Received: 27 September 2015 | Revised: 31 July 2022 | Accepted: 18 August 2022 | Published on: 19 September 2022

DOI: 10.1002/smj.3453

RESEARCH ARTICLE

# The “CEO in context” technique revisited: A replication and extension of Hambrick and Quigley (2014)

**Tobias Keller**[1] | **Martin Glaum**[1] | **Andreas Bausch**[2] | **Thorsten Bunz**[2]

[1]WHU – Otto Beisheim School of Management, Vallendar, Germany

[2]Justus-Liebig-Universität Giessen, Gießen, Germany

**Correspondence**
Martin Glaum, WHU – Otto Beisheim School of Management, Burgplatz 2, 56179 Vallendar, Germany.
Email: martin.glaum@whu.edu; www.whu.edu/facct.

## Abstract

**Research Summary:** Hambrick and Quigley's (2014) “CEO in context” (CiC) technique leads to a much larger CEO effect than traditional ANOVA or multi-level modeling. We replicate H&Q's study, apply their CiC technique to a much more comprehensive U.S. sample, and assess the sensitivity of the model findings to variations in method and data. We generally confirm H&Q's finding of a high CEO effect, but find a smaller industry effect and a larger firm effect in our much larger sample. Applying the CiC technique with adjusted $R^2$s has only a moderate impact on year, industry, and firm effects, but markedly reduces the CEO effect. We also document that CiC model findings are sensitive to sample characteristics, namely firm size and CEO tenure.

**Managerial Summary:** Hambrick and Quigley (2014) introduced a new method to analyze the influence of CEOs on firm performance. The study's empirical analysis focused on large U.S. firms. We replicate the original study and extend it to a much larger, comprehensive sample of U.S. firms that is composed of 33,996 firm-year observations, compared to 4,866 firm-years in the original study. Controlling for the number of variables used in the estimations, the model attributes about a third of the total variance of firm performance (ROA) to the CEO. Further analyses show that the model findings differ for firms of different size and CEO tenure; the larger the firms and the longer the CEO tenures, the smaller tends to be the percentage of variance explained by the CEO.

**KEYWORDS**

CEO effect, CEOs, firm performance, managerial discretion, variance partitioning

## 1 | INTRODUCTION

The impact of top executives on organizational outcomes has been a topic of scientific inquiry for decades (e.g., Crossland & Hambrick, 2007; Hambrick & Finkelstein, 1987; Lieberson & O'Connor, 1972). Some theorists have argued that top executives have a considerable impact on organizational outcomes (Child, 1972; Hambrick & Mason, 1984). In contrast, others have suggested that the impact of executives is negligible because their actions are greatly constrained by external pressures (DiMaggio & Powell, 1983; Hannan & Freeman, 1977).

Various scholars have attempted to analyze the influence of executives empirically by partitioning total variance in firm performance and attributing it to different levels: the industry, the firm, and, finally, the CEO. Earlier studies apply sequential ANOVA (e.g., Lieberson & O'Connor, 1972); later work mostly relies on multilevel modeling (e.g., Crossland & Hambrick, 2011). More recently, Hambrick and Quigley (2014; hereafter H&Q) proposed the "CEO in context," or CiC, technique to more accurately contextualize the impact of CEOs on firm performance. A key criticism of customary variance partitioning methods is that the common use of industry and firm indicators (or grand averages on those levels) leads to misspecifications. To capture these contextual influences more accurately, H&Q introduce benchmark variables that aim to separate the effects of the CEO from those of the firm or industry. Based on data for large U.S. firms for the years 1992–2011, their study yields estimates for the CEO effect that are much higher than those of most previous studies using ANOVA or multilevel modeling.

In this article, we replicate H&Q's (2014) original work; extend the study to a much larger, comprehensive sample of U.S. stock-listed firms; and analyze the sensitivity of the study's inferences to an important change in its method, the use of adjusted $R^2$, and to two important sample characteristics, firm size and CEO tenure. Our work responds to a call for replication studies in strategic management (Bettis, Ethiraj, Gambardella, Helfat, & Mitchell, 2016; Hubbard, Vetter, & Little, 1998). Revisiting and testing the generalizability of extant findings is particularly relevant for CEO effect studies, for several reasons: the importance of top executives is one of the central questions in strategic management; the question continues to attract controversy (e.g., Fitza, 2014, 2017; Quigley & Graffin, 2017); and empirical results vary with measurement techniques and samples. We focus on H&Q (2014) because of the study's importance. It introduces a sophisticated and arguably superior technique to isolate the CEO's impact on firm performance, and it generates a high CEO effect estimate, which, according to the authors, is "much more in line with what would be expected from accepted theory about CEO influence on performance" (p. 475).

We start our work by providing what Bettis, Helfat, and Shaver (2016) call a "narrow replication" of H&Q's study (also see Tsang & Kwan, 1999), using the same data and the same research design as H&Q (2014). Subsequently, we proceed with what Bettis et al. call "quasi-replications"—that is, we assess the sensitivity and robustness of H&Q's findings to variations in the method and the data. More specifically, H&Q focused on a sample of large firms, "roughly the 1,500 largest U.S. corporations" (Hambrick & Quigley, 2014, p. 480). This raises the question of whether their findings also hold for a broader sample that is more representative of the economy as a whole—more specifically, a sample that also includes smaller firms, which tend to have a higher performance variance than larger firms (for a similar line of argument see Chang & Sing, 2000). Thus, following the guidelines developed by Bettis, Helfat, and Shaver (2016), we step by step enlarge H&Q's sample and analyze the effects on the model estimations. Our largest sample, which exploits as completely as possible the currently available data on CEOs of stock-listed firms in the U.S., comprises 33,996 firm-year observations, 1,983 unique firms, and 5,191 CEOs, as compared to H&Q's 4,866 firm-years, 315 firms, and 830 CEOs. Our sample also spans a longer time horizon (50 years vs. H&Q's 20 years), includes more industries, and reflects more diversity in firm size and profitability.

