one method considered to be a check for the endogeneity problem regarding the relationship of leadership heterogeneity and firm financial performance for publications in any given year within our sample of 163 papers published from 1996 to 2024

As shown in Table 5, nearly two-thirds of papers in our overall sample recognize endogeneity problems regarding the subject matter of leadership demographic heterogeneity and firm financial performance. Only seven of these papers recognize the issue and do not employ some kind of methodical remedy to alleviate concerns (e.g.,

**Table 3** Number of techniques utilized by sample studies checking methodically for endogeneity concerns

# of methods utilized	# of papers	% of all papers	% of papers utilizing method to check for endogeneity issues
0	40	25	NA
1	10	6	8%
2	38	23	31%
3	29	18	24%
4	31	19	25%
5	12	7	10%
6	2	1	2%
7	1	1	1%

This Table describes the number and share of techniques utilized by studies checking methodically for endogeneity concerns from our sample of 163 papers. Instrumental variables are counted as a separate type of variable, even in cases where a sub form of IV is counted as another variable (e.g. 2SLS)

Abdullah et al. 2016). In turn, somewhat more papers (75% of our total sample) employ methodical remedies for endogeneity, than the number of studies that recognize the problem. In our further evaluation, we decided to include the papers that don't mention endogeneity but handle it somehow (e.g. Naseem et al. 2017), given that the methods employed help to check for endogeneity, regardless of whether they were explicitly utilized for this reason.

As also shown in Table 4, the methods utilized to account for endogeneity, are heavily concentrated on a handful of general approaches (all described in detail in Chapter 3). Two-thirds of papers that account methodically for endogeneity employ some sort of fixed effects for at least one regression in their research. These time, industry and/or firm fixed effects were often not explicitly utilized as an answer to endogeneity problems, but as a general tool in the given regression models.

Well over half (58%) of studies that utilized a technique to mitigate endogeneity concerns used some form of a lagged variable. This commonly meant including the performance of a given firm (the dependent variable) one or multiple periods after the relevant independent variables in question (firm characteristics, e.g., leadership heterogeneity) to account for the effects of leadership decision-making to arrive at the financial performance level. For System GMM models, the lagged variable means including a lagged performance variable (a lagged version of the dependent variable) as an independent variable in the regression equation. Other cases of lagged variables included, e.g., lagged instrumental variables.

GMM models were used in 31% of studies using methodical approaches to account for endogeneity concerns. Over half of these papers utilized the System GMM methodology, with the rest being simple Difference GMM models. System GMM models were especially widespread in more recent studies (e.g., Arora 2022; Arvanitis et al. 2022; Tejerina-Gaite and Fernández-Temprano 2021), underpinning the statement from Chapter 3 that this method is currently seen as a sophisticated approach to account for endogeneity concerns. Nearly as much, namely, 29%

**Table 4** The share of papers finding a significant positive or negative, or no significant relationship in the overall sample of 163 papers and in the subsample of 123 papers checking for endogeneity concerns

	# of papers in over- all sample	% of papers in over- all sample	# of papers in over- check for endogeneity issues	% of total of those that utilize method to mitigate endogene- ity
Studies finding a significant positive connection	122	75	92	75
Studies finding a significant negative connection	49	30	34	28
Studies finding no significant connection	88	53	58	47

This Table describes the number and share of papers within our sample finding a positive or negative significant, or no significant relationship between leadership heterogeneity and firm financial performance. The Table additionally includes the number and share of papers within our sample checking methodically for endogeneity concerns finding a positive or negative significant, or no significant relationship between leadership heterogeneity and firm financial performance

**Table 5** Share of papers utilizing specific methods to check for endogeneity concerns and study relationship test outcomes

	# of papers	% of papers that utilize a method to check for endogeneity issues in sample	% of papers that found a positive connection	% of papers of papers that found a negative connection	% of papers of papers that found no connection
Share of papers recognizing endogeneity issues	104	93	75	21	48
Share of papers utilizing method to check for endogeneity issues	123	100	75	28	47
Lagged variable(s) used	71	58	76	23	49
Fixed effects	82	67	77	29	45
Random effects	24	20	75	29	42
Instrumental Variable (IV)	55	45	78	25	36
2SLS	33	27	79	18	33
3SLS	2	2	50	0	100
GMM	38	31	82	29	40
System GMM	23	19	87	22	48
Difference (or other) GMM	15	12	73	40	33
Diff-in-diff	4	3	25	50	50
Propensity score matching	8	7	63	25	38
Others	6	5	67	0	67

This Table shows the overall number of papers mentioning endogeneity concerns, the number of papers utilizing a method recognized as a check for endogeneity, and the number of papers utilizing a given mitigating technique. The Table additionally shows the share of papers in each given category finding significant positive, negative, or no significant results. Individual studies in some instances do not reach a given result counted here based on a specific technique. E.g. System GMM could have been resulting in negative significant relationships found, while another technique employed by the paper does not find any significant relationship, also accounted for under System GMM in the table above. Further, due to the only insignificant representation in our sample of 163 studies, we excluded variations of U-shaped relationships from this representation

of research papers in our sample that utilized a method to account for endogeneity, used the 2SLS or 3SLS technique.

Further methods were much less utilized, with random effect models often not employed after utilizing the Hausman test to decide whether to employ fixed or random effects (e.g. Gordini and Rancati 2017). Diff-in-diff methods, propensity score matching, and various other approaches are even more marginal in our sample.

