Author(s): Rönkkö, Mikko; Maheshwaree, Pardeep; Schmidt, Jens

Title: The CEO effect and performance variation over time

Year: 2023

Version: Published version

Please cite the original version:

Rönkkö, M., Maheshwaree, P., & Schmidt, J. (2023). The CEO effect and performance variation over time. Leadership Quarterly, 34(5), Article 101733. https://doi.org/10.1016/j.leaqua.2023.101733

The Leadership Quarterly 34 (2023) 101733

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

Full length article

The CEO effect and performance variation over time☆

Mikko Rönkkö [a,*], Pardeep Maheshwaree [b], Jens Schmidt [b]

[a] University of Jyväskylä, Jyväskylä University School of Business and Economics, P.O. Box 35, FI-40014 Jyväskylä, Finland
[b] Aalto University, Department of Industrial Management and Engineering, Institute of Strategy and Venturing, P.O. Box 11000, FI-00076 Aalto, Finland

ARTICLE INFO

Keywords:
Autocorrelation
CEO effect
CEOs
Multilevel modeling
Performance trends

ABSTRACT

While CEO effect scholars agree that variation in firm performance tends to persist over time and that CEOs' performance contribution should be gauged against a changing context, recent CEO effect studies addressing these issues have made extreme but opposite claims concerning the magnitude of the CEO effect. We show why recent findings that indicate a much larger CEO effect are spurious. We also argue and empirically demonstrate that a multilevel model that includes autocorrelation can properly gauge CEOs' performance contribution against a changing context while simultaneously avoiding confounding. Our empirical result shows that the opposite claim positing that the CEO effect is nearly indistinguishable from chance is likewise unwarranted. These findings have implications for CEO effect studies and other studies that use longitudinal firm performance data.

Introduction

How important are CEOs to the performance of the firms they lead? To answer this question, there is a long tradition in which variance decomposition studies have been conducted to compare the variation in firm performance among the tenures of different CEOs. These studies have found that typically approximately 15–20% of a firm's performance variance can be attributed to the tenures of its different CEOs (Hambrick & Quigley, 2014; Quigley & Graffin, 2017). Recently, however, two studies have challenged this consensus. At one extreme, by comparing the CEO against an industry- and firm-level context that changes over time, Hambrick and Quigley (2014) found that the CEO effect was as high as 38.5%. At the other extreme, Fitza (2014, 2017) argued that due to methodological problems, the CEO effect of 15–20% that had commonly been found in previous studies is nevertheless indistinguishable from chance. That these studies reach such dramatically different conclusions is paradoxical as they do so despite starting from the same premise; both studies argue that the CEO effect should not be estimated by simply comparing the levels of performance under different CEOs against the average performance of the firm, however, the fact that performance differences (firm-level context) persist over time must also be accounted for.

Resolving this apparent paradox is important because CEO effect studies have a large influence on the field.[1] Empirically, CEO effect studies are important because they provide baseline evidence for the relevance of CEOs (Chiu & Walls, 2019; Fitzsimmons & Callan, 2016; Schepker, Kim, Patel, Thatcher, & Campion, 2017). If the CEO effect is much greater than the firm and industry effects, as Hambrick and Quigley (2014) claim, CEO studies should become a centerpiece of strategic management research. In contrast, if the CEO effect is indistinguishable from chance, as Fitza (2017) claims, much of the prior CEO effect studies would be invalid, and one could even go so far as to question the meaningfulness of studying how CEOs affect firm performance.[2] Methodologically, Hambrick and Quigley (2014) has been replicated by Keller, Glaum, Bausch, and Bunz (2023), and their work has also served as a template for gauging the level of CEO impact on

☆ This research was supported in part by a grant from the Academy of Finland (grant 311309), Liikesivistysrahasto (grant 14-7574), and Marcus Wallenberg Foundation. We acknowledge the computational resources provided by the Aalto Science-IT project.

* Corresponding author.
E-mail address: mikko.ronkko@jyu.fi (M. Rönkkö).

[1] At the time of writing, according to the ISI Web of Science, the CEO effect studies conducted by Crossland and Hambrick (2011) and Quigley and Quigley and Hambrick (2015) were the second most cited empirical studies published by SMJ during those years, and the study by Hambrick and Quigley (2014) was among the ten most cited articles.

[2] As Hambrick and Quigley (2014) note, examining the CEO effect "has the benefit of gauging the overall impact of CEOs" and thus "attention to overall CEO effects and attention to specific CEO attributes […] can be thought of as highly symbiotic" (p. 475). Indeed, in explaining how performance varies under different CEOs with distinct CEO attributes (e.g., Chatterjee and Hambrick (2007); Nadkarni and Herrmann (2010) is only meaningful in the context of demonstrable variation in performance.

https://doi.org/10.1016/j.leaqua.2023.101733
Received 19 May 2022; Received in revised form 16 August 2023; Accepted 19 August 2023
Available online 28 August 2023

other outcomes, such as corporate social responsibility (CSR) (Wernicke, Sajko, & Boone, 2022). CEO effect studies also have direct policy implications. For example, the increase in CEO pay over recent decades could be explained if the magnitude of the CEO effect has indeed increased over time (Quigley & Hambrick, 2015).

The apparent paradox we noted above can be resolved by first recognizing that the disagreement is a methodological disagreement rather than a conceptual or empirical one. Both Hambrick and Quigley (2014) and Fitza (2014, 2017) use essentially the same data and start from the same conceptual definition of the CEO effect as the average difference in performance that is attributable to a CEO of a given firm. While Hambrick and Quigley (2014) argument that CEO effect studies should consider the evolving contexts of CEOs is conceptually sound, we show that their empirical approach still has confounding issues that lead to an overestimation of the CEO effect. We further argue that instead of using Hambrick & Quigley's (2014) CEO in context approach, an evolving context is best modeled using a mean-reverting autocorrelation structure (Fitza, 2017), which allows for consideration of the well-established fact (e.g., Fama & French, 2000; Waring, 1996) that "shocks" to firm performance in one period tend to persist into the following periods (i.e., a firm with exceptionally high performance in one year tends to have above average performance in the following year). We develop and validate a new empirical approach that demonstrates that the use of a multilevel model with such an autocorrelation structure can disentangle the effects of firm, CEO, and time on firm performance (Blettner, Chaddad, & Bettis, 2012).

Using this empirical setup, we find a CEO effect of 11.5%, which is considerably smaller than the effect range found in previous studies (Hambrick & Quigley, 2014; Quigley & Graffin, 2017) and approximately half the size of the CEO effect found in recent comparable studies (e.g., Quigley and Graffin (2017), who measure a CEO effect of 21.8%). This result resolves the apparent paradox noted above and implies that neither the conclusion that "the effect of CEO leadership is almost indistinguishable from the effect of luck and chance" (Fitza, 2017, p. 809) nor the conclusion that "Our sizeable CEO effect indicates that some CEOs are able to alter the trajectories of their firms—perhaps appreciably more than previously thought" (Hambrick & Quigley, 2014, p. 488) are warranted when the CEO effect is properly contextualized. Given that the existence of a positive CEO effect has consequences for both policy decisions and research in the field of strategy, our results are important because they establish there remains a measurable and non-negligible CEO effect even after accounting for the methodological problems that have been raised by prior studies (Blettner et al., 2012; Fitza, 2014, 2017)..

To further demonstrate that the use of our analysis approach not only produces a smaller CEO effect but can also alter the substantial conclusions that have been reached in published papers measuring the CEO effect, we replicated the analysis of Quigley and Hambrick (2015), which found that the CEO effect has substantially increased over time. However, when accounting for autocorrelation in firm performance, we do not find conclusive evidence that the CEO effect has substantially increased over time. In contrast, we find that changes in the CEO effect over time are much smaller once we consider persistence in firm performance, and some of the differences that have been detected over the years in CEO effects across different time periods are not statistically significant. Using our approach, we thus show that there is reason to doubt the conclusion reached by Quigley and Hambrick (2015) that the CEO effect has increased over time. This is important because an increasing CEO effect could be used to justify recent increases in attention or compensation being awarded to CEOs, which our result challenges.

Our study also contributes to recent discussions concerning methodological rigor (e.g., Bergh, Sharp, Aguinis, & Li, 2017; Bettis, Ethiraj, Gambardella, Helfat, & Mitchell, 2016) in strategic management studies in general and leadership studies in particular that go beyond CEO effect studies. By using Monte Carlo simulation, we demonstrate that our analysis approach is effective at estimating the CEO effect, which has not been demonstrated by prior studies proposing novel techniques for estimating the CEO effect or by any other variance decomposition studies (e.g., Guo, 2017). Our study highlights the importance of providing methodological evidence when proposing novel methodological approaches to address key leadership research questions.

Performance variation and the CEO effect

The challenges of measuring the CEO effect

Starting with the seminal study by Lieberson and O'Connor (1972), strategy and leadership scholars have been interested in examining the degree to which CEOs matter to firm performance both in absolute terms and as relative to stable industry and firm-level influences (Hambrick & Quigley, 2014). The "leadership effect" identified by Lieberson and O'Connor (1972), which has subsequently become known as the "CEO effect," measures the portion of the variance in firm performance that occurs between the tenures of different CEOs of the same firm. Current CEO effect studies typically decompose the overall variance in observed firm-level performance (typically in return on assets [ROA]) into three aspects of interest (industry, firm, and CEO) while controlling for year-specific effects and including an observation level error term using the following models:

$$\begin{array}{llll} Performance_{tijk} = & \beta_0 + \beta_{0k} + & \text{Level } 4: & \text{Industries } k \\ & \beta_{0jk} + & \text{Level } 3: & \text{Firms } j \\ & \beta_{0ijk} + & \text{Level } 2: & \text{CEOs } i \\ & \beta_1 Time_1 + \cdots + \beta_n Time_n + \epsilon_{tijk} & \text{Level } 1: & \text{Year s } t \end{array} \quad (1)$$

where $\beta_n Time_n$ are the fixed effects of time shared between all firms and industries, β_0, β_{0k}, β_{0jk}, and β_{0ijk} are the grand, industry, firm, and CEO intercepts that measure the share of variance attributable to time-invariant factors, respectively, and ϵ_{tijk} is the observation-specific error term. Variance decomposition studies using Eq. (1) or the equivalent are thus based on the fundamental assumption of stable and time-invariant effects at the level of the industry, the firm, and the CEO after accounting for macroeconomic trends that exert the same impact on every firm (Hambrick & Quigley, 2014). Thus, the only two firm-specific factors that vary over time are the error term, which takes different values for each year, and the CEO effect, which takes different values for the different CEO tenures.