In the second part of our investigation, building on contributions by Fitza (2014) and Quigley and Graffin (2017), we assess the sensitivity of the CiC model to the use of adjusted $R^2$. Fitza (2014) argues that the commonly applied variance decomposition methods overstate the CEO effect because part of the variance in performance that is attributed to the CEO is effectively random. In a reply to Fitza (2014), Quigley and Graffin (2017) argue, inter alia, that the "random chance" element of variance decomposition can be accounted for by basing inferences on adjusted $R^2$, instead of unadjusted $R^2$. Thus, while H&Q (2014) introduced their CiC method with unadjusted $R^2$s, in the second part of our empirical analysis we investigate in detail how using adjusted $R^2$s affects the findings from their model.

In the third part of our investigation, we examine the impact of firm size and CEO tenure on the CiC model findings, following H&Q's own suggestion to assess "when and where CEOs matter most (and least)" (Hambrick & Quigley, 2014, p. 488). We do this mainly by splitting our total sample into quartiles for each of the two characteristics and applying the CiC method to the resulting subsamples. This part of our analysis benefits from our comprehensive dataset, which includes a large and highly diverse set of firms, and a large number of CEOs, among them many CEOs with long tenures.

Our findings can be summarized as follows. First, while we generally confirm H&Q's main finding of a high CEO effect, extending H&Q's original sample does affect the contextual effects of the CiC model. More specifically, H&Q's sample focused on the largest firms in the U.S. stock market. When we add data for smaller and more volatile firms, the CEO effect remains high in all steps of our analysis, but the industry effects are very small and the firm-specific effect becomes more important than in Hambrick and Quigley (2014).

Second, applying H&Q's CiC model with adjusted $R^2$s instead of unadjusted $R^2$s does not have a strong impact on the estimates for the year, industry, and firm effects. In contrast, the CEO effect is markedly lower with adjusted $R^2$s, because of the large number of CEO indicator variables that are introduced in the last stage of the CiC model. However, comparisons of year, industry, firm, and CEO effects across samples are not materially affected by the adjustment of $R^2$s, because the adjustment leaves differences across samples largely intact.

Third, we document that the CiC model findings are sensitive to sample characteristics, namely firm size and CEO tenure. More concretely, the CEO effect is larger, and both firm and industry effects are smaller, in smaller companies than in larger ones. As regards tenure, our investigation reveals that CEO tenure is negatively related to both the CEO and the firm effects. Some of the revealed associations between the sample characteristics and the model findings are conceptually founded and intuitive. Other associations are less intuitive and may be driven by technical aspects of the CiC method, for example, the size-weighting of the industry benchmarks.

## 2 | STUDIES INVESTIGATING THE CEO EFFECT

Over the past decades, numerous studies have attempted to empirically gauge the CEO's influence on firm performance, by using variance partition methods that attribute the total variance of a dependent variable (usually return on assets [ROA]) to factors such as the year, the industry, the firm, and the CEO. We provide an overview of the most important CEO effect studies in Table S1.[1]

In an early paper, Lieberson and O'Connor (1972) find that the CEO accounts for 14.5% of the variance in performance. However, subsequent studies have generated very diverse findings, with CEO effect estimates ranging from about 2% (Bertrand & Schoar, 2003) to more than 40% (Weiner & Mahoney, 1981). Possible reasons for the diverse results could be the different time periods and samples of the studies, but also the different estimation methods (analysis of variance [ANOVA], maximum likelihood estimation, or multilevel modeling [MLM]).

Hambrick and Quigley (2014) criticize another common aspect of traditional CEO effect studies, the use of simple indicator ("dummy") variables to measure the contextual effects of industries and firms:

> Nominal indicators of context do not specify the pertinent, proximal conditions in which individual CEOs are located, which [...] causes substantial blurring of contextual effects and CEO effects. The use of nominal predictors is especially problematic because it treats some of the CEO's own impact as part of the context in which he or she is operating, thus systematically underestimating overall CEO influence. (Hambrick & Quigley, 2014, p. 474).

As we explain in detail below, H&Q instead introduce year-specific benchmarks for the performance of firms' industries and CEO-specific variables for firms' "inherited" health and performance. Applying their refined model to a sample of large U.S. firms for the years from 1992 to 2011 yields a CEO effect of 38.5%, considerably higher than their results from sequential ANOVA (16%) or MLM (20%). In the first part of our own empirical analysis, we replicate the original H&Q study and then investigate whether their findings also hold in a much broader sample that also includes medium-sized and smaller firms.

[1]In a related stream of research, starting with Schmalensee (1985) and Rumelt (1991), studies using variance decomposition investigate the relative importance of industry, business line, and corporate effects. For an overview of earlier studies, see Bowman and Helfat (2001); for more recent studies, see McGahan and Porter (2002), Adner and Helfat (2003), Misangyi, Elms, Greckhamer, and Lepine (2006), and Guo (2017).

In a critical assessment of the CEO effect literature published in parallel to Hambrick and Quigley (2014), Fitza (2014) argues that part of the variance in performance that variance-decomposition studies attribute to the CEO is in fact random. As Fitza further points out, the CEO effect is the more inflated the shorter the CEOs' tenures, and the lower the number of individual CEOs in a sample. In a reply, Quigley and Graffin (2017) argue that the "random chance" element of variance decomposition can be controlled by basing inferences on adjusted $R^2$, instead of unadjusted $R^2$.[2] We take account of this methodological debate by carefully examining, in the second part of our empirical investigation, how using adjusted $R^2$s affects the findings from H&Q's CiC model.