As also displayed in Table 4, the distribution of findings from studies based on different methods employed varied strongly in some cases. From papers that employed some sort of statistical methodical remedy for endogeneity concerns, 75% found at least some statistically positive connection between firm leadership heterogeneity and firm financial performance. Only 28% found a significant negative relationship, and nearly half of papers in at least some instances (regression models run) did not find any significant relationship.<sup>3</sup> While 87% of research papers utilizing System GMM models found positive significant results, only one out of the four studies in our sample employing diff-in-diff found a significantly positive connection between some leadership demographic heterogeneity factors and firm financial performance. Negative significant relationships found between a heterogeneity factor and firm financial performance were more equivalently represented throughout methods employed, with, e.g., papers utilizing fixed effects in models finding negative relationships in 29% of these studies. Only a third of studies employing the 2SLS method found at least in some cases no relationship, while close to half of System GMM employing papers in some regression instances did not reach conclusive results.

A first simplified conclusion from this is that the methods employed to check for endogeneity may influence outcomes, with more sophisticated statistical tools (e.g., System GMM) finding more often positive and significant relationships than various other tools to check for the issue.<sup>4</sup> It is clear from our data that methods that are supposed to account for endogeneity in research might have an influence on model outcomes for the leadership heterogeneity and firm financial performance nexus, which is a key hypothesis to be validated by future more detailed meta-research.

As shown in Fig. 5,<sup>5</sup> we looked deeper into our sample studies and accounted for individual outcomes of papers on a heterogeneity dimension level for further insights. The share of papers in our large sample, finding a significant positive

<sup>3</sup> A study can have simultaneous findings of positive- and negative significant relationships, and no significant relationships found, due to more than one regression run, and/or multiple demographic heterogeneity variables tested separately in the same paper (e.g. Fraga and Silva 2012; Lückcrath-Rovers 2013; Tarigan et al. 2018; Ararat et al. 2010).

<sup>4</sup> This simplified view is somewhat dampened by the fact that our data evaluation is only respective paper-based regarding the relationship outcomes, meaning it is theoretically possible, that e.g. regression models finding a positive and significant connection for a relationship are not performed by utilizing a method from our categories above, but are e.g. outcomes from a simple regression also run in the same study.

<sup>5</sup> We excluded the demographic heterogeneity factors of Religion and Tenure, as well as the Aggregate demographic heterogeneity factor for this evaluation, as these all had a very small sample sizes and hence could not be analyzed thoroughly here.

or negative relationship, or no significant relationship in some instances, varies strongly based on the demographic heterogeneity dimension tested.<sup>6</sup> Between 42 (Education) AND 58% (Gender) of papers of a given dimension found at least some positive and significant relationship between heterogeneity in leadership and firm performance. The representation of negative significant findings was between 5 (Ethnicity) and 32% (Education), depending on dimension studies, while studies that found at least in some instances no significant relationship were represented between 40 (Ethnicity) and 53% (Age). This shows that demographic heterogeneity effects will depend on what dimensions are considered. The variation also does not correspond to whether a heterogeneity factor is categorized as readily observable or non-observable since, e.g., there is no clear order of non-observable factors (nationality, education, functional) dominating the lows or highs of positive or negative findings' representation.

Regarding the delta in findings between the overall sample (including papers not mitigating the endogeneity problem with any methodical tools) and our subsample of studies employing endogeneity mitigating techniques, this also varied depending on the given dimension of heterogeneity tested. The largest delta for positive and significant relationship findings was the Nationality heterogeneity dimension, with a 6% higher share of papers finding positive connections in the subsample checking for endogeneity. The direction of this delta, however, was not always positive; e.g., the factors of age (-6%) and ethnicity (-5%) had lower shares of positive findings than in our overall sample when only considering endogeneity mitigating papers. For negative significant findings, the outlier was the Education dimension, where 10% fewer papers (a still comparatively high 23%) found a negative significant relationship between heterogeneity in leadership and firm financial performance.

Endogeneity mitigating techniques do not systematically produce more significant positive or negative findings in papers. However, they might somewhat reduce the chances of studies, at least for some tests run, finding no significant relationship between leadership heterogeneity and firm financial performance. In the overall sample and all dimensions examined here, the subsample, including endogeneity mitigating techniques, had a lower share of papers finding, in some cases, no connection.

As shown in Table 6, the most cross-cited papers within our endogeneity mitigating subsample are focused on firms from the United States, with the frontrunners of European countries introducing quota laws (e.g., Norway) also prominently represented. Although the most cross-cited study in our sample (Adams and Ferreira 2009) finds a negative and significant connection between leadership heterogeneity (focusing on the gender dimension) and firm financial performance, most top-cited work within our subsample of papers checking for endogeneity finds positive significant relationships. Regarding the dimensions of demographic heterogeneity covered,

---

<sup>6</sup> We recognize that a comparison between a "pure" sample not including any papers considering endogeneity and our subsample of studies methodically checking for it would have been more prudent. However, as only a limited number of studies does not employ any method or technique to check for endogeneity, this comparison would have been significantly less representative.

all of the mainly cross-cited studies above at least cover the gender heterogeneity dimension, with ethnicity heterogeneity also tested in some instances. This represents the overall distribution of studies well, given that most studies in our sample focus on the demographic heterogeneity dimension of gender (102 out of 123 papers in our subsample of papers checking for endogeneity), with other factors represented and tested in far fewer instances. The following most tested dimensions, nationality, age, and education heterogeneity, were merely tested 36, 34, and 31 times within our subsample.