Recently, variance decomposition studies have faced severe criticism from Blettner et al. (2012), Hambrick and Quigley (2014), and Fitza (2017). The central argument in these three articles is that performance is not independent between years but rather tends to persist or show trends over time at both the industry and the firm level. For example, as Hambrick and Quigley (2014) note, actions by CEOs that lead to "enhancing or impairing their companies' brands, technology pipelines, or cultures" (p. 479) may have an effect that lasts over time even beyond their specific tenures. Indeed, it is well established that firm performance is affected by both stable (time-invariant) and changing (time-variant) factors[3] (the latter are often called "transient factors"; e.g., Rumelt (1991); Misangyi et al. (2006); similarly Guo (2017) distinguishes "stable" from "dynamic" variance). Conceptually, whereas stable firm-level differences are used to capture the effect of firm-specific characteristics that are immutable over time, changing firm-level differences are used to capture the effect of changes in firm-specific characteristics that are neither immutable nor completely random but rather exhibit a certain *persistence* over time. However, while the three studies

[3] Furthermore, assumptions about which factors are stable and which vary over time may lead to drastically different results, as demonstrated by the classic debate between Rumelt (1991) and Schmalensee (1985).

mentioned above agree that firm performance depends on evolving contextual factors, they disagree about the implications of this fact for the CEO effect: Hambrick and Quigley (2014) find a much greater CEO effect than prior studies had found, whereas Fitza (2017) suggests that this effect is much smaller than that found by prior studies. Blettner et al. (2012) provide a pessimistic assessment that the “triple confounding of firm, CEO, and time […] will be difficult or impossible to overcome with current statistical technology” (p. 990).

Triple confounding, changing firm-level differences and autocorrelation

To understand what Blettner et al. (2012) meant by the problem of “triple confounding” in CEO effect studies, it is useful to consider that both industry- and firm-level effects can be decomposed into stable and changing effects whereas the CEO effect is assumed to be stable by definition. Switching to a different notation to emphasize that what follows is a conceptual discussion of what we would like to estimate instead of a statistical discussion on how exactly we do the estimation, we can decompose the industry and firm terms in Eq. (1) as:

$$\begin{aligned} Performance_{tijk} = \; & Baseline\ level + \\ & Stable\ industry\ effect_k + \\ & Changing\ industry\ effect_{tk} + \\ & Stable\ firm\ effect_{jk} + \\ & Changing\ firm\ effect_{tjk} + \\ & CEO\ effect_{ijk} + \\ & Year\ effect_t + \\ & Year\ to\ year\ random\ variation_{tijk} \end{aligned} \tag{2}$$

Eq. (2) reveals three potentially confounding sources and thus three ways in which the CEO effect could be measured incorrectly. First, because the observed variance of performance between firms is the sum of the seven sources of variance shown in Eq. (2), omitting any of these sources from the analysis means that the effect of one or more of the sources that are included in the model will be overestimated. This effect, for example, has been demonstrated by Fitza (2017), who simulated data with changing firm-level effects (i.e., an evolving context; Hambrick and Quigley (2014)) and demonstrated that if these effects were to be ignored in the estimated model, this source of variance would be erroneously attributed to the CEO effect (see our Appendix B for details on the underlying mechanism). Second, even if all sources of variance were included in the model, there would be a substantial risk of confounding because of overlap between the effects. Most importantly, because CEOs are nested in firms, there is a complete overlap between the stable firm effect and the CEO effect, which leads to confounding when ANOVA and other fixed effects models are used (Blettner et al., 2012; Misangyi et al., 2006; Quigley & Graffin, 2017). Third, because the number of observations per CEO is generally small, variance due to year-to-year random variation can be incorrectly attributed to the CEO effect in fixed-effects models such as ANOVA (Fitza, 2014).

The use of multilevel modeling has been shown to avoid the confounding of the CEO effect with both the stable firm effect and random year-to-year variation while also addressing the problem of the typically small number of CEO observations (Quigley & Graffin, 2017). However, empirically taking changing firm-level differences into account has proven to be a more challenging task. The key challenge is that one cannot assume that performance is simply the sum of time-invariant effects (as in Eq. (1)) and year-to-year random variation (Hambrick & Quigley, 2014). Rather, the presence of changing firm-level differences means that firm performance correlates over time, which gives rise to a unique performance trajectory over time for each firm (e.g., Henderson, Raynor, and Ahmed (2012); Mueller (1977)). This outcome is often referred to as autocorrelation (Blettner et al., 2012; Fitza, 2017), which captures the empirical regularity of persistence in performance between two adjacent time periods and can be formally expressed as:

$$\begin{aligned} Changing\ firm\ effect_{tjk} = \; & \rho \cdot Changing\ firm\ effect_{(t-1)jk} \\ & + Random\ shock_{tjk} \end{aligned} \tag{3}$$

where ρ measures the extent to which the effect of random “shocks” on performance (again, typically reflected in ROA) persists between two adjacent time points.[4] The impact of such shocks does not persist indefinitely, however, but generally diminishes over time, thus exhibiting a mean-reverting random walk (Blettner et al., 2012), where ρ measures how quickly the effect reverts to the mean ($\rho = 0$ implies no autocorrelation, whereas $\rho = 1$ is a pure random walk, which diverges over time (Denrell, 2004). There is indeed substantial empirical evidence across disciplines to indicate that firm performance is autocorrelated (Fama & French, 2000; Fama & French, 2006; McGahan & Porter, 1999; Short, Ketchen, Bennett, & du Toit, 2006; Wiggins & Ruefli, 2002) and follows a distinct, mean-reverting trajectory over time (Blettner et al., 2012; Waring, 1996).[5]

Note that in Eq. (3), the variable *Changing firm* effect$_{tjk}$ is indexed by j (firm) and k (industry), which means that performance is modeled as a firm-specific time series exhibiting a mean-reverting random walk; thus, each firm is allowed to have its own performance trajectory (this is the case even if ρ is assumed to be the same for all firms).[6] Note that this approach is similar to but more general than Guo’s (2017) approach, which imposes the restrictive assumption of a U-shape to the changing firm-level differences in his study. To simply illustrate such a persistent shock, if a firm had an exceptionally good year, it would be more likely to also have an above average performance in the following year. Therefore, ignoring the fact that shocks to firm performance persist and rather assuming that the firm-level context is constant throughout the period under study (Hambrick & Quigley, 2014) would obviously be a mistake and would cast doubt on the results (Fitza, 2017).

Empirically distinguishing CEO effect and autocorrelation

The fact that both the CEO effect and the changing firm effect vary over time makes empirically distinguishing between the two

[4] Technically, autocorrelation parameter can be defined for any random effect term that varies over time. In most guidelines and applications, the only time-varying random effect is the lowest level error term. This error term and the corresponding autocorrelation parameter would then reflect all time-varying effects that are not explained by the fixed part of the model regardless of the source. While it would technically be possible to define and estimate a model using two “error terms” so that the time-varying effect is further decomposed into firm an industry levels, this solution is not very practical in the context of CEO effect studies even without applying autocorrelation because it requires a model with crossed random effects. To add the industry level “error term” to Equation (1) would involve adding a new level, namely, industry-year, as another level 3 unit crossed with firms. Such models are not only difficult but may be impossible to estimate with most multilevel software (Hox, 2010, Chapter 9). This problem would be further compounded in CEO effect studies because the CEO level also spans over multiple years and would need to become crossed with industry-year as well. As explained later, including cluster (i.e., industry-year) means as fixed effects to account for industry level trends is a much more practical solution for this problem.

[5] Scholars have also estimated the degree of profitability persistence ρ, with estimates varying between $\rho = 0.38$ (McGahan & Porter, 1997) and $\rho = 0.62$ (Fitza, 2017).

[6] Formally, all firms are assumed to have the same long run values of ρ and $var(Random\ shock)$ in the population, but the variance and degree of autocorrelation can differ among the firms in a sample because of the short time periods that are typical set for CEO effect studies (e.g. 20–30 observations per firm). While it would be possible to us a mixed effects scale-location model (Wang, Hamaker, & Bergeman, 2012) that allows both $var(Random\ shock)$ and ρ to vary between industries or even firms, these kind of models are difficult to estimate in short panels. Moreover, in the context of CEO effect studies, using a more parsimonious model where all firms share the same ρ may avoid the potential confounding of CEOs and the changing firm effect (Fitza, 2017).

challenging, which has led some CEO effect scholars to be pessimistic about whether it is even possible to do so. Specifically, Blettner et al. (2012) note that "The problems introduced by the mean reverting random walk nature of returns are at least partially beyond the capabilities of current statistical technology," continuing that this is "especially so for CEO fixed effects" (p. 989). Similarly, Fitza's (2017) study shows that the problem of confounding the CEO and the changing firm effects is not unique to fixed effects models but also applies to random effect models, which led him to the conclusion that "*on average* the effect of CEO leadership is almost indistinguishable from the effect of luck and chance" (p. 809).

However, the fact that these two effects are easily confounded does not imply that the effects cannot be empirically distinguished. Upon closer inspection, the two effects produce distinct empirical patterns, as illustrated in Fig. 1. The figure shows simulated ROA time series for one firm over five consecutive CEOs, each with six years of tenure (i.e., there are 30 firm-level observations but only six per CEO), with different combinations depending on whether autocorrelation is present (left vs. right; we use $\rho = 0.9$) and whether there is a CEO effect (top vs. bottom; we use a CEO effect of 75%). The values we use are larger than those observed in empirical studies (which is the case in particular for the CEO effect); however, using extreme values is useful for demonstration purposes, as it makes the difference between plots more clearly visible. The dashed lines represent the expected ROA, which is calculated as the sum of the stable effect parameters and uses a LOESS curve for the autocorrelated error term.