Over the years, several authors have argued that researchers should not content themselves with the question of whether CEOs, in general, have an impact on firm performance, but should also investigate under which sets of conditions CEOs matter more, or less (e.g., Fitza, 2017; Hambrick & Quigley, 2014; Quigley & Graffin, 2017; Wasserman, Anand, & Nohria, 2010). For example, Lieberson and O'Connor (1972) and Wasserman et al. (2010) point out that the magnitude of the CEO effect differs across industries; Crossland and Hambrick (2007, 2011) identify country differences; and the findings of Quigley and Hambrick (2015) suggest that the importance of CEOs has increased over time. In the third part of our own empirical investigation, we examine more closely how firm size and CEO tenure moderate the CiC model findings. We focus on these two factors because Hambrick and Quigley (2014) limited their original study to large U.S. firms and because tenure is closely connected to the CEO effect and plays an important role in the argumentation of Fitza (2014).

## 3 | METHOD

Our study replicates and extends H&Q's 2014 study. H&Q define a set of nested equations to analyze firm performance and to isolate the portion of its variance that is attributable to the CEO. As in most studies on the CEO effect, the performance variable is ROA, calculated as net income divided by total assets for each firm-year.

Like previous researchers, in a first step H&Q use indicator variables for calendar years to capture macroeconomic effects [see Equation (1) below]. Previous studies also used indicator variables to measure the influence of the industry. It is a key insight of H&Q that this approach is imperfect because the firm itself contributes to the mean performance in its industry, and because it assumes that the industry influence on performance is constant over the time horizon of the study. Therefore, H&Q replace the industry dummies with an industry benchmark, calculated as the size-weighted mean ROA per industry in a given year, excluding the focal firm (and hence its CEO).[3] Similar arguments apply at the firm level. That is, just as the firm itself contributes to the mean performance in its industry, each CEO contributes to the firm's average performance across the study's data panel. In addition, H&Q point out that using firm indicator variables for the entire time span of a study not only assumes constancy of the firm effect but also implies that CEOs are assessed against an average firm performance that includes the years beyond their tenures (Hambrick & Quigley, 2014, p. 479). To address these issues, H&Q replace the firm dummies with two variables meant to account for the condition of the company at the start of each CEO's tenure: "inherited

[2]More recently, Fitza (2017) suggests that CEO effect estimates in variance decomposition studies can be affected not only by random noise, but also by the autocorrelation of firm performance.

[3]In accordance with H&Q, when calculating the industry benchmarks we include all firms in Compustat for which the necessary basic data (i.e., net income, total assets) are available.

profitability," proxied by the mean ROA for the 2 years preceding the CEO's tenure, and "inherited health," measured as the company's market-to-book ratio (MTB) divided by the median MTB ratio of the industry excluding the focal firm, at the start of the CEO's tenure. In the fourth and last of the regression equations, in accord with previous studies, H&Q (2014) use indicator variables for individual CEOs to capture their aggregate effect.

The following equations summarize Hambrick and Quigley's (2014) CiC model:

$$ROA_{yiko} = \sum DYEAR_y + \varepsilon^1_{yiko} \tag{1}$$

$$ROA_{yiko} = \sum DYEAR_y + I - BENCH_{yik} + \varepsilon^2_{yiko} \tag{2}$$

$$ROA_{yiko} = \sum DYEAR_y + I - BENCH_{yik} + INPROF_{ko} + INHEALTH_{ko} + \varepsilon^3_{yiko} \tag{3}$$

$$\varepsilon^3_{yiko} = \sum DCEO_o + \varepsilon^4_{yiko}, \tag{4}$$

where subscripts $y$, $i$, $k$, and $o$ indicate the year, industry, firm, and CEO, respectively.

The year effect is estimated as the $R^2$ from Equation (1). The industry and firm effects are estimated as the incremental $R^2$s of Equations (2) and (3) respectively. The CEO effect is calculated in two steps[4]: first, by calculating the $R^2$ from Equation (4), which indicates how much of the remaining unexplained performance variance in Equation (3) is explained by the CEO indicator variables; and second, by multiplying this $R^2$ value by the percentage of unexplained variance from Equation (3), which yields the incremental total variance explained by the CEO indicator variables.[5]

In the first part of our own analysis, we "narrowly" replicate and then extend H&Q's 2014 study by progressively enlarging their sample in five steps, which we describe in detail in Section 4. As we enlarge the sample, we compare year, industry, firm, and CEO effects across the steps. We use Fisher's $z$-transformation to estimate 95% confidence intervals for all effect estimates, and in our discussion, we focus on step-to-step changes that are clearly statistically distinguishable because the confidence intervals of the two effects do not overlap.[6]

---

[4]According to H&Q (2014, p. 482), "using the residuals from Equation (3) as the dependent variable rather than adding the CEO dummy variables as further explanatory variables to Equation (3), assures that the fixed effect coefficients for individual CEOs [...] can be meaningfully interpreted as each CEO's net effect after completely controlling for contextual factors." They further explain, "adding the CEO dummies to the full model generates the same amount of variance explained, or R-squared, but yields less stable estimates for the individual CEOs."

[5]We follow H&Q (2014) and winsorize continuous variables in all regression estimations at the top and bottom 2.5 percentiles. To take account of potential serial correlation in the panel data, H&Q use generalized estimating equations (GEE). Their specific GEE model assumes a normal distribution and a first-order autoregressive structure for the within-panel correlation (for details of GEE model estimation, see Hardin & Hilbe, 2013). In our own empirical investigation, we mainly use OLS estimation. The reason for this is that a major part of our analyses is concerned with the use of adjusted $R^2$, and we are not aware of an established method to compute adjusted $R^2$s for GEE.