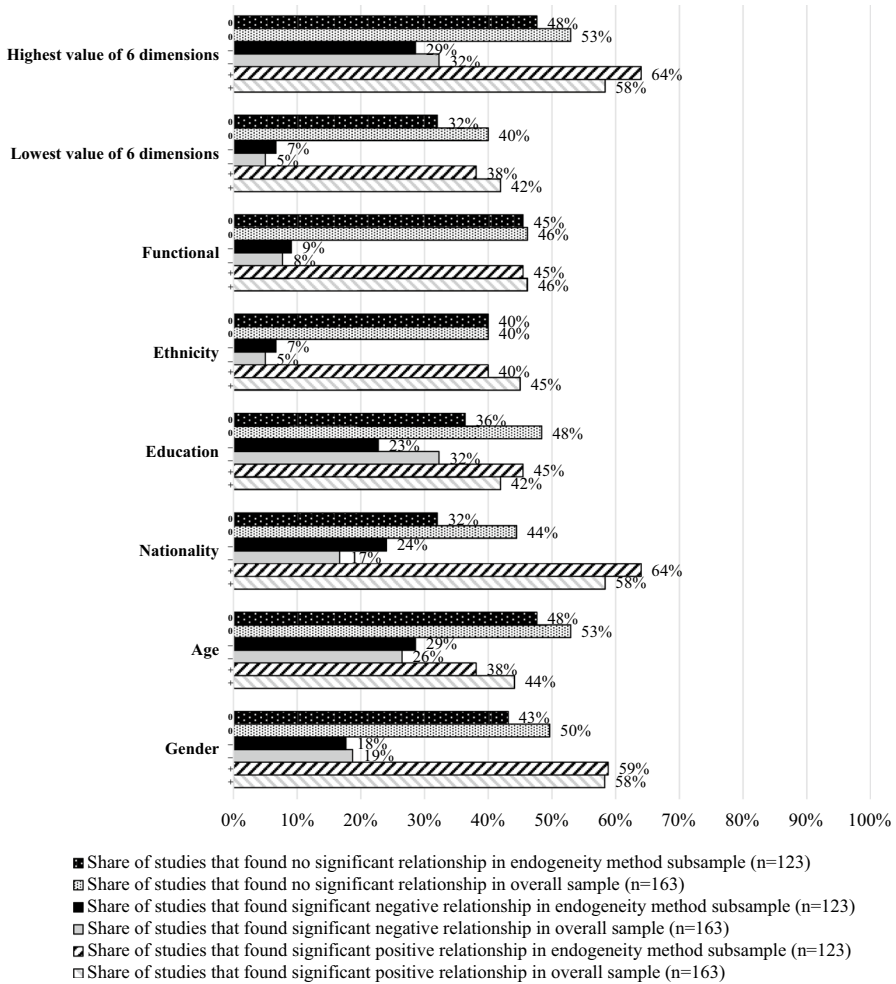
## 5 Conclusion

Based on our sample of empirical studies examining the connection between companies' leadership demographic heterogeneity and financial performance, we can state that a broad share of studies recognizes endogeneity problems and attempts to mitigate these by employing various methodical approaches. We also showed some variability regarding the recognition of endogeneity, but empirical papers covering various dimensions of heterogeneity are overwhelmingly checking for this issue. The demographic diversity field least often checking for endogeneity is age heterogeneity related empirical research, with still 62% of papers employing checks for the issue. Further, the share of papers mentioning and checking for this issue is growing, showing the rising awareness of endogeneity problems in the empirical literature.

Three-quarters of papers in our sample of 163 studies found positive relationships between some respective leadership heterogeneity dimension and financial performance – with the caveat that some of these studies also found negative significant or no significant relationships (30% and 53% respectively in our sample). This ambiguity was possible because a single research paper might have found a specific relationship for one heterogeneity dimension while not finding a relationship for another or due to multiple types of regression analyses performed for the very same variables with differing outcomes. The findings within our subsample of papers that controlled for endogeneity did not materially differ from these overall numbers.

Nevertheless, the relationship findings for individual heterogeneity dimensions differed in some instances between the sample, including papers not checking for endogeneity and the subsample concentrating on endogeneity mitigating studies only. However, there is no clear direction on how endogeneity mitigating techniques influence performance outcomes in a significant positive or negative way between various diversity dimensions.