Plot 1 shows how the value a purely random variable varies around its trend. As can be easily seen, there are nevertheless differences between the mean ROA during the tenure of each of the five CEOs due to the small sample size for each CEO; this is the confounding effect when using fixed-effects models such as ANOVA, which was demonstrated by Fitza (2014). Plot 2 in the first column demonstrates a scenario where performance varies systematically between the tenures of different CEOs but is otherwise random. The extremely large CEO effect used in our simulation makes the difference between the first and second plots clearly visible.

The second column in Fig. 1 shows an example of an autocorrelated time series in Plot 3 and a combination of autocorrelation and CEO effect in Plot 4. When comparing Plot 2 (only CEO effect) and Plot 3 (only autocorrelation), it is interesting to note that both exhibit systematic differences over time but with clearly different patterns. Fitza (2017) demonstrated that these two patterns are easily confused and that the effects can be confounded even when random effect models are used. The empirical challenge to avoid confounding the CEO effect and the autocorrelation is in statistically distinguishing between these patterns. This is illustrated in Plot 4, where both effects are present. While it is clear that we cannot empirically distinguish between these two effects by simply separately estimating each from the same dataset[7] (Fitza, 2017), the fact that the empirical patterns in the different plots clearly differ implies that it should be possible to distinguish between them in principle.

Accounting for changing firm-level differences but overestimating the CEO effect

So how can this be accomplished? The study by Hambrick and Quigley (2014) advances what is thus far the only serious attempt at addressing the issue of changing firm-level differences in a CEO effect study. Their empirical approach starts from the premise that to properly evaluate the influence of CEOs on firm performance, one should "contextualize" CEOs by gauging their performance against the current industry trend and their firm-level "inherited performance" rather than using the sum of long-run averages of industry and firm performance (as in Eq. (1)). To control for changing industry-level differences, Hambrick and Quigley (2014) model the industry effects (both stable and changing effects) as industry means excluding the focal firm. This approach effectively eliminates both stable and changing effects on the industry level from the data, simply leaving the question of how the triple confounding of firm, CEO, and time can be avoided. To control for firm effects (both stable and changing), Hambrick and Quigley (2014) operationalize the changing firm-level context by constructing a lagged dependent variable as a mean of the ROA over the two years prior to the new CEO taking office as a firm's health measure at the point of CEO succession.[8] However, in their empirical analysis, the effect of firm health was small, and thus, in practice, their approach was simply to control for persistence of performance by using a lagged dependent variable as a control.

Using this empirical setup, Hambrick and Quigley (2014) measure a surprisingly large CEO effect of 38.5%, which is a counterintuitive result; if firm-level performance variation is a function of the seven variance sources shown in Eq. (2), then omitting any of these sources from the estimated model would inflate the estimates of the other variance sources, as demonstrated by Fitza's (2017) study. Thus, as the prior CEO effect studies that did not control for a changing context produced estimates in the range of 15–20% (Hambrick & Quigley, 2014; Quigley & Graffin, 2017), we should expect a model that includes appropriate controls for a time-varying context to produce a *smaller* CEO effect estimate. However, the 38.5% effect found by Hambrick and Quigley (2014) is approximately twice as large as that typically found, which suggests that important sources of confounding remain in their modeling approach.

We can identify two sources of confounding in Hambrick and Quigley's (2014) study. First, the use of a lagged dependent variable can only control for the changing firm effect but produces results that are biased and can be severely misleading if there are stable firm effects in the data (Hamaker, Kuiper, & Grasman, 2015). The fact that stable firm level differences are not controlled for in their model leads to a major inflation of the CEO effect because omitting the firm fixed effect in fixed effects models leads to the CEO fixed effects also capturing all (stable) firm level variance, as noted in previous research (Blettner et al., 2012; Misangyi et al., 2006; Quigley & Graffin, 2017). Second, by estimating a separate mean ROA level for each CEO (i.e., a fixed effect), their study confounds the CEO effect with random noise due to the small number of observations per CEO, which leads to inflating the CEO effect as noted above (Fitza, 2014; Quigley & Graffin, 2017).

To summarize, while Hambrick and Quigley's (2014) premise for contextualizing the CEO effect is conceptually sound, their empirical approach unfortunately leads to a substantially inflated CEO effect.

[7] To demonstrate that these effects are easily confounded, we conducted a series of ANOVAs using the data shown in Plots 2–4 of Fig. 1. The R2 varied between 0.65 and 0.75, thus showing no substantial differences between the datasets (i.e., we measure a CEO effect in approximately the same range regardless of whether there is a true CEO effect in the data or not). We also estimated autocorrelations by fitting ARIMA models to the same three datasets, producing autocorrelation estimates between 0.73 and 0.78 (i.e., we measure a large autocorrelation parameter ρ regardless of whether there is autocorrelation in the data or not).

[8] We note that Hambrick and Quigley (2014) also used GEE estimation in which the error covariance matrix has been parameterized with autocorrelation but did not explain the reason for doing so. While GEE estimation does produce more accurate fixed effects estimates than normal OLS regression analysis when the error correlation structure is correctly specified (McNeish, Stapleton, & Silverman, 2017), it does not eliminate the estimation errors of individual CEO fixed effects and therefore still produces inflated variance estimates, as explained later in the article.

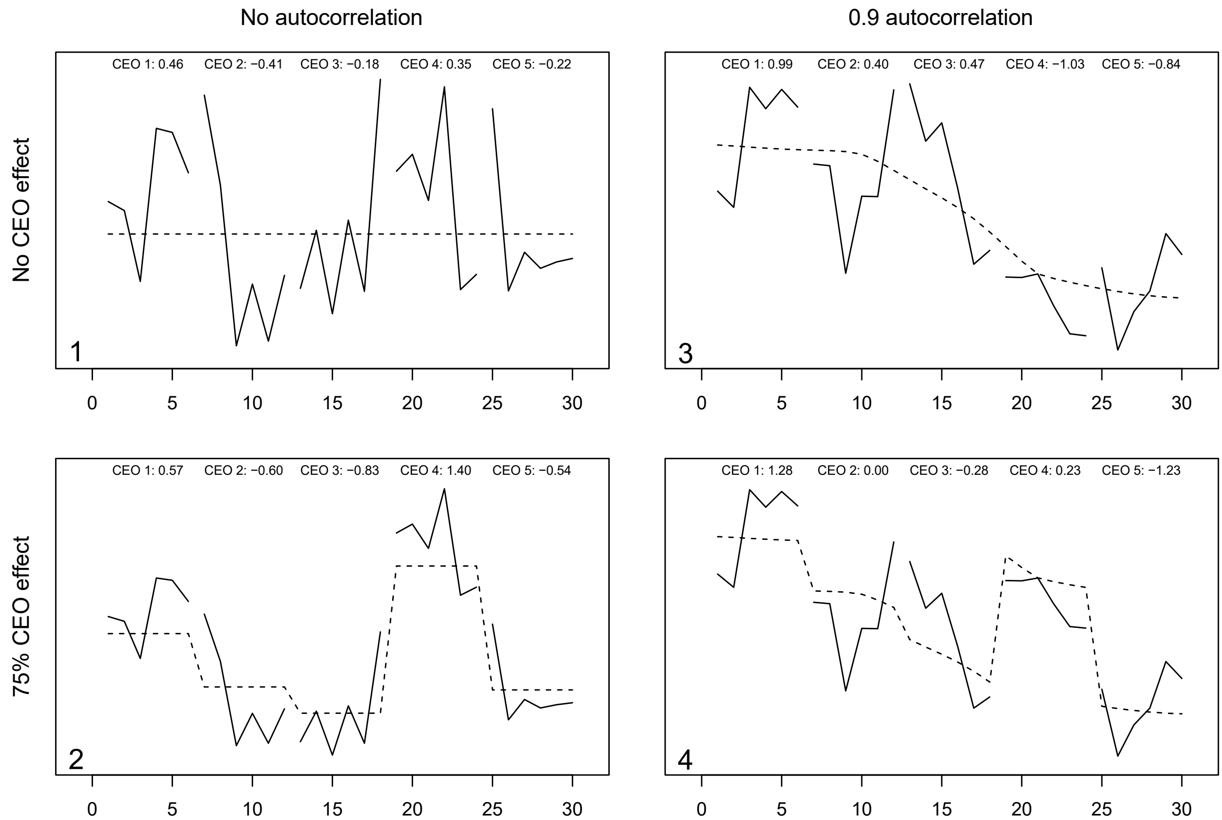

Fig. 1. Four simulated firm performance trajectories over five different CEOs over 30 years as organized by the level of autocorrelation and CEO effect. *Note.* Solid lines show the simulated performance data, and the discontinuities in those lines indicate a change in CEO. The dashed lines represent the expected ROA, calculated as the sum of the stable effects parameters and a LOESS curve for the autocorrelated error term. The numbers after the CEO labels indicate the average performance of a CEO expressed as standard deviations from the mean performance of their firm.

Blettner et al.'s (2012) conclusion that resolving the triple confounding issues may not be possible could be correct for fixed effects models. However, both the bias from small samples and the problem of confounding due to the overlap between the firm and CEO fixed effects can be addressed through the application of multilevel modeling (Quigley & Graffin, 2017), and thus, the only remaining part of the triple-confounding issue left to account for is the changing firm effect. The two options for accomplishing this are to either include a lagged dependent variable or to include an autocorrelated error term (as in Eq. (3)). While the former approach has been recommended in prior literature (Bergh et al., 2016; Semadeni, Withers, & Trevis Certo, 2014), it is not applicable if there are stable firm-level differences in the data (which would require the inclusion of firm-fixed effects), because the error term and the lagged dependent variable will be correlated, resulting in an endogeneity problem known as "dynamic panel bias" (Bou & Satorra, 2018; Dishop & DeShon, 2022; Hamaker et al., 2015).[9] Therefore, the use of a lagged dependent variable does not provide an effective solution for modeling changing firm effects. On the other hand, directly including an autocorrelated error term in the model is a much more attractive approach. Indeed, this approach is recommended in many multilevel modeling guidelines (Bliese & Ployhart, 2002; Short et al., 2006) and has also recently been specifically recommended for variance decomposition studies that measure performance (Guo, 2017).