[6]We are grateful to one of the reviewers for suggesting the use of Fisher's z-transformation. Thus, following Quigley and Hambrick (2015), we interpret year, industry, firm, and CEO effects as partial $R^2$s and use Fisher's z-transformation to estimate confidence intervals. We use the Stata command *corrcii* with the *fisher* option to calculate Fisher's z (see Cox, 2008). In line with general *SMJ* policy, we use the confidence intervals descriptively and not to identify specific cutoff levels of statistical significance. Focusing on confidence intervals that do not overlap is a conservative approach, because some differences in $R^2$s across steps could be "statistically significant" in the traditional sense even when confidence intervals overlap (for details see Schenker & Gentleman, 2001). For detailed discussions on the estimation of confidence intervals of $R^2$, see Olkin and Finn (1995) and Zou (2007).

In the second part of our analysis, building on the work of Fitza (2014) and Quigley and Graffin (2017), we investigate how using adjusted, rather than unadjusted, $R^2$s affects the model's findings. In the third and final part of our analysis, we investigate how sensitive the H&Q model is to important sample characteristics, namely firm size and CEO tenure. We do this mainly by splitting our sample into quartiles and running the model separately for the resulting subsamples. We again use confidence intervals based on Fisher's $z$-transformation to assess differences between the quartiles.

## 4 | DATA

The starting point of our analysis is H&Q's data on CEOs of large stock-listed U.S. corporations for the years 1992 to 2011 from the Execucomp database, matched to key financial data for these firms from the Compustat database: 4,866 firm-year observations, pertaining to 830 CEOs and 315 unique firms in 44 industries (4-digit SIC), with at least four firms in each industry.[7] We follow H&Q (2014) and exclude financial institutions, public sector entities, and unclassified firms, as well as firms with only one CEO and CEOs who stayed only 1 year or less.[8]

We then extend the sample in steps. In Step 1, we focus on the same firms as H&Q (2014) and use the longest time series for which we can find data in Execucomp and Compustat (1971–2019). In Step 2, we add data for all further available firms in the same industries as in H&Q (2014), and in Step 3 we extend the sample to additional industries, that is, we make full use of the available data in Execucomp and Compustat.[9] In Step 4, we broaden our sample further by adding data from the Refinitiv Eikon database (formerly ThompsonReuters Eikon), which provides both CEO data and financial data (i.e., it includes the formerly separate databases Datastream and Worldscope). At this stage, the sample has 29,318 firm-year observations and comprises data for 4,515 CEOs, 1,725 unique firms, and 164 industries.

While we thus have markedly expanded our sample in comparison to H&Q (2014), we are still constrained by requiring at least four benchmark firms in each industry, measured at the four-digit SIC level, the finest available industry classification.[10] We could attain a still broader

[7]We thank Tim Quigley for having made available to us the final dataset of H&Q's (2014) study as well as his Stata code for estimating the variance partitioning.

[8]In line with H&Q (2014), we use the Execucomp variable "COPEROL" to identify the CEOs.

[9]Here and in what follows, in order for firms to be included in our sample we require them to have nonnegative equity and at least US$ 1 m in assets.

[10]A factor limiting the estimation of industry effects in variance decomposition studies in general is the quality of the data, that is, the firms' industry classification provided in databanks such as Compustat and Refinitiv Eikon. Like H&Q (2014) and previous researchers, we assign all firms in our samples to the industries of their respective primary product segments. However, the assignment of firms to industries can be complicated and may lead to errors in the databases (Jacobs & O'Neill, 2003; Kahle & Walkling, 1996). Moreover, firms' product portfolios change over time, so using static SIC industry affiliations, as is common in variance partitioning studies, is problematic, especially over longer horizons. Also problematic is the industry classification of multisegment firms, where the data providers assign "primary codes" to indicate the firms' major areas of activities. For all of these reasons, industry effect estimates from variance partitioning studies need to be interpreted with care.

representation of firms and CEOs if we conducted the analysis on the broader three-digit SIC level instead.[11] Thus, in an intermediate step, we examine whether the CiC model estimations are sensitive to the granularity of the industry categorization by estimating the model for the three-digit SIC level, using the same firm-year observations as in Step 4. Finding that the estimation results are largely the same as those based on four-digit SIC industry benchmarks, in the fifth and last step we include all firms for which CEO and financial data are available and for which we can find at least four firms in the same industry, defined at SIC level 3. The resulting sample, the largest possible sample to which we can apply the H&Q (2014) CiC model, has 33,996 firm-year observations, with 5,191 individual CEOs for 1,983 firms in 140 SIC level-3 industries.

## 5 | FINDINGS

### 5.1 | Part 1: Narrow replication of H&Q (2014) and stepwise sample extensions

The findings of the first part of our analysis are summarized in Table 1. The first column of Table 1 presents H&Q's (2014) original estimation. The next two columns present our narrow replication of their estimation, and the subsequent columns present the findings from quasi-replications in which we extend our sample step by step. Panel A describes the samples; Panel B shows the estimation results for the performance variance components. We present point estimates for all year, industry, firm, and CEO effects, and we present confidence intervals based on Fisher's $z$-transformations for all OLS estimations. Panel C presents descriptive statistics on sales, total assets, ROA, and CEO tenure for the samples.