For papers tackling the problem of endogeneity, an overwhelming majority utilize more than one method to check for this issue – the most prevalent methods are lagged variables and fixed effects. Somewhat less prominently represented but methodically deeper are IV techniques, particularly 2SLS and GMM regressions. The outcomes from respective methods differ, with papers employing the system GMM technique finding significant positive relationships in 87% of studies utilizing this method, with the same share being only 76% for papers utilizing lagged variables. This aligns with our larger sample, including studies that do not check for endogeneity. Similarly, only 33% of studies utilizing the 2SLS technique find, at



**Fig. 5** Demographic heterogeneity in top management teams effects on financial performance by relationship outcome and endogeneity checks employed. This figure shows relationship outcomes by papers in our sample of 163 papers and our subsample of 123 papers, which check methodically for endogeneity concerns. Studies can find significant positive or negative, or no significant relationship, or a combination of these categories within one paper and respective demographic heterogeneity dimension. The demographic heterogeneity dimensions of tenure and religion, as well as studies utilizing an aggregated form of a heterogeneity measure combining multiple heterogeneity dimensions into one factor, were excluded from this Figure due to the only insignificant representation in our sample. Further, due to the only insignificant representation in our sample of 163 studies, we excluded variations of U-shaped relationships from this representation

least in some cases, no significant relationships, while the same statistic for papers using lagged variables is 49% (more aligned with our overall sample). Thus, the detailed effects of each method utilized are a highly relevant direction for further research, as identified by our analysis. The relevance of empirical outcomes is predominantly legitimate, considering that endogeneity can arguably be managed by

**Table 6** Most cross-cited research from within study subsample checking for endogeneity concerns methodically

Title	# Quoted by	Country	Time-frame	Fin-dings	Endogeneity check methods employed	Dimensions covered	Publication
Adams and Ferreira (2009) – Women in the boardroom and their impact on governance and performance	81	United States	1996–2003	–	Fixed Effects, Lagged Variables, Difference GMM, IV	Gender	Journal of Financial Economics
Carter et al. (2003) – Corporate Governance, Board Diversity, and Firm Value	69	United States	1999	+	2SLS, IV	Gender, Ethnicity/Race	The Financial Review
Campbell and Minguez-Vera (2008) – Gender Diversity in the Boardroom and Firm Financial Performance	48	Spain	1995–2000	+	2SLS, IV, Fixed Effects	Gender	Journal of Business Ethics
Ahern and Dittmar (2011) – The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation	48	Norway	2001–2009	–	Fixed Effects, IV	Gender	Quarterly Journal of Economics
Carter et al. (2010) – The Gender and Ethnic Diversity of US Boards and Board Committees and Firm Financial Performance	42	United States	1998–2002	0	3SLS, IV, Fixed Effects, Lagged Variables	Gender, Ethnicity/Race	Corporate Governance
Farrell and Hersch (2005) – Additions to corporate boards: the effect of gender	39	United States	1990–1999	+	Fixed Effects	Gender	Journal of Corporate Finance

Table 6 (continued)

Title	# Quoted by	Country	Time-frame	Fin-dings	Endogeneity check methods employed	Dimensions covered	Publication
Smith et al. (2006) – Do women in top management affect firm performance? A panel study of 2,500 Danish firms	32	Denmark	1993–2001	+ / 0	Fixed Effects, IV	Gender, Education	International Journal of Productivity and Performance Management
Miller and Del Carmen Triana (2009) – Demographic Diversity in the Boardroom: Mediators of the Board Diversity-Firm Performance Relationship	27	United States	2002–2005	+ / 0	Lagged Variables	Gender, Ethnicity/Race	Journal of Management Studies
Liu et al. (2014) – Do women directors improve firm performance in China?	26	China	1999–2011	+	2SLS, IV, Fixed Effects, Lagged Variables, System GMM	Gender	Journal of Corporate Finance
Terjesen et al. (2015) – Does the presence of independent female directors impact firm performance? A multi-country study of board diversity	23	Global	2010	+	Lagged Variables, Difference GMM	Gender	Journal of Management & Governance

This table shows the most cross-referenced empirical research within our subsample of a total of 123 papers employing a method to check for endogeneity concerns. We hereby only strictly accounted for references coming from other papers within this subsample. Besides author(s) and title, we included the number of references coming from other paper within the sample (# Quoted by), the origin of the data set (# country) of respective empirical study, the observation period of the respective empirical study (Timeframe), and the techniques employed and considered to be a check for endogeneity concerns (Endogeneity check methods employed). We also show the findings regarding the relationship between tested diversity dimensions and financial performance measures (Findings), “+” depicting a negative significant relationship, “+” depicting a positive significant relationship, and “0” accounting for no significant relationship found. Studies with more than one depiction (e.g. “+ / 0”) have multiple test runs with different relationship outcomes. Also covered are the respective publications (Publications) and diversity dimensions which are addressed by the authors (Dimensions covered)

combining multiple advanced testing methods, e.g., by employing System GMM in empirical studies.

Endogeneity is one of the major critical factors of demographic diversity research and deserves the methodical attention it has received in recent years in empirical studies. Our overall literature review of demographic diversity research of the past decades shows that scientific results of the entire demographic diversity field should not be disqualified due to existing presumptions regarding reverse causality or omitted variables. Endogeneity is a potential conflict but not an effect bound to happen. Due to the high practical implications of demographic diversity research outcomes for multiple stakeholder groups, the tradeoff between rigor and relevance tilts towards the findings' relevance factor. Investors, policymakers, and other parties involved with corporate leadership teams have a vast interest in advancing the understanding of leadership team dynamics and their effects on performance factors.

In future studies, the employment of novel statistical tools and econometric techniques, e.g., Bayesian methods or the seldom used diff-in-diff technique, should be considered by empirical scholars to advance the existing findings of demographic diversity effects on performance factors. This could especially raise the quality of more stable demographic factors in firm leadership teams with more consistent and well-documented prior distribution attributes. Additionally, Bayesian models in the future may better account for endogeneity and variability in demographic diversity effects on performance by industry, geography, or other parameters. Novel statistical tools overall could support stakeholders in better comprehending the interaction between different diversity measures (e.g., joint effects of multiple heterogeneity factors like gender, age, and nationality) while mitigating endogeneity problems. Advanced computational techniques based on even never pioneering approaches, e.g., machine learning in combination with traditional empirical modeling, could help identify causal relationships more efficiently in the presence of endogeneity.