In conclusion, to account for changing firm-level differences in a way that avoids the potential triple confounding of firm, CEO, and time effects (Blettner et al., 2012) and thus does not lead to an (upward) biased CEO effect estimate, one should use a multilevel model that includes an autocorrelated error term, which can account for both changing firm effects and random year-to-year noise in addition to stable firm effects. Given that applying industry-year means as controls eliminates all industry-level trends (Hambrick & Quigley, 2014), such a model should in principle be able to empirically distinguish between the CEO effect and autocorrelation in firm performance while also avoiding confounding CEO and stable firm effects. In the next section, we set up and estimate just such a multilevel model including autocorrelation. Not only has this type of model not previously been applied to estimate CEO effects, but it also amounts to appropriately contextualizing the CEO effect by measuring the level of CEO contribution to firm performance against an evolving and changing firm-specific context rather than a long-run average of firm performance, as has been argued by Hambrick and Quigley (2014).

Methods and results

Here, we specify and run a multilevel model (also known as hierarchical linear model; HLM) that, in addition to industry-, firm-, and CEO-level effects, allows for autocorrelation in firm performance. By adding an autoregressive component, we can differentiate persistence in performance from other effects (and, in particular, from the CEO effect) to address the points raised by Blettner et al. (2012) and later by Fitza (2017).

Sample

To make our results comparable to those of prior studies, we rely on data aligned with data used in recent studies, such as that of Fitza (2014) and that of Quigley and Graffin (2017). We collected CEO information from the ExecuComp database and company financial information from

[9] For an accessible demonstration, see https://statisticalhorizons.com/lagged-dependent-variables.

the Compustat database covering the years of 1992–2015. Following earlier research on the CEO effect (Fitza, 2014; Mackey, 2008; Quigley & Graffin, 2017), we eliminated all financial companies and government organizations (SIC codes 6000–6999 and 9000–9999) and retained only those firms that had at least 20 million dollars of assets in each year across the entire observation period. We also followed existing practice and dropped the largest and smallest 1% year observations of ROA. To estimate industry effects, we eliminated all companies that were solely observed by their industries each year.[10] This resulted in an unbalanced panel dataset containing 28,026 observations covering 5,136 CEOs of 2,407 companies in 196 industries (three-digit SIC).

Modeling approach

The initial model consisted of yearly observations (L1) nested in CEOs (L2) nested in firms (L3) nested in industries (L4), and it used ROA as the dependent variable. Prior research on the CEO effect has eliminated economy-level trends by including year dummies in the analysis (Mackey, 2008; Wasserman, Anand, & Nohria, 2010), which, in this scenario, is equivalent to mean centering the data by year, which is known as the with-in transformation and is the first step in the fixed effects estimator (Certo, Withers, & Semadeni, 2017; Wooldridge, 2009, Chapter 14). We opted for this latter approach because it simplifies our modeling. Following Hambrick and Quigley (2014), we modeled the industry trends using their "CEO in Context" (CiC) technique. In this approach, an estimated value of industry mean ROA is calculated for each firm-year observation using all other firms from the same industry and year.[11] This new variable is then used as a fixed effect in the multilevel model. In the case of CEO effect studies, using industry-year means as a control has the effect of eliminating all industry-level trends from the data and thus allows for the interpretation of the remaining variance components as firm-level effects (Antonakis, Bastardoz, & Rönkkö, 2021).

To properly account for an evolving firm-level context, we specified that the error term follows an autocorrelation (AR1) pattern over time within firms, as shown in Eq. (3), following the recommendations given in the methodological literature (Bliese & Ployhart, 2002; Short et al., 2006).[12] The fully specified model is thus:

$$ROA_{tijk} = \beta_0 + \beta_1 \overline{ROA}_{tjk} + a_t + u_{jk} + u_{ijk} + \epsilon_{tjk} \quad (4)$$

where β_0 is the intercept, $\beta_1 \overline{ROA}_{tjk}$ are the industry effect following the CiC approach, a_t are the time fixed effects that are eliminated by centering, u_{jk} is the firm level random intercept, u_{ijk} is the CEO level random intercept ("CEO effect"), and ϵ_{tjk} is the error term autocorrelated at ρ_{AR1}. Because the industry effect ($\beta_1 \overline{ROA}_{tjk}$) effect varies over time, it captures both the stable and changing firm effects in the conceptual Eq. (2), the random effects u_{jk} and u_{ijk} capture the stable firm effect and the CEO effect respectively, and because the error term ϵ_{tjk} autocorrelates within firm, it captures both changing firm effect and random year to year variations. (See Appendix A for more discussion on the level of the autocorrelation parameter.).

While considering different error structures is a standard practice in multilevel modeling (Bliese & Ployhart, 2002; Short et al., 2006), applying these techniques in CEO effect studies is difficult because autocorrelated errors are allowed only *within the lowest-level groups* in commonly used multilevel modeling software (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011, sec. 9.4; StataCorp LP, 2015, pp. 386–388),[13] which is the CEO level in CEO effect studies rather than the firm level. This means that the error terms are still constrained to be independent among the lowest-level groups. In the context of CEO effect studies, these models thus allow autocorrelation of performance within a CEO tenure but constrain the errors to be uncorrelated across the tenures of the distinct CEOs of a firm, making it impossible to test Fitza's (2017) alternative explanation that the observed CEO effect is simply an outcome of performance trends that span the tenures of multiple CEOs.

The lack of support for a particular model parameterization by statistical software can be overcome by using general maximum likelihood estimation tools that allow for the maximization of a user-defined likelihood function. Therefore, we wrote our own likelihood function in R and used the *maxLik* package (Henningsen & Toomet, 2011) for estimation. The starting values for estimation were the same as those used by Stata's *mixed* command, which were obtained using the expectation maximization algorithm in Stata.[14] We validated our estimation routine by applying it to simplified versions of our main model that can be estimated with Stata. The log likelihood and all estimates were identical to the results produced by Stata's *mixed* command. Additionally, we performed a series of Monte Carlo simulations to rule out the possibility that our results are idiosyncratic to a particular sample. We used a 5x3 full factorial design varying the CEO effect between 0% and 25% at increments of 5% and the AR1 at 0, 0.25, and 0.50 with 1000 replications for each cell. The results demonstrated that our estimation approach was effective in capturing the CEO effect and autocorrelation parameter from the data without confounding the two. The variance estimates of the CEO effects as well as any differences in the CEO effect were calculated using cases bootstrap (Leeden, Meijer, & Busing, 2008) resampling at the firm level. The analysis was run on a computer cluster for computational feasibility. The analysis files and the results of the Monte Carlo study are included in Appendix A, and the Compustat dataset is available from the first author by request.

Results

The main results using the Compustat data are presented in Table 1. We started by estimating a random intercept model using all levels of data but excluded the autoregressive term (Model 1). Using this model specification, the CEO effect estimate is 23.8%, which is close to the 21.8% estimated by Quigley and Graffin (2017). We obtained this result

[10] To validate our dataset, we excluded all CEOs with single year tenure and we limited our data to the period of 1993–2012, following Fitza (2014) and Quigley and Graffin (2017). A comparison of the number of observations and other descriptive statistics revealed only trivial differences between the datasets. However, these exclusions were not applied to the version of the data used in our main analysis.

[11] ROA is calculated as $ROA = \frac{Income\ \ Before\ \ Extraordinary\ \ Items}{Total\ \ Assets}$. Industry mean ROA excluding the focal firm is calculated as $\overline{ROA}_{tjk} = \frac{\sum_{x\in\{J_k - j\}} ROA_{txk}}{|J_k|-1}$, where j is the focal firm that is excluded from the mean, and J_k is the set of firms in industry k.

[12] Given that we assume that firm performance can have different trajectories under different CEOs, it is reasonable to ask if there are some typical trajectories that firm performance follows. To understand this, we applied cluster analysis and latent class analysis to ROA within CEO tenures. These analyses did not reveal any typical patterns. We report these analyses in Appendix C.

[13] Multilevel models can be equivalently estimated by converting the data from long to wide format and applying SEM (Bou & Satorra, 2018). This approach allows specifying dynamic models that overcome the endogeneity problem which arises when a lagged dependent variable is used together with a random intercept in a mixed model, as explained earlier. While specifying a stable and changing firm effects model would be straightforward using this approach, modeling the CEO effect would be complicated because the pattern of CEO succession differs among firms. In theory, this could be handled by specifying a multigroup model where each firm forms its own group. Indeed, this approach would be equivalent to our multilevel model, but much more difficult to implement either programmatically or computationally.

[14] When estimating models with autoregression parameter using mixed, Stata initially sets the autoregression parameter to be zero when calculating the starting values.

Table 1
Mixed effects regressions of ROA.

	Model 1	Model 2	Model 3
Fixed effects			
Industry mean	.376	.329	.328
	(.0115)	(.0121)	(.0120)
Intercept	−.555	−.359	−.421
	(.107)	(.104)	(.0994)
Random effects			
Firm	15.58	15.19	12.90
	(.425)	(.402)	(.417)
CEO	14.35		6.81
	(.322)		(.299)
Error variance	3.49	43.09	39.40
	(.145)	(.286)	(.292)
AR1 (within firm)		.527	.459
		(.00915)	(.00793)
Log likelihood	−91711.5	−90266.2	−90164.3
CEO effect	23.7%		11.5%
	[20.9, 26.5]		[8.4, 14.7]
CEO effect difference			12.2%
			[10.3, 14.1]

Note. N = 28,026 firm-years; 196 three-digit SIC industries, 2,407 firms, and 5,136 distinct CEOs. The AR1 parameter is a correlation metric. Models 1 and 2 estimated with Stata's *mixed*. Model 3 estimated with our R code. The delta method standard errors are displayed in parentheses. The CEO effect is calculated as the ratio of the variance of the CEO level random intercept to the sum of all variance components, and its variance estimate is calculated with case bootstrapping. 95% bootstrap percentile confidence intervals in brackets. The ROA estimate has been multiplied by 100 to make the variance components easier to interpret.

using similar data[15] and analysis methods, so the results are comparable. Similarly, the effect of industry means is comparable to the estimates presented by Hambrick and Quigley (2014). Next, Model 2 adds an AR1 error structure while omitting the CEO effect so that the firm level becomes the lowest-level group, and the AR1 error structure therefore applies across all errors within a firm. The large autocorrelation coefficient (AR1) provides evidence for the existence of strong performance trends at the firm level. While Model 1 and Model 2 are non-nested and therefore cannot be directly tested for improved fit, comparing the log likelihoods of each shows that Model 2, which includes the autocorrelation term, explains the data substantially better than Model 1 does, including the CEO effect.