We begin by replicating H&Q's original estimation as closely as possible, using H&Q's set of sample firms and method (GEE), but our own data and calculations (Stata code). As the results in column 2 of Table 1 document, our estimation yields nearly the same results as those of H&Q—the industry effect is 7.22% (vs. 6.9% in H&Q [2014]), the firm effect is 11.51% (vs. 12.1%), and the CEO effect is 39.01% (vs. 38.5%). The minor differences between our point estimates and those of H&Q may be related to the industry benchmarks, that is, $I-BENCH_{yik}$ in Equations (2) and (3). To calculate these benchmarks, H&Q use all firms available in Compustat, and we follow their procedure. But it is likely that the set of firms covered by Compustat, and the industry affiliation of some firms, have changed over time.[12]

In the second column, we provide the findings from OLS estimations with the same data, that is, the original H&Q sample. These findings are very similar to those with GEE. In fact, the firm effect is now slightly higher than with GEE, and the CEO effect slightly lower, bringing the estimates yet closer to the original H&Q (2014) results. Having thus established that there is

[11]Weiner (2005) has shown that the degree of homogeneity of firms belonging to the same industry according to Compustat or Datastream Worldscope (now Refinitiv Eikon) is very similar for the four-digit and the three-digit SIC classification levels. Weiner argues that narrow industry classifications, such as the four-digit SIC level, lead to small populations, so that results can easily be biased through outliers. Broader industry definitions do not suffer as much from outliers but tend to become inhomogeneous. Examining this trade-off further, Weiner concludes that the three-digit SIC level yields the most accurate value predictions. For a more formal treatment of the trade-off between the efficiency gains from pooling and the possible bias resulting from heterogeneity see Wang, Zhang, and Paap (2019).

[12]Using Tim Quigley's final dataset and his Stata code yields exactly the results reported by H&Q (2014). Furthermore, we also get exactly the same results when we use his dataset and our code.

**TABLE 1** Narrow replication of H&Q (2014) and stepwise sample extensions

| | Original results | Narrow replication | | Quasi-replications: Stepwise sample extensions | | | | |
|---|---|---|---|---|---|---|---|---|
| | H&Q 2014 | GEE | OLS | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 |
| *Panel A: Sample characteristics* | | | | | | | | |
| Time period | 1992–2011 | 1992–2011 | 1992–2011 | 1971–2019 | 1970–2019 | 1970–2019 | 1970–2019 | 1969–2019 |
| No. of firm-years | 4,866 | 4,866 | 4,866 | 7,528 | 11,588 | 17,692 | 29,318 | 33,996 |
| No. of unique firms | 315 | 315 | 315 | 315 | 570 | 890 | 1,725 | 1,983 |
| No. of CEOs | 830 | 830 | 830 | 1,007 | 1,667 | 2,569 | 4,515 | 5,191 |
| No. of years | 20 | 20 | 20 | 49 | 50 | 50 | 50 | 51 |
| SIC level | SIC 4 | SIC 4 | SIC 4 | SIC 4 | SIC 4 | SIC 4 | SIC 4 | SIC 3 |
| No. of industries | 44 | 44 | 44 | 44 | 44 | 100 | 164 | 140 |
| Estimation method | GEE | GEE | OLS | OLS | OLS | OLS | OLS | OLS |
| Database | Execucomp/ Compustat | Execucomp/ Compustat | Execucomp/ Compustat | Execucomp/ Compustat | Execucomp/ Compustat | Execucomp/ Compustat | Execucomp/ Compustat and Refinitiv Eikon | Execucomp/ Compustat and Refinitiv Eikon |
| *Panel B: Estimation results* | | | | | | | | |
| Year effect | 2.5 | 2.56 | 2.59 | 3.07 | 2.75 | 2.58 | 2.47 | 2.57 |
| | | | 1.78–3.54 | 2.35–3.88 | 2.20–3.37 | 2.14–3.06 | 2.14–2.84 | 2.25–2.91 |
| Industry effect | 6.9 | 7.22 | 7.39 | 7.29 | 3.49 | 2.91 | 0.45 | 0.37 |
| | | | 6.02–8.85 | 6.19–8.45 | 2.86–4.17 | 2.44–3.41 | 0.31–0.62 | 0.26–0.51 |
| Firm effect | 12.1 | 11.51 | 12.18 | 11.27 | 14.37 | 14.45 | 19.22 | 18.65 |
| | | | 10.49–13.94 | 9.95–12.64 | 13.20–15.57 | 13.50–15.41 | 18.41–20.03 | 17.90–19.39 |
| CEO effect | 38.5 | 39.01 | 38.18 | 34.92 | 35.73 | 36.68 | 40.72 | 39.92 |
| | | | 36.01–40.30 | 33.17–36.64 | 34.32–37.12 | 35.54–37.80 | 39.85–41.58 | 39.11–40.72 |
| Unexplained variance | 40.0 | 39.70 | 39.66 | 43.45 | 43.66 | 43.38 | 37.14 | 38.49 |

**TABLE 1** (Continued)

| | | Original results | Narrow replication | | Quasi-replications: Stepwise sample extensions | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | H&Q 2014 | GEE | OLS | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 |
| *Panel C: Descriptive statistics for samples* | | | | | | | | | |
| Sales (m US$) | Mean | 10,355.1 | 10,355.1 | 10,355.1 | 10,729.4 | 7,868.1 | 7,439.4 | 5,082.0 | 4,910.6 |
| | Median | 2,532.1 | 2,532.1 | 2,532.1 | 2,576.7 | 1,619.0 | 1,649.0 | 821.3 | 781.0 |
| | *SD* | 30,835.8 | 30,835.8 | 30,835.8 | 32.660.4 | 26,673.0 | 23,899.8 | 19,107.3 | 18,143.7 |
| Tot assets (m US$) | Mean | 12,745.4 | 12,745.4 | 12,745.4 | 14,262.8 | 10,742.1 | 9,319.7 | 6,344.9 | 6,090.4 |
| | Median | 3,337.3 | 3,337.3 | 3,337.3 | 3,575.7 | 2,234.2 | 1,857.5 | 840.2 | 787.8 |
| | *SD* | 32,540.1 | 32,540.1 | 32,540.1 | 35,616.8 | 29,801.6 | 26,944.9 | 21,709.7 | 20,759.6 |
| ROA (%) | Mean | 0.04 | 0.04 | 0.04 | 0.04 | 0.03 | 0.04 | 0.00 | 0.00 |
| | Median | 0.04 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.04 | 0.04 |
| | *SD* | 0.13 | 0.13 | 0.13 | 0.12 | 0.18 | 0.16 | 0.32 | 0.40 |
| CEO tenure (years) | Mean | 5.86 | 5.86 | 5.86 | 7.48 | 6.95 | 6.89 | 6.35 | 6.38 |
| | Median | 5.00 | 5.00 | 5.00 | 7.00 | 6.00 | 6.00 | 5.00 | 5.00 |
| | *SD* | 3.17 | 3.17 | 3.17 | 4.43 | 4.53 | 4.61 | 4.49 | 4.51 |