Still, more calculation techniques and data may never solve a statistical problem if it has a real-life fundamental, such as the case of endogeneity. All rigorous scientific methods used to account for endogeneity can have flaws, e.g., it is practically hard to find suiting instrumental variables without autocorrelation issues. Statistical methods discussed here may have their issues and should not be considered an ultimate remedy. As Myles Shaver (2020) describes, it is essential to regard endogeneity as a comprehensive problem, where a part of the solution incorporates rival theories to our own in research designs, and another part employs a broad selection of mitigating techniques simultaneously. For the demographic diversity field, especially relevant in the context of possible endogeneity is the theoretical framework that guides hypothesis development and, in the aggregate, suggests a positive relationship between decision-making outcomes and decision-maker group heterogeneity.

Our paper provides important implications on how the relevance of endogeneity problems in empirical social and economic research can be adequately considered and how negative or distorting effects in empirical evaluation can be reduced. Future empirical research on the relationship between demographic diversity and company performance metrics should consider certain methodological safeguards

to alleviate endogeneity concerns as a major criticism of the field. As a first step, hypothesis development in diversity research should always check for current decision-outcome-related theories utilized in the broader field. Regarding statistical techniques, it is crucial to consider endogeneity as more than a one-trick problem easily solved by any specific quantitative method or tool. Instead, future research should employ several of the techniques introduced above, emphasizing robustness checks and multiple versions of variables utilized regarding dependent and independent variables (e.g., ROA, Tobin's  $q$ , and share price development, all utilized in separate tests as dependent variables). Statistical designs should operate with a lagged dependent variable and ideally test by utilizing both fixed effects (after an appropriate Hausmann-test) and sophisticated instrumental variable regression models (e.g., system GMM) to validate the developed hypotheses. If feasible, based on the data available, quasi-experimental research designs, e.g., diff-in-diff or propensity score matching, are advised to be added to the process. Following the above steps should provide researchers with an appropriately delicate statistical answer to existing concerns regarding endogeneity in demographic diversity research and, in total, enrich the field of demographic diversity research with sound and highly relevant empirical outcomes.

**Funding** Open Access funding enabled and organized by Projekt DEAL. No funding was received.

**Data availability** The data supporting this study's findings are available from the corresponding author upon reasonable request.

## Declarations

**Conflict of interest** None; there are no conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- \*Abdullah SN, Ismail KNIK, Nachum L (2016) Does having women on boards create value? The impact of societal perceptions and corporate governance in emerging markets. *Strategic Manag J* 37(3):466–476. <https://doi.org/10.1002/smj.2352>
- \*Adams RB, Ferreira D (2009) Women in the boardroom and their impact on governance and performance. *J Financ Econ* 94(2):291–309. <https://doi.org/10.1016/j.jfineco.2008.10.007>
- \*Ahern KR, Dittmar AK (2011) The Changing of the boards: the impact on firm valuation of mandated female board representation. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.1364470>