Finally, Model 3, which we estimated using our own likelihood function in R, includes both CEO and firm effects as well as the within-firm autocorrelated error structure in the same model and thus accounts for an evolving firm-level context. Here, we find a CEO effect of 11.5% (compared to 23.7% in Model 1, which omitted autoregression). This empirically confirms that omitting autocorrelation leads to confounding and in fact inflates the CEO effect. Likelihood ratio tests further show that Model 3 fits the data significantly better than either Model 1 ($\chi^2_{(1)} = 1547.17$) or Model 2 ($\chi^2_{(1)} = 101.87$). This result formally establishes the fact that there is a statistically significant CEO effect even after accounting for autocorrelation. Comparing the coefficients between Model 1 and Model 3 also reveals that while adding the autoregression parameter to the model decreases both CEO and firm effects, the effect on the CEO level variance component is larger (−17% vs. −53%). The results thus demonstrate that a misspecified model that does not take autocorrelation into account largely attributes the data trends to the CEO effect.

Robustness checks

We performed four analyses to check the robustness of our above results, which are shown in Table 2. First, as stated above, we trimmed the most extreme 1% of ROA values from both ends of the distribution in our main analyses to be comparable with prior research. To conduct our first robustness check, we ran our models (Models 4–6) without trimming extreme ROAs. Unexpectedly, this increased all variance estimates, but it primarily increased the error term variance in such a way that the estimated CEO effect was minimized across the board (11.3% in Model 4 and 7.8% in Model 6). Both the effects of industry mean and the AR1 parameter were similarly reduced by approximately half, which is expected when noise is added to the data. We did not estimate the confidence intervals for CEO effects of the robustness checks to save computational time.

Second, following Certo, Busenbark, Kalm, and LePine (2020), many strategy scholars have recently begun to challenge the use of ratios as dependent variables. Both Certo et al. (2020) and Wiseman (2009) raised the concern that using ratios in models can lead to incorrect conclusions about regression coefficients. This specific concern does not apply to variance decomposition studies where the regression coefficients are not of primary interest. In fact, whether using ratios in variance decomposition analysis is problematic has yet to be methodologically examined. Nevertheless, we decided to replicate our study using the "Use unscaled variables/control for scale" strategy (Certo et al., 2020, p. 231). Following this strategy, we re-estimated our models by using net income as the dependent variable and added assets as a control variable. The results from the net income analysis, shown in Models 7–9, show that all variance components are much larger than those in the ROA models and the CEO effects are approximately twice as large as those in the original models, but the general pattern of results is the same as that derived from the ROA models: the model including autocorrelation but no CEO effects fits the data better than the model that includes CEO effects but no autocorrelation. However, the model with both parameters fits better than either of these two models. Modeling the autocorrelation of net income also reduces the CEO effect to approximately half, as with our main model. Note that these results are not technically comparable to prior CEO effect studies (which predominantly use ROA as the dependent variable) in terms of the magnitude of effect. While controlling for assets may be an effective strategy for controlling for scale when estimating the regression coefficients, it does not necessarily control for scale when estimating the variance components, as the large size of the variance estimate values show.

Third, we ran our analysis using Tobin's Q as the dependent variable. While ROA is a commonly used performance measure in CEO effect studies, it is not the only performance variable that can be used.[16] Particularly, ROA has the problem of focusing on the current performance and not considering the future outlook of the company. To address this issue, some researchers have used market-based measures (Quigley & Hambrick, 2015). Following this approach, we performed a variance decomposition analysis using Tobin's Q as the dependent variable (Models 10–12). The results of this analysis are interesting. Using a traditional variance decomposition analysis, the CEO effect is approximately the same as that when using ROA (Model 1 in Table 1 vs. Model 10 in Table 2). When autocorrelation is added to the model, the CEO effect vanishes, and the resulting model fits the data substantially better. To ensure that this result did not arise from a computational error, we estimated the model multiple times from different starting

[15] Our dataset differs from that used by Quigley and Graffin (2017) in that we included more years in the panel and did not remove companies with a single CEO from our analyses. While single CEO companies do not provide any information on the CEO effect, they nevertheless contain useful information for estimating the year-to-year performance variations and the persistence of performance over time in multilevel models. As a robustness check, we also estimated our models omitting the single CEO companies with similar results.

[16] We thank a reviewer for pointing this out.

Table 2
Robustness checks and additional mixed effects regressions.

	No trimming of ROA			Returns controlling for Assets			Tobin's Q used as thedependent variable			Alternative formula for the ROA		
	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15
Fixed effects												
Assets				.0521	.0486	.0487						
				(.000600)	(.000707)	(.000680)						
Industry mean	.212	.190	.187	22.08	17.50	17.99	.445	.412	.412	.495	.440	.446
	(.0123)	(.0126)	(.0125)	(1.652)	(1.745)	(.0270)	(.0108)	(.0111)	(.0116)	(.0100)	(.0103)	(.0103)
Intercept	−.569	−.333	−.429	−41.73	−12.59	−15.49	−.176	2.818	2.82	−.471	−.269	−.272
	(.186)	(.186)	(.186)	(16.58)	(16.46)	(.0210)	(1.583)	(1.602)	(.0474)	(.120)	(.116)	(.0439)
Random effects												
Firm	30.2	29.2	23.5	2628767	284650	264401	4251	3265	3260	22.57	14.99	14.3
	(.891)	(.842)	(.923)	(72467)	(7349)	(7086)	(91.13)	(102.1)	(126)	(.525)	(.568)	(.491)
CEO	21.2		14.9	5517623		274480	2193		2.39	15.08		2.72
	(.833)		(.953)	(8151)		(526.6)	(41.30)		(.196)	(.289)		(.0895)
Error variance	136	158	151	604660	1083700	815671	3218	6695	6690	21.64	42.77	41.2
	(.652)	(.839)	(.886)	(2818)	(7585)	(6690)	(15.23)	(80.10)	(70.3)	(.103)	(.484)	(.472)
AR1 (within firm)		.280	.272		.693	.495		.747	.747		.728	.726
		(.00821)	(.00879)		(.00879)	(.00834)		(.00623)	(.00546)		(.00629)	(.00642)
Log likelihood	−113367.6	−112942.1	−112903.1	−231229.7	−229634.6	−229355.6	−155996.0	−151676.2	−151676.2	−87317.8	−83000.2	−82960.56
CEO effect	11.3%		7.8%	38.9%		20.3%	22.7%		.0%	25.4%		4.7%

Note. N = 28,625/28,026 firm-years; 196 three-digit SIC industries, 2,409/2,407 firms, and 5,188/5,136 distinct CEOs for Models 3–6/7–15. The AR1 parameter is a correlation metric. Models 4, 5, 7, 8, 10, 11, 13, and 14 are estimated using Stata's *mixed*. Models 6, 9, 12, and 15 were estimated with our R code. The delta method standard errors are presented in parentheses. The CEO effect is calculated as the ratio of the variance of the CEO level random intercept to the sum of all variance components. ROA and Tobin's Q have been multiplied by 100 to make the variance components easier to interpret.

values and using different optimization algorithms and obtained the same result. One possible explanation for this finding is that because Tobin's Q is a forward-looking measure, investors might already be pricing in a new CEO before he or she steps in. As such, the differences in performance between different CEOs are much less clear with this metric.

Fourth and finally, we used an alternative formulation for ROA. Following prior CEO effect research (Hambrick & Quigley, 2014; Keller et al., 2023; Withers & Fitza, 2017), we defined ROA as the ratio $\frac{net\ income}{total\ assets}$. Specifically, we used income before extraordinary items (as defined in Compustat (ib)) as the numerator. However, it can be argued that this method is not an ideal measure of performance[17] because it focuses only on the returns to shareholders and thus confounds performance with capital structure decisions. In other words, the numerator focuses only on returns to shareholders, but the denominator, total assets, also includes debt that needs to be serviced. An alternate way to calculate ROA could be $\frac{EBIT\ or\ EBITDA}{total\ assets}$, which does not subtract interest from the numerator and thus considers the cash flow available for servicing debt as well as that added to equity or given out as dividends. We expected these two measures to be nearly perfectly correlated, but this was not the case. The variables correlated only at 0.65. A scatter plot indicated a few outliers. After these were eliminated by trimming the first and last 1% from both variables, the correlation remained at just 0.80. When operationalizing this way, the traditional variance decomposition leads to a similar result (Model 1 in Table 1 vs. Model 13 in Table 2). However, the autocorrelation parameter is much higher, and the CEO effect is just half that of our main result (Model 3 in Table 1 vs. Model 15 in Table 2). This is probably because capital structure evolves slowly and is thus more stable over time than the commonly used measures of ROA. Decisions made to take on or reduce debt also likely span multiple CEO tenures, and as such, they do not contribute to CEO effects. Even so, it is notable that we still find a CEO effect, even if it is indeed smaller, when using this setup.

In summary, the CEO effect estimates for different choices of performance measures range from 0% to 20.3%. Importantly, all four robustness checks confirm the result from our main analysis, which shows that ignoring autocorrelations produces an inflated CEO effect. Thus, our main conclusion remains valid regardless of the choice of measures and the accompanying methodological approach.