*Note*: The table presents results from H&Q CiC model estimations for different samples. The first column of the table presents the original estimation from H&Q (2014). The second column presents a narrow replication of H&Q (2014), using the same data and the same estimation method (GEE). The third column presents a replication again using the same data, but OLS estimation instead of GEE. The subsequent columns present findings from our analyses based on extended samples using OLS. The largest sample (Step 5) comprises 33,996 firm-year observations. Panel A describes the samples used in the consecutive estimations (time period, number of firm-years, unique firms, CEOs, industries), the databases, and the estimation method. Panel B reports the results of the model estimations, that is, point estimates and, for the OLS estimations, 95% confidence intervals based on Fisher's $z$-transformations; all effects are calculated using unadjusted $R^2$s, as in H&Q (2014). Panel C presents key descriptive statistics for the samples (mean, median, and standard deviation of sales, total assets, ROA, CEO tenure).

no meaningful difference between the GEE and the OLS estimates, we continue to use OLS in the following parts of our investigation.[13]

In Step 1, the first of our sample extensions, we confine our analysis to the same firms (and hence the same industries) as H&Q (2014), but include all firm-year observations that are currently available for those firms in Execucomp and Compustat, the databases used by H&Q. As Panel A of Table 1 shows, we now use 7,528 firm-year observations, pertaining to 1,007 CEOs who were active between 1971 and 2019.[14] One effect of the longer time horizon is that the average CEO tenure is longer than in H&Q (2014), because we do not artificially cut longer tenures at the lower and upper limits of 1992 and 2011. While the mean (median) CEO tenure in H&Q's original sample is 5.86 years (5 years), it now is 7.48 years (7 years). As Panel B of Table 1 reveals, the estimation results for the year, the industry, and the firm effects in Step 1 are quite similar to those in Step 0 and those in H&Q (2014). However, we now find a somewhat lower CEO effect (34.92%) and a higher unexplained residual variance (43.45%).[15]

In Step 2, we still restrict the sample to the same 44 industries as in H&Q (2014), but now add data for all additional firms available in Execucomp and Compustat. The new sample comprises 11,588 firm-year observations for 570 unique firms and 1,667 individual CEOs. Panel C of Table 1 reveals that we add predominantly smaller firms.[16] The mean (median) sales of the sample firms now is US$ 7.9 bn (US$ 1.6 bn), whereas it was more than US$ 10 bn (US$ 2.5 bn) for the samples used in H&Q (2014) and in Step 1 of our analysis. The mean ROA of the new sample is a little lower than in Step 1 (0.03% vs. 0.04%), and the standard deviation of ROA increases from 0.12% in Step 1 (0.13% in H&Q [2014]) to 0.18% in the new sample, indicating that the smaller firms that we add in Step 2 display more heterogeneous performances than the larger ones in H&Q's sample. In consequence, the industry factors explain less variance and the firm-specific factors explain more; that is, the industry effect becomes weaker (3.49%), while the firm effect becomes stronger (14.37%).[17] The CEO effect (35.73%) and the remaining, unexplained part of ROA variance (43.66%) are similar to those in Step 1.

In Step 3, we extend the sample to additional industries. In other words, we now use all data currently available from Execucomp and Compustat. At this stage, the sample has 17,692 observations covering 100 industries, more than twice as many industries as in H&Q's study, and there are 2,569 individual CEOs, compared to 1,667 in the previous step and 830 in H&Q's study. The firms in the newly added industries do not differ much from the Step 2 firms in size, performance, or CEO tenure (see Panel C), and the year, industry, and firm effect estimates are similar to those in Step 2. The CEO effect now explains 36.68% of ROA variance, about one percentage point more than in Step 2.

[13]To investigate further whether the regression technique (OLS vs. GEE) affects our findings, we also estimate the H&Q model for all stepwise sample extensions in Table 1 using GEE. The results of these untabulated estimations are very similar to those of the tabulated OLS estimations, leading us to conclude that our findings and inferences do not depend on the choice of the regression technique.

[14]Execucomp's coverage starts in 1992. Hence, firm-year observations for previous years pertain to CEOs who were serving in 1992 but started their tenure earlier. In some cases, Execucomp reports the earlier starting dates; in all other cases we hand-collected this information, wherever possible. We acknowledge that in cases where we cannot identify an earlier starting date, the CEO tenure data are censored. Our CEO tenure data are also right-censored to some degree, because some of the CEOs serving in 2019 continue in their positions in 2020 and beyond.

[15]Given the reported 95% Fisher's z confidence intervals, the intervals for the CEO effects in Step 0 and Step 1 clearly overlap. They would still (just) overlap, if we were instead using a 90% confidence level; the lower and upper limits would then be 36.37% to 39.97% for Step 0 and 33.46% to 36.37% for Step 1.

[16]Execucomp appears to have extended its coverage among smaller firms since H&Q's study.

[17]In both cases, the confidence intervals do not overlap with those in Step 1.