- Akram F, Abrar ul Haq M, Natarajan VK, Chellakan RS (2020) Board heterogeneity and corporate performance: an insight beyond agency issues. *Cogent Bus & Manag* 7(1):1809299. <https://doi.org/10.1080/23311975.2020.1809299>
- \*Ali M, Ng YL, Kulik CT (2014) Board age and gender diversity: a test of competing linear and curvilinear predictions. *J Bus Ethics* 125(3):497–512. <https://doi.org/10.1007/s10551-013-1930-9>
- \*Ararat M, Aksu MH, Cetin AT (2010) The impact of board diversity on boards monitoring intensity and firm performance: evidence from the istanbul stock exchange. *SSRN Electro J*. <https://doi.org/10.2139/ssrn.1572283>
- Arellano M, Bond S (1991) Some tests of specification for panel data: monte carlo evidence and an application to employment equations. *Rev Econ Stud* 58(2):277. <https://doi.org/10.2307/2297968>
- Arellano M, Bover O (1995) Another look at the instrumental variable estimation of error-components models. *J Econom* 68(1):29–51. [https://doi.org/10.1016/0304-4076\(94\)01642-D](https://doi.org/10.1016/0304-4076(94)01642-D)
- \*Arioglu E (2020) Female board members: the effect of director affiliation. *Gender Manag: Int J* 35(2):225–254. <https://doi.org/10.1108/GM-05-2019-0080>
- \*Arora A (2022) Gender diversity in boardroom and its impact on firm performance. *J Manag Gov* 26(3):735–755. <https://doi.org/10.1007/s10997-021-09573-x>
- \*Arvanitis SE, Varouchas EG, Agiomirgianakis GM (2022) Does board gender diversity really improve firm performance? evidence from greek listed firms. *J Risk Financ Manag* 15(7):306. <https://doi.org/10.3390/jrfm15070306>
- Baker HK, Pandey N, Kumar S, Haldar A (2020) A bibliometric analysis of board diversity: current status, development, and future research directions. *J Bus Res* 108:232–246. <https://doi.org/10.1016/j.jbusres.2019.11.025>
- \*Bayly N, Breunig RV, Wokker C (2023) Female board representation and corporate performance: a review and new estimates for Australia. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.4639529>
- Bellemare MF, Masaki T, Pepinsky TB (2017) Lagged explanatory variables and the estimation of causal effect. *J Polit* 79(3):949–963. <https://doi.org/10.1086/690946>
- \*Bennouri M, Chtioui T, Nagati H, Nekhili M (2018) Female board directorship and firm performance: what really matters? *J Bank Financ* 88:267–291. <https://doi.org/10.1016/j.jbankfin.2017.12.010>
- Berle AA, Means GC (1932) *The Modern Corporation and Private Property*. Macmillan, New York
- Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. *J Econom* 87(1):115–143. [https://doi.org/10.1016/S0304-4076\(98\)00009-8](https://doi.org/10.1016/S0304-4076(98)00009-8)
- Campbell K, Mínguez-Vera A (2008) Gender diversity in the boardroom and firm financial performance. *J Bus Ethics* 83(3):435–451. <https://doi.org/10.1007/s10551-007-9630-y>
- \*Cannella AA, Park J-H, Lee H-U (2008) Top management team functional background diversity and firm performance: examining the roles of team member colocation and environmental uncertainty. *Acad Manag J* 51(4):768–784. <https://doi.org/10.5465/amr.2008.33665310>
- \*Carter DA, Simkins BJ, Gary Simpson W (2003) Corporate governance, board diversity, and firm value. *Financ Rev* 38(1):33–53. <https://doi.org/10.1111/1540-6288.00034>
- \*Carter DA, D'Souza F, Simkins BJ, Simpson WG (2010) The gender and ethnic diversity of US boards and board committees and firm financial performance. *Corp Gov: Int Rev* 18(5):396–414. <https://doi.org/10.1111/j.1467-8683.2010.00809.x>
- \*Chapple L, Humphrey JE (2014) Does board gender diversity have a financial impact? evidence using stock portfolio performance. *J Bus Ethics* 122(4):709–723. <https://doi.org/10.1007/s10551-013-1785-0>
- \*Chijoke-Mgbame AM, Boateng A, Mgbame CO (2020) Board gender diversity, audit committee and financial performance: evidence from Nigeria. *Account Forum* 44(3):262–286. <https://doi.org/10.1080/01559982.2020.1766280>
- Cross JL, Hoogerheide L, Labonne P, van Dijk HK (2024) Bayesian mode inference for discrete distributions in economics and finance. *Econ Lett* 235:111579. <https://doi.org/10.1016/j.econlet.2024.111579>
- \*Đặng R, Houanti L, Reddy K, Simioni M (2020) Does board gender diversity influence firm profitability? a control function approach. *Econ Model* 90:168–181. <https://doi.org/10.1016/j.econmod.2020.05.009>
- \*Dezsö CL, Ross DG (2012) Does female representation in top management improve firm performance? a panel data investigation. *Strategic Manag J* 33(9):1072–1089. <https://doi.org/10.1002/smj.1955>
- Farrell KA, Hersch PL (2005) Additions to corporate boards: The effect of gender. *J CorpFinan* 11(1–2):85–106. <https://doi.org/10.1016/j.jcorpfin.2003.12.001>