Replication of Quigley and Hambrick (2015)

We next show that accounting for autocorrelation not only changes the magnitude of the CEO effect but also has an impact on the conclusions that can appropriately be drawn from CEO effect studies. In particular, we replicate the highly cited study conducted by Quigley and Hambrick (2015), which found that the CEO effect has increased over time. This study is important because it tests whether the increased levels of attention given to CEOs in recent years is due to their increased impact on firm performance. The study also has a clear policy implication because if the CEO effect has increased over time, the increases in CEO compensation that we have seen over recent decades could be justified by the increasing influence that CEOs exert over firm performance. The original study did not account for autocorrelation. However, there is an argument to be made that autocorrelation of performance provides an alternative explanation for the finding of an increasing CEO effect over time. It is simply possible that the within-firm variation of performance has increased over time (e.g., due to hypercompetition (Wiggins & Ruefli, 2005)), and because firm performance is autocorrelated, the increased variation is erroneously attributed to the CEO effect. If this alternative explanation is true, then it would imply that the increased level of attention given to CEOs in recent years is only loosely connected to their actual impact on firm performance and might have to do with other factors not yet explored (cf. Meindl, Ehrlich, & Dukerich, 1985).

To test whether autocorrelation can provide an alternative explanation for the findings of Quigley and Hambrick (2015), we applied the autocorrelation model to their datasets.[18] We started by replicating their models in Stata to ensure that the model and data were correct. Thereafter, we estimated the same models with an autocorrelated error term using a modified version of the R code used for the main analysis.[19] The results in Table 3 show a much smaller statistically significant CEO effect between the first (1950–1969) and second periods (1970–1989) but not between the second and the third periods (1990–2009). The results from our replication are consistent with our main result, which shows that ignoring the autocorrelation of performance data inflates the CEO effect and that the year-to-year random variation varies between contexts or over time, and that this variation is incorrectly attributed to the CEO effect by any model that does not consider autocorrelation. This implies that an increasing volatility of firm-level performance, rather than an actual increase in the variation in performance attributable to CEOs, can explain the results of Quigley and Hambrick (2015). Specifically, because these firm-level variations are not fully captured by the year dummies, which estimate the macroeconomic trends shared by all firms, they are incorrectly attributed to the CEO effect unless modeled with an autocorrelation. In contrast to the original study, our analysis suggests that the CEO effect has not significantly increased over time, at least not from the 1970–1989 period to the 1990–2009 period.

Discussion and conclusions

CEO effect studies have been highly influential in the strategic management and leadership literature and have often been used as a basis for examining the impact of a variety of leadership attributes on firm performance. However, efforts to quantify the baseline CEO effect have not been straightforward and have faced both conceptual (Blettner et al., 2012) and statistical challenges (Fitza, 2014, 2017; Hambrick & Quigley, 2014). We have discussed both the challenges in quantifying the CEO effect and developed and validated a modeling approach that addresses these challenges. Using this approach, we found a positive CEO effect, albeit a smaller one than what had been found in recent studies. In addition, our work provides clarification and important qualifications to the prior literature on the CEO effect, which we discuss in the following section.

[17] We thank a reviewer for pointing this out.

[18] We thank Tim Quigley for providing us with the dataset.

[19] Our calculation of the variance percentage differs from that of Quigley and Hambrick (2015). Quigley and Hambrick calculated the total variance using a model without year dummies, and then divided the variance components from a model with year dummies to estimate the total variance. The year effect was estimated as 1 – the sum of the percentages of all the variance components in the model. Hambrick and Quigley then tested the difference by assuming that the variance percentages can be analyzed as R^2 values, which could then be converted to correlations and tested using Fisher's z. While this seems reasonable, it is not a valid test for this purpose; the sampling distribution of a correlation depends only on the population correlation and the sample size, whereas the sampling distribution of the variance components is much more complex, depending on, for example, the level of model complexity. Therefore, we used a different approach for testing and calculated the standard errors (delta method) of all the variance components, and we used these to calculate z tests. The calculation of standard errors in this approach requires that all the effects be calculated from a single model. Therefore, we opted to calculate the full variance using the variance components from the final model, which included dummies. We also estimated the same model without year dummies with nearly identical results. Given that the year dummies explained at most 4% of variance in the corresponding analysis in Hambrick and Quigley (2015), we believe that our results are fully comparable even if the year dummies were not included when calculating the shares of explained variance.

Table 3
Proportion of variance in the ROA explained by different components used in the work of Quigley and Hambrick (2015).

Variance component	Period 11950–1969		Period 21970–1989		Period 31990–2009		Z tests for difference between periods		
							Period1-2	Period2-3	Period1-3
Year-dummies	Included		Included		Included				
Industry (%)	10.7	(3.7)	2.1	(1.0)	2.0	(1.0)			
Company (%)	37.0	(4.0)	18.0	(2.6)	13.6	(2.3)			
CEO (%)	2.6	(1.4)	8.0	(2.0)	10.7	(2.1)	.026	.354	.001
Unexplained (%)	49.7	(3.9)	72.0	(2.3)	73.7	(2.3)			
AR1 (within firm)	66.6	(2.0)	48.1	(1.5)	43.4	(1.6)			

Note. Maximum likelihood estimates. The delta method standard errors are displayed in parentheses.

The CEO effect after accounting for performance variation over time

Our study resolves the apparent paradox that caused Fitza (2017) and Hambrick and Quigley (2014) to reach dramatically different conclusions from the same basic premise that firm performance varies systematically over time. On the one hand, Fitza (2017) is right that the CEO effect can be confounded with the changing firm effect. Our contribution for that work is to show that the two effects can be empirically distinguished. On the other hand, Hambrick and Quigley (2014) are also right in that context matters and the persistence of past performance must be accounted for in CEO effect studies. We demonstrate that their analysis approach confounds the stable firm and CEO effects, and their use of a fixed-effects model misattributes random noise to the CEO effect as a result of the small number of observations per CEO (Fitza, 2014). Thus, even though the issue of persistence highlighted by Fitza (2017) is a serious concern for CEO effect studies, we show that it can be addressed by contextualizing CEOs in the manner suggested by Hambrick and Quigley (2014). However, doing so effectively requires that we can avoid the trap of triple confounding against which Blettner et al. (2012) warned. We show that this is possible by using a multilevel modeling approach that includes autocorrelation to account for the unique context inherited by a new CEO.

As a result, we can establish the existence of a CEO effect beyond mere chance, which strengthens the confidence in those prior empirical studies that have found a CEO effect.[20] However, we also find that the magnitude of the CEO effect (which is approximately 11.5%) is lower than previously understood based on conventional best practices using multilevel models (Quigley & Graffin, 2017). As a direct comparison, our measured CEO effect is approximately half as large as the 21.8% CEO effect found by Quigley and Graffin (2017), who use the same data and analysis technique (i.e., multilevel modeling) but do not model the contextual effects that vary over time.

Additionally, we find that including autocorrelation changes the firm or industry effect to a much lesser degree than the CEO effect. This means that those CEO effect studies that neglect autocorrelation incorrectly attribute autocorrelated performance trends primarily to the CEO effect, which has been a major source of confusion in efforts to quantify the CEO effect. Indeed, some of the prior studies report a CEO effect that is even larger than the firm effect (Hambrick & Quigley, 2014). In contrast, our results indicate that the CEO effect is much lower than the firm effect when modeled properly. While the result showing that the firm effect is larger than the CEO effect seems reasonable given what we know about inertia and the sustainability of performance differences among competing firms (e.g., Hannan & Freeman, 1984; Henderson et al., 2012), the existence of a positive and significant CEO effect also points to the fact that CEOs are able to affect the performance trajectories of the firms that they lead (Hambrick & Quigley, 2014), thus providing justification for those studies that examine the attributes of CEOs and the mechanisms through which CEOs can affect firm performance, whether for better or for worse.

Finally, by failing to replicate the result by Quigley and Hambrick (2015) showing that the CEO effect has substantially increased over recent decades, our study also shows that accounting for autocorrelation not only affects the size of the CEO effect but may also put cast doubt on the conclusions drawn from prior studies. By modeling the evolving firm context, we show that there is reason to be doubtful that the CEO effect has substantially increased over recent years. As such, current CEO effect studies cannot be used to justify the recent increases in attention or compensation that have been awarded to CEOs.

Contextualizing CEOs and assessing whether they "make a difference"

Comparing our study with that of Hambrick and Quigley (2014) raises a subtle but important point concerning how the CEO effect should be thought of conceptually, or in other words, what it specifically is that the CEO effect estimates as a percentage means. Hambrick and Quigley (2014) correctly point out that comparing the performance of a CEO who turned around a company against the long-run average performance of that company may produce a misleading estimate regarding the influence of an individual CEO. In their example, the performance of IBM under Lou Gerstner was −4.4 percentage points below the average level of the company over all years covered by their data but 6 points above the average performance of the company for the two years before his tenure. Based on these data, they argue that comparing a CEO's performance to firm performance under his or her predecessor provides a more valid assessment of CEO performance than a comparison against the long-run average performance of the company does. However, it is not difficult to find counterexamples. Under the first five years of Tim Cook's tenure (2012–2016), Apple has become the world's most valuable company and boasts an average ROA of 18.2%, which is well over (+9.4 points) the company's long-run average of 8.8% from our data. When comparing Tim Cook against Apple's long-run average performance, it is difficult to judge him as nothing less than a successful CEO. However, if we compare his performance to the two last years of Steve Jobs (2010, 2011), when the average ROA was 20.5%, he would be viewed much less favorably (−2.3 points).

As these examples illustrate, comparing an individual CEO against their immediate predecessor is a futile exercise in the context of CEO effect studies. More fundamentally, these examples focus on individual CEOs, whereas the CEO effect captures the variance that can be explained by CEO tenure on average across a large sample. In fact, CEO effect studies using variance decomposition approaches *cannot* be used to assess the skills or quality of different CEOs, but rather to assess whether "CEOs 'matter' or 'make a difference' to the extent that they

[20] The result of a positive CEO effect is immune against Fitza (2014, 2017) arguments concerning chance effects inflating the estimated CEO effect. That is, even though our estimated CEO effect is below the null hypothesis thresholds established by Fitza (2014, 2017), these thresholds do not apply to our results for the following reasons: First, we use multilevel modeling which, as noted by Quigley & Graffin, 2017, avoids the issue caused by the potential of random noise to inflate the CEO effect (Fitza, 2014). Second, Fitza (2017) subsequent CEO effect threshold – when random noise is subject to autocorrelation – was calculated for misspecified models that do not account for autocorrelation, so that critique does not apply to our results because we correctly model using autocorrelation.

exhibit performance tendencies that deviate […] from what would be predicted by their contexts" (Hambrick & Quigley, 2014, p. 474). This question can be formalized through a counterfactual (Durand & Vaara, 2009; Morgan & Winship, 2007) as *how much does firm performance under a specific CEO differ from what would be expected had that specific CEO not been the CEO.* The answer to this question forms the individual CEO causal effect, and CEO effect studies essentially attempt to estimate the variance in this effect across all CEOs.