In the next step of the analysis (Step 4), we add data for further firms and industries from Refinitiv Eikon.[18] The sample again increases substantially, now to 29,318 firm-year observations (+65.7% compared to the previous step), 1,725 unique firms (+93.8%), and 4,515 individual CEOs (+75.7%). In Panel C of Table 1, we see that the mean (median) sales of the new sample is now only US$ 5.1 bn (US$ 0.8 bn), the mean (median) ROA is zero (4%), and the standard deviation of ROA is now 32%. In other words, once again, the additional firms are, on average, smaller, and they exhibit a much higher performance variance. Accordingly, the industry effect is now much weaker—so weak that it almost disappears (0.45%). A much higher portion of the variance is explained by the firm effect (19.22%)—and the CEO effect is now even higher than in H&Q's original study (40.72%).[19] The differences between the industry, firm, and CEO effects and the corresponding effects in the previous step are all statistically meaningful, as indicated by the nonoverlapping confidence intervals.[20]

In the fifth and last step, we further increase our sample by using SIC-3 instead of SIC-4 industry codes, as this gives us more observations with at least four firms per industry. Using SIC-3 industry benchmarks leads to similar but more generalizable conclusions than using the more specific and thus more restrictive SIC-4 industry benchmarks.[21] The descriptive statistics in Panel C of Table 1 document that we again add somewhat smaller and more volatile firms. However, Panel B reveals that the estimation results do not change substantially. More precisely, in the largest possible sample to which we can apply the H&Q (2014) CiC model, the year effect is 2.57%, the industry effect is 0.37%, the firm effect is 18.65%, and the CEO effect is 39.92%, slightly lower than the CEO effect estimate in Step 4 and slightly higher than the estimate in H&Q's original study.

[18]We use Execucomp as our primary data source for CEO data; that is, we add data for CEOs from the Refinitiv Eikon database only if the required data are not available from Execucomp. Compustat is our primary data source for financial data, and we add financial data from Refinitiv Eikon only for firm-years for which they are not available from Compustat. The "bottleneck" for the availability of data is usually the CEO data (i.e., Execucomp), not the financial data.

[19]This difference is, however, not pronounced, because the confidence interval of the CEO effect in Step 4 overlaps with the confidence interval of our CEO effect estimate for the original H&Q (2014) sample; see column 3 of Table 1.

[20]In Step 4, the results may be influenced by several changes in the data. First, in this step, we add data from Refinitiv Eikon for additional firms in the 100 industries that were already included in Step 3, as well as data for firms in 64 further industries. In an untabulated test, we restrict the new sample to the 100 industries from Step 3 and find that the differences are not driven by the new industries added in Step 4. Second, the new data not only increase the sample size, but also affect the industry benchmarks, because their calculation now relies on the combined Compustat and Refinitiv Eikon firm-years. To check whether the changes in the benchmark affect the findings, we also reestimate Step 3 using benchmarks that include the Refinitiv Eikon data and find that this has no material effect on the results. In a further test, we rerun Step 4 using only Compustat data, again with very similar results. We conclude that the lower industry effects in Step 4 are not the result of the broader benchmark but are caused by adding firms that differ from their industry peers.

[21]In an untabulated intermediate step, we examine whether the CiC model estimations are sensitive to the granularity of the industry categorization, by estimating the model for the three-digit SIC level using the same firm-year observations as in Step 4. We now distinguish 125 SIC-3 industries, instead of 164 SIC-4 industries. The findings are qualitatively very similar to those reported for Step 4 in Table 1. The explanatory power of the broader SIC-3 industry benchmark is somewhat lower than that of the more specific level-4 benchmarks, albeit on a generally very low level (0.25% vs. 0.45%). The year effect is independent of the industry benchmarks, and the remaining effects are almost identical to those in Step 4. Consequently, we continue to use SIC-3 industry benchmarks in the third part of our analysis. However, we always check whether our findings for SIC level 3 benchmarks also hold when we use the four-digit SIC definition. While using SIC level 4 benchmarks generally reduces the sample size, the estimation results are very similar and inferences remain unchanged.

To sum up, our study confirms H&Q's main finding of a high CEO effect on firm performance, in the narrow replication of H&Q (2014) as well as in our stepwise sample extensions. However, H&Q's original study focused on the largest firms in the U.S. stock market, and when we add data for smaller firms to their sample, we arrive at much smaller estimates for the industry effect and higher estimates for the firm effect.[22]

## 5.2 | Part 2: Sensitivity of H&Q (2014) model findings to method—Adjustment of $R^2$

In what follows, we examine in detail how using adjusted, rather than unadjusted, $R^2$s affects the H&Q model findings. The coefficient of determination, $R^2$, that is, the proportion of variance in the dependent variable that is explained by a regression equation, increases when independent variables are added to a model, even if the increase in the explained variance is spurious and happens only by chance. In contrast, the adjusted $R^2$ takes the number of explanatory variables (in relation to the sample size) into account and increases only if a new variable improves the $R^2$ more than would be expected by pure chance (see, e.g., Kennedy, 2008).

H&Q (2014) introduced their CiC method with unadjusted $R^2$s. There are several possible reasons for this. First, most previous studies of the CEO effect had not used adjusted $R^2$. Second, unadjusted $R^2$s directly represent the percentage of the variance explained in a regression, facilitating the interpretation of the (incremental) $R^2$s and the associated year, industry, firm, and CEO effects in the nested CiC model structure. Third, H&Q used generalized estimating equations (GEE) to estimate their CiC model, and while there are procedures to calculate (pseudo-) $R^2$s for GEE models (e.g., Ballinger, 2004; Hardin & Hilbe, 2013; Zhang, 2017), no established method exists for adjusting $R^2$s from GEE.