- Fama EF, Jensen MC (1983) Separation of ownership and control. *J Law Econ* 26(2):301–325. <https://doi.org/10.1086/467037>
- \*Fraga J, Silva V (2012) Board diversity and firm performance: an empirical investigation in the Brazilian market. *Braz Bus Rev sp.ed(1)*:55–77. <https://doi.org/10.15728/bbrconf.2012.3>
- Freeman RE (1984) *Strategic management: a stakeholder approach*. Pitman, Boston
- \*Gordini N, Rancati E (2017) Gender diversity in the Italian boardroom and firm financial performance. *Manag Res Rev* 40(1):75–94. <https://doi.org/10.1108/MRR-02-2016-0039>
- \*Gruszczynski M (2020) Women on boards and firm performance: a microeconomic search for a connection. *J Risk Financ Manag* 13(9):218. <https://doi.org/10.3390/jrfm13090218>
- \*Halcro K, Noamene TB, Chaheer D, Talib A (2021) Is there a business case for board diversity? An investigation of UK's FTSE 100 companies. *Human Syst Manag* 40(3):421–434. <https://doi.org/10.3233/HSM-190849>
- Hambrick DC, Mason PA (1984) Upper Echelons: the organization as a reflection of its top managers. *Acad Manag Rev* 9(2):193. <https://doi.org/10.2307/258434>
- Hermalin B, Weisbach M (2003) *Boards of Directors as an Endogenously Determined Institution: A Survey of the Economic Literature*. Cambridge
- Hermalin BE, Weisbach MS (1998) Endogenously chosen boards of directors and their monitoring of the CEO. *Am Econ Rev* 88(1):96–118
- Hesarzadeh R (2020) Difference-in-differences design and propensity score matching in top accounting research: a short guide for Ph.D. students in Iran. *Iran J Account, Audit & Financ* 4(3):35–47
- Hill AD, Johnson SG, Greco LM, O'Boyle EH, Walter SL (2021) Endogeneity: a review and agenda for the methodology-practice divide affecting micro and macro research. *J Manag* 47(1):105–143. <https://doi.org/10.1177/0149206320960533>
- Hillman AJ, Withers MC, Collins BJ (2009) Resource dependence theory: a review. *J Manag* 35(6):1404–1427. <https://doi.org/10.1177/0149206309343469>
- Jensen MC, Meckling WH (1976) Theory of the firm: managerial behavior, agency costs and ownership structure. *J Financ Econ* 3(4):305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)
- Kagzi M, Guha M (2018) Board demographic diversity: a review of literature. *J Strategy Manag* 11(1):33–51. <https://doi.org/10.1108/JSMA-01-2017-0002>
- Kanter RM, Bluedorn A (1977) *Men and Women of the Corporation*. Basic Books, New York
- Konrad AM, Kramer V, Erkut S (2008) Critical mass: the impact of three or more women on corporate boards. *Organ dyn* 37(2):145–164
- Lee DS, Lemieux T (2010) Regression discontinuity designs in economics. *J Econ Lit* 48(2):281–355. <https://doi.org/10.1257/jel.48.2.281>
- \*Liu Y, Wei Z, Xie F (2014) Do women directors improve firm performance in China? *J Corp Finance* 28:169–184. <https://doi.org/10.1016/j.jcorpfin.2013.11.016>
- \*Lückerath-Rovers M (2013) Women on boards and firm performance. *J Manag & Gov* 17(2):491–509. <https://doi.org/10.1007/s10997-011-9186-1>
- \*Mastella M, Vancin D, Perlin M, Kirch G (2021) Board gender diversity: performance and risk of Brazilian firms. *Gender Manag: Int J* 36(4):498–518. <https://doi.org/10.1108/GM-06-2019-0088>
- \*Masulis RW, Wang C, Xie F (2012) Globalizing the boardroom—The effects of foreign directors on corporate governance and firm performance. *J Account Econ* 53(3):527–554. <https://doi.org/10.1016/j.jacceco.2011.12.003>
- \*Miletkov MK, Poulsen AB, Wintoki M, Babajide (2012) *Importing governance: a multinational study of the determinants and effects of having foreign independent directors*. European finance association annual meeting, Copenhagen, Denmark, 2012.
- \*Miller T, del Carmen Triana M (2009) Demographic diversity in the boardroom: mediators of the board diversity–firm performance relationship. *J Manag stud* 46(5):755–786
- Milliken FJ, Martins LL (1996) Searching for common threads: understanding the multiple effects of diversity in organizational groups. *Acad Manag Rev* 21(2):402. <https://doi.org/10.2307/258667>
- Myles Shaver J (2020) Endogeneity in International Business Research: A Commentary. In: Eden L, Nielsen BB, Verbeke A (eds) *Research Methods in International Business*. Springer, Cham, pp 377–382. [https://doi.org/10.1007/978-3-030-22113-3\\_19](https://doi.org/10.1007/978-3-030-22113-3_19)
- \*Naseem MA, Xiaoming S, Riaz S, Rehman RU (2017) Board attributes and financial performance: the evidence from an emerging economy. *J Dev Areas* 51(3):281–297. <https://doi.org/10.1353/jda.2017.0073>
- Nash R, Patel A (2019) Instrumental variables analysis and the role of national culture in corporate finance. *Financ Manag* 48(2):385–416. <https://doi.org/10.1111/fima.12248>