A causal interpretation of the CEO effect naturally raises concerns about endogeneity. Endogeneity is present to some degree in all non-experimental research; therefore, we should focus on what the most severe sources of endogeneity are and the degree to which they present a problem (Ketokivi & McIntosh, 2017). In the context of CEO effect studies, the most pressing concerns relate to the confounding of CEO effects and firm effects, the persistence of context (Blettner et al., 2012; Fitza, 2014, 2017; Hambrick & Quigley, 2014; Quigley & Graffin, 2017), the selection effect caused by high-performing firms attracting better CEOs, and the selection effect created when CEOs are dismissed based on firm performance. The confounding that can occur between the firm and CEO level is effectively addressed by using multilevel modeling, and we argue that the autocorrelation model addresses the problem of inherited performance. However, the selection effects remain a concern, and we address that next.

To address the potential endogeneity problem caused by good CEOs being attracted by high-performing firms, we need to focus on the counterfactual, or *the level of firm performance had the specific CEO not been the CEO.* We argue that this counterfactual should be constructed based on the *typical* CEO for that firm. Constructing the counterfactual this way provides a solution to the problem that an average firm with a series of superstar CEOs is empirically indistinguishable from a superstar firm with a series of average CEOs. If all CEOs of a firm are superstars, then being a superstar in that firm does not make a difference. As such, framing the question this way addresses the endogeneity concern that high-performing firms attract better CEOs. The endogeneity generated by the dismissal or resignation of a prior CEO based on poor performance still remains a concern, but these are rare events, as the vast majority of CEO departures are attributed to non-performance reasons (Gentry, Harrison, Quigley, & Boivie, 2021) and thus unlikely to severely compromise the variance decomposition results.

Methodological sophistication in strategic management research

Methodologically, we show that the downbeat assessment by Blettner et al. (2012) concerning the inability of current statistical technologies to avoid the "triple confounding of firm, CEO, and time effects" is not warranted. Our approach not only avoids triple confounding but also addresses the concerns that the CEO effect must be gauged against an evolving firm context (Hambrick & Quigley, 2014). Given that "the conclusions drawn within strategic management research are only as solid as the methodological practices that underlie the research" (Ketchen, Boyd, & Bergh, 2008, p. 643), our study raises a fundamental issue regarding the manner in which novel methodological solutions are introduced to management and leadership research.

Both Hambrick and Quigley (2014) and our own work propose approaches for estimating the CEO effect under an evolving firm context. However, while both approaches sound reasonable based on the verbal arguments used to justify their use, they produce dramatically different results (11.5% vs. 38.5%). This raises the question of how we can know which statistical technique to trust or, more generally, how management and leadership researchers can ensure that their statistical techniques are effective at answering the questions that they are modeling. There are complex interdependencies between the methodological choices in CEO effect studies (Blettner et al., 2012), so how do we know which are the right choices to make?

Because knowing what the true parameter values are is generally impossible with any empirical dataset, demonstrations of the validity of new analysis techniques must instead rely on simulated datasets where the estimated population quantities are defined by the researcher and are thus known. Our full Monte Carlo simulation, presented in Appendix A, demonstrates that our analytical approach is effective at correctly estimating the CEO effect and autocorrelation over a range of conditions. In contrast, the empirical examples used in prior studies can only show that a technique produces plausible estimates but cannot inform us about the correctness of those estimates because the true population values are not known. The same applies to the simple simulation reported by Hambrick and Quigley (2014). They demonstrate that individual CEO fixed effects behave as expected when the data are manipulated. However, this unfortunately does not show whether the overall variance (the CEO effect) is estimated correctly in the first place. Overall, our study highlights the need for validating modeling approaches through the application of simulations to demonstrate that the modeling approaches can accurately estimate known population values.

Limitations of variance decomposition studies

It is important to note that CEO effect studies, such as ours, only answer the question of how much CEOs matter to firm performance *on average* while saying nothing about the specific attributes or mechanisms by which CEOs affect firm performance or about whether particular CEOs obtain proper credit for their contributions to firm performance. Therefore, the policy implications that CEO effect studies can have are limited to higher-level questions such as whether the increased levels of attention given to CEOs in recent years is a result of actual CEO impact on firm performance increasing dramatically (Quigley & Hambrick, 2015) or whether this is due to the increasing romanticization of leadership that has occurred in recent years (Meindl et al., 1985). Nevertheless, as suggested by Hambrick and Quigley (2014), CEO effect studies are important for setting a baseline for research that gauges the impact of different kinds of CEOs and their characteristics in different contexts and that may ultimately produce results describing optimal CEO characteristics that can be used to inform CEO selection.

Finally, our study shares a limitation with prior variance decomposition studies because of the Compustat data we use. The dataset contains only relatively large US-based firms that are not in the government or financial sectors. This limits the generalizability of our results to the broader set of geographies and, organizations including financial services and utilities, which we had to exclude to make our result comparable with the prior literature.

Conclusions and further research

The aim of our article was to clarify an apparent paradox in the recent literature regarding the impact of CEOs on firm performance. We demonstrate that a multilevel model that includes autocorrelation makes it possible to avoid the problem of the triple confounding of firm, CEO and time noted by Blettner et al. (2012) and the methodological problems of overestimation raised by Fitza (2014, 2017), while at the same time contextualizing CEOs in the way that Hambrick and Quigley (2014) argue that they should be contextualized.

Our result indicating that the autocorrelation of performance changes the estimates for other effects, such as the CEO effect, potentially has implications for several past and future studies on CEOs. The most direct of these implications concern those studies that apply variance decomposition methods to subsamples to assess whether the CEO effect has increased over time (e.g., Quigley & Hambrick, 2015) or whether it varies between contexts (e.g. Clark, Murphy, & Singer, 2014; Crossland & Hambrick, 2007; Crossland & Hambrick, 2011). Our result stating that ignoring the autocorrelation of performance data inflates the CEO effect suggests that the results of these previous studies may be alternatively explained by the fact that the year-to-year random variation changes between contexts and over time, and that this variation has been incorrectly attributed to the CEO effect by models that do not consider autocorrelation. This alternative explanation is particularly relevant for Quigley and Hambrick's (2015) study, which shows that both the estimated CEO effect (ignoring autocorrelation) and the unexplained year-to-year variance in performance increase over time. An alternative explanation would be that just the year-to-year variance in performance has increased over time, thus causing a spurious increase in the CEO effect estimate. Our study also suggests that studies assessing the effects of specific CEO attributes might be more effective if these attributes were used as independent variable(s) in a multilevel model rather than following Hambrick and Quigley's (2014) recommended approach and using the calculated CEO fixed effects as a dependent variable due to the problems that we have noted in detail.

Admittedly, outside the scope of our research and not examined directly here, the autocorrelation of performance may potentially affect a number of studies in management research areas beyond the CEO effect or CEO studies more generally (e.g. Guo, 2017), as noted by Bergh and colleagues (Bergh, 1993a, 1993b; Bergh & Holbein, 1997) more than twenty years ago. Investigating the possible confounding factors arising from ignoring autocorrelation provides a potentially fruitful avenue for further research. The issues related to autocorrelation raised in this article may be particularly relevant for other variance decomposition studies in the strategic management literature using longitudinal performance data and those that typically do not account for autocorrelation or apply ineffective analysis regarding this issue. The relevant topics examined include strategic groups (Short, Ketchen, Palmer, & Hult, 2007), firm vis-à-vis industry effects in specific contexts (Bamiatzi, Bozos, Cavusgil, & Hult, 2016), corporate and industry interaction effects (Guo, 2017), board chair effect (Withers & Fitza, 2017), ownership effect (Xia & Walker, 2015), and dynamic managerial capabilities (Adner & Helfat, 2003). In addition, the issue of ignoring autocorrelation also raises model misspecification concerns for studies that go beyond simple variance decomposition by including explanatory variables in a multilevel model, such as the corporate effect that occurs in capital allocation competency (Arrfelt, Wiseman, McNamara, & Hult, 2015; Fitza & Tihanyi, 2017). The empirical demonstration presented in this article can hopefully serve as a prompt reminder for upcoming studies of the perils of ignoring autocorrelation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendices A–C. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.leaqua.2023.101733.

References

Adner, R., & Helfat, C. E. (2003). Corporate effects and dynamic managerial capabilities. *Strategic Management Journal, 24*, 1011–1025.

Antonakis, J., Bastardoz, N., & Rönkkö, M. (2021). On ignoring the random effects assumption in multilevel models: review, critique, and recommendations. *Organizational Research Methods, 24*, 443–483.

Arrfelt, M., Wiseman, R. M., McNamara, G., & Hult, G. T. M. (2015). Examining a key corporate role: the influence of capital allocation competency on business unit performance. *Strategic Management Journal, 36*, 1017–1034.

Bamiatzi, V., Bozos, K., Cavusgil, S. T., & Hult, G. T. M. (2016). Revisiting the firm, industry, and country effects on profitability under recessionary and expansion periods: a multilevel analysis. *Strategic Management Journal, 37*, 1448–1471.

Bergh, D. D. (1993a). Don't "waste" your time! The effects of time series errors in management research: the case of ownership concentration and research and development spending. *Journal of Management, 19*, 897–914.

Bergh, D. D. (1993b). Watch the time carefully: the use and misuse of time effects in management research. *Journal of Management, 19*, 683–705.