- \*Nguyen T, Locke S, Reddy K (2015) Does boardroom gender diversity matter? Evidence from a transitional economy. *Int Rev Econ Financ* 37:184–202. <https://doi.org/10.1016/j.iref.2014.11.022>
- \*Niikura H, Seko M (2020) The effect of inside and outside female directors on firm performance: comparison of the first section, second section, mothers, and Jasdaj in the Tokyo stock exchange market. *Int J Econ Policy Stud* 14(1):123–166. <https://doi.org/10.1007/s42495-019-00025-x>
- \*Oldford E, Ullah S, Hossain AT (2021) A social capital view of women on boards and their impact on firm performance. *Manag Financ* 47(4):570–592. <https://doi.org/10.1108/MF-02-2020-0091>
- Papies D, Ebbes P, Van Heerde HJ (2017) Addressing Endogeneity in Marketing Models. In: Leefflang PSH, Wieringa JpE, Bijmolt THA, Pauwels KH (eds) *Advanced Methods for Modeling Markets*. Springer, Cham, pp 581–627. [https://doi.org/10.1007/978-3-319-53469-5\\_18](https://doi.org/10.1007/978-3-319-53469-5_18)
- Park S, Gupta S (2012) Handling Endogenous regressors by joint estimation using copulas. *Market Sci* 31(4):567–586. <https://doi.org/10.1287/mksc.1120.0718>
- Pfeffer J, Salancik GR (1978) *The external control of organizations. A resource dependence perspective*. New York: Harper & Row.
- Pickering C, Byrne J (2014) The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Educ Res & Dev* 33(3):534–548. <https://doi.org/10.1080/07294360.2013.841651>
- Post C, Byron K (2015) Women on boards and firm financial performance: a meta-analysis. *Acad Manag J* 58(5):1546–1571. <https://doi.org/10.5465/amj.2013.0319>
- \*Pucheta-Martínez MC, Bel-Oms I, Olcina-Sempere G (2018) Female institutional directors on boards and firm value. *J Bus Ethics* 152(2):343–363. <https://doi.org/10.1007/s10551-016-3265-9>
- Reeb D, Sakakibara M, Mahmood IP (2012) From the editors: endogeneity in international business research. *J Int Bus Stud* 43(3):211–218. <https://doi.org/10.1057/jibs.2011.60>
- Reinwald M, Hüttermann H, Kröll J, Boerner S (2015) Gender diversity in Führungsteams und Unternehmensperformanz: Eine Metaanalyse. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung* 67(3):262–296. <https://doi.org/10.1007/BF03372925>
- Roodman D (2009) How to do Xtabond2: an introduction to difference and system GMM in Stata. *Stata J: Promot commun stat Stata* 9(1):86–136. <https://doi.org/10.1177/1536867X0900900106>
- \*Rose C, Munch-Madsen P, Funch M (2013) Does board diversity really matter? gender does not, but citizenship does. *Int J Bus Sci Appl Manag* 8(1):16–27. <https://doi.org/10.69864/ijbsam.8-1-96>
- \*Sani A (2021) Board diversity and financial performance of the nigerian listed firms: a dynamic panel analysis. *JABE (J Account Bus Edu)* 6(1):1. <https://doi.org/10.26675/jabe.v6i1.18817>
- Selznick P (1957) *Leadership in administration*. Harper and Row, New York
- \*Simionescu LN, Gherghina ŞC, Tawil H, Sheikha Z (2021) Does board gender diversity affect firm performance? Empirical evidence from standard & Poor's 500 information technology sector. *Financ Innov* 7:1–45
- \*Smith N, Smith V, Verner M (2006) Do women in top management affect firm performance? A panel study of 2,500 Danish firms. *Int J Prod Perform Manag* 55(7):569–593. <https://doi.org/10.1108/17410400610702160>
- Tajfel H, Turner JC (1986) *The Social Identity Theory of Intergroup Behavior*. In: Worchel S and Austin WG (eds.): *Psychology of Intergroup Relations*. Chicago
- \*Talavera O, Yin S, Zhang M (2018) Age diversity, directors personal values, and bank performance. *Int Rev Financ Anal* 55:60–79. <https://doi.org/10.1016/j.irfa.2017.10.007>
- \*Tarigan J, Hervindra C, Hatane SE (2018) Does board diversity influence financial performance? *Int Res J Bus Stud* 11(3):195–215. <https://doi.org/10.21632/irjbs.11.3.195-215>
- \*Tejerina-Gaite FA, Fernández-Temprano MA (2021) The influence of board experience on firm performance: does the director's role matter? *J Manag Gov* 25(3):685–705. <https://doi.org/10.1007/s10997-020-09520-2>
- \*Terjesen S, Couto EB, Francisco PM (2015) Does the presence of independent and female directors impact firm performance? A multi-country study of board diversity. *J Manag Gov* 20:447–483. <https://doi.org/10.1007/s10997-014-9307-8>
- Tomas G, Hult M, Hair JF, Proksch D, Sarstedt M, Pinkwart A, Ringle CM (2018) Addressing endogeneity in international marketing applications of partial least squares structural equation modeling. *J Int Market* 26(3):1–21. <https://doi.org/10.1509/jim.17.0151>
- \*Unite AA, Sullivan MJ, Shi AA (2019) Board diversity and performance of Philippine firms: do women matter? *Int Adv Econ Res* 25(1):65–78. <https://doi.org/10.1007/s11294-018-09718-z>

- van de Schoot R, Kaplan D, Denissen J, Asendorpf JB, Neyer FJ, van Aken MAG (2014) A gentle introduction to bayesian analysis: applications to developmental research. *Child Dev* 85(3):842–860. <https://doi.org/10.1111/cdev.12169>
- Velte P (2017) Do women on board of directors have an impact on corporate governance quality and firm performance? A literature review. *Int J Sustain Strateg Manag* 5(4):302. <https://doi.org/10.1504/IJSSM.2017.089125>
- \*Von Meyerinck F, Niessen-Ruenzi A, Schmid M, & Davidoff Solomon S (2021). As California goes, so goes the nation? The impact of board gender quotas on firm performance and the director labor market. *European Corporate Governance Institute–Finance Working Paper Series*. 785, pp1–99.
- Westphal JD, Milton LP (2000) How experience and network ties affect the influence of demographic minorities on corporate boards. *Adm Sci Q* 45(2):366–398. <https://doi.org/10.2307/2667075>
- Wooldridge JM. (2010) *Econometric analysis of cross section and panel data*. 2. ed. Cambridge, Mass.: MIT Press. Online: <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10453042>.
- \* denotes 45 out of the 163 papers that are part of the systematic literature review sample and are cited in the text.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Daniel Tubik<sup>1</sup>  · Tim Alexander Herberger<sup>1,2,3,4</sup> 

✉ Tim Alexander Herberger  
tim.herberger@andrassyuni.hu; tim-alexander.herberger@uni-bamberg.de

Daniel Tubik  
daniel.tubik@andrassyuni.hu

- <sup>1</sup> Andrassy University Budapest (AUB), Budapest, Hungary
- <sup>2</sup> Department of Finance, Bamberg University, Bamberg, Germany
- <sup>3</sup> University of Erlangen Nuremberg (FAU), Erlangen, Germany
- <sup>4</sup> Salzburg University of Applied Sciences, Salzburg, Austria