Bergh, D. D., Aguinis, H., Heavey, C., Ketchen, D. J., Boyd, B. K., Su, P., … Joo, H. (2016). Using meta-analytic structural equation modeling to advance strategic management research: guidelines and an empirical illustration via the strategic leadership-performance relationship. *Strategic Management Journal, 37*, 477–497.

Bergh, D. D., & Holbein, G. F. (1997). Assessment and redirection of longitudinal analysis: demonstration with a study of the diversification and divestiture relationship. *Strategic Management Journal, 18*, 557–571.

Bergh, D. D., Sharp, B. M., Aguinis, H., & Li, M. (2017). Is there a credibility crisis in strategic management research? Evidence on the reproducibility of study findings. *Strategic Organization, 15*, 423–436.

Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchell, W. (2016). Creating repeatable cumulative knowledge in strategic management. *Strategic Management Journal, 37*, 257–261.

Blettner, D. P., Chaddad, F. R., & Bettis, R. A. (2012). The CEO performance effect: statistical issues and a complex fit perspective. *Strategic Management Journal, 33*, 986–999.

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

Bou, J. C., & Satorra, A. (2018). Univariate versus multivariate modeling of panel data: model specification and goodness-of-fit testing. *Organizational Research Methods, 21*, 150–196.

Certo, S. T., Busenbark, J. R., Kalm, M., & LePine, J. A. (2020). Divided we fall: how ratios undermine research in strategic management. *Organizational Research Methods, 23*, 211–237.

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

Chatterjee, A., & Hambrick, D. C. (2007). It's all about me: narcissistic chief executive officers and their effects on company strategy and performance. *Administrative Science Quarterly, 52*, 351–386.

Chiu, S.-C., & Walls, J. L. (2019). Leadership change and corporate social performance: the context of financial distress makes all the difference. *The Leadership Quarterly, 30*, Article 101307.

Clark, J. R., Murphy, C., & Singer, S. J. (2014). When do leaders matter? Ownership, governance and the influence of CEOs on firm performance. *The Leadership Quarterly, 25*, 358–372.

Crossland, C., & Hambrick, D. C. (2007). How national systems differ in their constraints on corporate executives: a study of CEO effects in three countries. *Strategic Management Journal, 28*, 767–789.

Crossland, C., & Hambrick, D. C. (2011). Differences in managerial discretion across countries: how nation-level institutions affect the degree to which CEOs matter. *Strategic Management Journal, 32*, 797–819.

Denrell, J. (2004). Random walks and sustained competitive advantage. *Management Science, 50*, 922–934.

Dishop, C. R., & DeShon, R. P. (2022). A tutorial on Bollen and Brand's approach to modeling dynamics while attending to dynamic panel bias. *Psychological methods, 27*, 1089–1107.

Durand, R., & Vaara, E. (2009). Causation, counterfactuals, and competitive advantage. *Strategic Management Journal, 30*, 1245–1264.

Fama, E. F., & French, K. R. (2000). Forecasting profitability and earnings. *The Journal of Business, 73*, 161–175.

Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics, 82*, 491–518.

Fitza, M. A. (2014). The use of variance decomposition in the investigation of CEO effects: how large must the CEO effect be to rule out chance? *Strategic Management Journal, 35*, 1839–1852.

Fitza, M. A. (2017). How much do CEOs really matter? Reaffirming that the CEO effect is mostly due to chance. *Strategic Management Journal, 38*, 802–811.

Fitza, M. A., & Tihanyi, L. (2017). How much does ownership form matter? *Strategic Management Journal, 38*, 2726–2743.

Fitzsimmons, T. W., & Callan, V. J. (2016). CEO selection: a capital perspective. *The Leadership Quarterly, 27*, 765–787.

Gentry, R. J., Harrison, J. S., Quigley, T. J., & Boivie, S. (2021). A database of CEO turnover and dismissal in S&P 1500 firms, 2000–2018. *Strategic Management Journal, 42*, 968–991.

Guo, G. (2017). Demystifying variance in performance: a longitudinal multilevel perspective. *Strategic Management Journal, 38*, 1327–1342.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological methods, 20*, 102–116.

Hambrick, D. C., & Quigley, T. J. (2014). Toward more accurate contextualization of the CEO effect on firm performance. *Strategic Management Journal, 35*, 473–491.

Hannan, M. T., & Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review, 49*, 149–164.

Henderson, A. D., Raynor, M. E., & Ahmed, M. (2012). How long must a firm be great to rule out chance? Benchmarking sustained superior performance without being fooled by randomness. *Strategic Management Journal, 33*, 387–406.

Henningsen, A., & Toomet, O. (2011). Maxlik: a package for maximum likelihood estimation in R. *Computational Statistics, 26*, 443–458.

Hox, J. J. (2010). *Multilevel analysis: techniques and applications*. New York: Routledge.

Keller, T., Glaum, M., Bausch, A., & Bunz, T. (2023). The "CEO in context" technique revisited: a replication and extension of Hambrick and Quigley (2014). *Strategic Management Journal, 44*, 1111–1138.

Ketchen, D. J., Boyd, B. K., & Bergh, D. D. (2008). Research methodology in strategic management: past accomplishments and future challenges. *Organizational Research Methods, 11*, 643–658.

Ketokivi, M., & McIntosh, C. N. (2017). Addressing the endogeneity dilemma in operations management research: theoretical, empirical, and pragmatic considerations. *Journal of Operations Management, 52*, 1–14.

Leeden, R. V. D., Meijer, E., & Busing, F. M. T. A. (2008). Resampling multilevel models. In J. D. Leeuw, & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 401–433). New York, NY: Springer New York.

Lieberson, S., & O'Connor, J. F. (1972). Leadership and organizational performance: a study of large corporations. *American Sociological Review, 37*, 117–130.

Mackey, A. (2008). The effect of CEOs on firm performance. *Strategic Management Journal, 29*, 1357–1367.

McGahan, A. M., & Porter, M. E. (1997). How much does industry matter, really? *Strategic Management Journal, 18*, 15–30.

McGahan, A. M., & Porter, M. E. (1999). The persistence of shocks to profitability. *Review of Economics and Statistics, 81*, 143–153.

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*, 114–140.

Meindl, J. R., Ehrlich, S. B., & Dukerich, J. M. (1985). The romance of leadership. *Administrative Science Quarterly, 30*, 78–102.

Misangyi, V. F., Elms, H., Greckhamer, T., & Lepine, J. A. (2006). A new perspective on a fundamental debate: a multilevel approach to industry, corporate, and business unit effects. *Strategic Management Journal, 27*, 571–590.

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference: methods and principles for social research* (1st ed.). New York: Cambridge University Press.

Mueller, D. C. (1977). The persistence of profits above the norm. *Economica, 44*, 369–380.

Nadkarni, S., & Herrmann, P. (2010). CEO personality, strategic flexibility, and firm performance: the case of the Indian business process outsourcing industry. *Academy of Management Journal, 53*, 1050–1073.

Quigley, T. J., & Graffin, S. D. (2017). Reaffirming the CEO effect is significant and much larger than chance: a comment on Fitza (2014). *Strategic Management Journal, 38*, 793–801.

Quigley, T. J., & Hambrick, D. C. (2015). Has the "CEO effect" increased in recent decades? A new explanation for the great rise in America's attention to corporate leaders. *Strategic Management Journal, 36*, 821–830.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., Jr., & du Toit, M. (2011). *HLM7: hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International Inc.

Rumelt, R. P. (1991). How much does industry matter? *Strategic Management Journal, 12*, 167–185.

Schepker, D. J., Kim, Y., Patel, P. C., Thatcher, S. M. B., & Campion, M. C. (2017). CEO succession, strategic change, and post-succession performance: a meta-analysis. *The Leadership Quarterly*.

Schmalensee, R. (1985). Do markets differ much? *The American Economic Review, 75*, 341–351.

Semadeni, M., Withers, M. C., & Trevis Certo, S. (2014). The perils of endogeneity and instrumental variables in strategy research: understanding through simulations. *Strategic Management Journal, 35*, 1070–1079.

Short, J. C., Ketchen, D. J., Bennett, N., & du Toit, M. (2006). An examination of firm, industry, and time effects on performance using random coefficients modeling. *Organizational Research Methods, 9*, 259–284.

Short, J. C., Ketchen, D. J., Palmer, T. B., & Hult, G. T. M. (2007). Firm, strategic group, and industry influences on performance. *Strategic Management Journal, 28*, 147–167.

StataCorp LP. (2015). *Stata multilevel mixed-effects reference manual* (Version 14 ed.). College Station, TX: Stata Press.

Wang, L., Hamaker, E., & Bergeman, C. (2012). Investigating inter-individual differences in short-term intra-individual variability. *Psychological Methods, 17*, 567–581.

Waring, G. F. (1996). Industry differences in the persistence of firm-specific returns. *The American Economic Review, 86*, 1253–1265.

Wasserman, N., Anand, B., & Nohria, N. (2010). *When does leadership matter?* Boston Massachusetts: Harvard Business Review Press.

Wernicke, G., Sajko, M., & Boone, C. (2022). How much influence do CEOs have on company actions and outcomes? The example of corporate social responsibility. *Academy of Management Discoveries, 8*, 36–55.

Wiggins, R. R., & Ruefli, T. W. (2002). Competitive advantage: temporal dynamics and the incidence and persistence of superior economic performance. *Organization Science, 13*, 82–105.

Wiggins, R. R., & Ruefli, T. W. (2005). Schumpeter's ghost: is hypercompetition making the best of times shorter? *Strategic Management Journal, 26*, 887–911.

Wiseman, R. M. (2009). On the use and misuse of ratios in strategic management research. *Research methodology in strategy and management*. Emerald Group Publishing Limited.

Withers, M. C., & Fitza, M. A. (2017). Do board chairs matter? The influence of board chairs on firm performance. *Strategic Management Journal, 38*, 1343–1355.

Wooldridge, J. (2009). *Introductory Econometrics: A Modern Approach* (4th ed.). Mason, OH: South Western, Cengage Learning.

Xia, F., & Walker, G. (2015). How much does owner type matter for firm performance? Manufacturing firms in China 1998–2007. *Strategic Management Journal, 36*, 576–585.