Using adjusted instead of unadjusted $R^2$s affects the year, industry, firm, and CEO effects very differently. In Equation (1), the year effect is estimated with the help of $y - 1$ year indicator variables, that is, 19 variables in H&Q (2014) and up to 50-year-indicator variables in our extended analysis. These numbers are small in relation to the sample sizes—4,866 observations in H&Q's study and 33,996 observations in our most comprehensive sample—and the effect of using adjusted instead of unadjusted $R^2$s will thus be modest. In Equation (2), the industry benchmarks are added to the CiC model. There are $i$ industries. Given that the industry benchmarks are year-dependent, in a traditional variance partitioning model one would use $iy - 1$ industry-year indicator variables. However, in the CiC model, only one additional variable, *I-BENCH*, is added, which means that the impact of using adjusted $R^2$ on the industry effect will

[22]A possible concern for our analysis could be that our findings are driven by very small firms that are economically not meaningful. This concern could apply, in particular, to Step 4, in which we add data from the Refinitiv Eikon database. However, while the firms that we add in Step 4 tend to be smaller than the firms composing the sample in Step 3, the average firm added in Step 4 still has US$ 1.8 billion in assets and US$ 1.5 billion in sales (medians: US$ 133.3 million and US$ 134.3 million, respectively). We also rerun our estimations with larger size filters—total assets of US$ 5 million or US$ 10 million, instead of the original US$ 1 million—and we still observe industry, firm, and CEO effects in Step 4 that are clearly different from those in Step 3 (i.e., confidence intervals do not overlap). To ensure that our findings are not unduly influenced by outliers, we also rerun the estimations with stricter winsorizing, at the top and bottom 5 percentiles rather than at H&Q's top and bottom 2.5 percentiles. Again, our results remain qualitatively unchanged.

**TABLE 2** Narrow replication of H&Q (2014) and stepwise sample extensions: Estimations based on unadjusted $R^2$s and adjusted $R^2$s

| | Narrow replication | Quasi-replications: Step-wise sample extensions | | | | |
|---|---|---|---|---|---|---|
| | | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 |
| *Panel A: Estimation results based on unadjusted $R^2$s* | | | | | | |
| Year effect | 2.59 | 3.07 | 2.85 | 2.58 | 2.47 | 2.57 |
| | 1.78–3.54 | 2.35–3.88 | 2.28–3.47 | 2.14–3.06 | 2.14–2.84 | 2.25–2.91 |
| Industry effect | 7.39 | 7.29 | 3.49 | 2.91 | 0.45 | 0.37 |
| | 6.02–8.85 | 6.19–8.45 | 2.86–4.17 | 2.44–3.41 | 0.31–0.62 | 0.26–0.51 |
| Firm effect | 12.18 | 11.27 | 15.13 | 14.45 | 19.22 | 18.65 |
| | 10.49–13.94 | 9.95–12.64 | 13.94–16.34 | 13.50–15.41 | 18.41–20.03 | 17.90–19.39 |
| CEO effect | 38.18 | 34.92 | 35.00 | 36.68 | 40.72 | 39.92 |
| | 36.01–40.30 | 33.17–36.64 | 33.59–36.39 | 35.54–37.80 | 39.85–41.58 | 39.11–40.72 |
| Unexplained variance | 39.66 | 43.45 | 43.53 | 43.38 | 37.14 | 38.49 |
| *Panel B: Estimation results based on adjusted $R^2$s* | | | | | | |
| Year effect | 2.21 | 2.45 | 2.43 | 2.31 | 2.31 | 2.43 |
| | 1.46–3.09 | 1.81–3.18 | 1.91–3.02 | 1.89–2.76 | 1.98–2.66 | 2.11–2.76 |
| Industry effect | 7.40 | 7.32 | 3.49 | 2.91 | 0.45 | 0.37 |
| | 6.03–8.86 | 6.23–8.49 | 2.86–4.18 | 2.44–3.42 | 0.31–0.62 | 0.25–0.51 |
| Firm effect | 12.20 | 11.32 | 15.18 | 14.48 | 19.25 | 18.67 |
| | 10.51–13.95 | 10.00–12.69 | 13.99–16.39 | 13.53–15.45 | 18.44–20.06 | 17.92–19.42 |
| CEO effect | 30.17 | 28.41 | 27.81 | 29.4 | 34.02 | 33.04 |
| | 28.00–32.31 | 26.69–30.13 | 26.42–29.19 | 28.27–30.52 | 33.13–34.90 | 32.22–33.85 |
| Unexplained variance | 48.02 | 50.50 | 51.09 | 50.90 | 43.97 | 45.49 |
| *Panel C: Differences between point estimation results based on unadjusted $R^2$s and point estimation results based on adjusted $R^2$s* | | | | | | |
| Year effect | 0.38 | 0.62 | 0.42 | 0.27 | 0.16 | 0.14 |
| Industry effect | −0.01 | −0.03 | 0.00 | 0.00 | 0.00 | 0.00 |
| Firm effect | −0.02 | −0.05 | −0.05 | −0.03 | −0.03 | −0.02 |
| CEO effect | 8.01 | 6.51 | 7.19 | 7.28 | 6.70 | 6.88 |
| Unexplained variance | −8.36 | −7.05 | −7.56 | −7.52 | −6.83 | −7.00 |

*Note*: The table presents results from H&Q CiC model estimations, using OLS estimation and unadjusted $R^2$s and adjusted $R^2$s, for different samples. The first column presents estimations for the original H&Q (2014) sample; the subsequent columns present findings from our analyses based on extended samples. The largest sample (Step 5) comprises 33,996 firm-year observations; for descriptive statistics of the various samples, see Table 1. Panel A presents model findings based on unadjusted $R^2$s, Panel B model presents findings based on adjusted $R^2$s. Both panels present point estimates for year, industry, firm, and CEO effects, and the respective 95% confidence intervals based on Fisher's $z$-transformations. Panel C presents the differences between the point estimates in Panel A (based on unadjusted $R^2$s) and Panel B (based on adjusted $R^2$s).

be very small.[23] The same holds for the next step, which adds two proxy variables for firms' inherited profitability and health: the impact of adjusting $R^2$ on the firm effect will also be small. The situation is different in Equation (4), where the residuals from Equation (3) are regressed on a set of $(o - 1)$ CEO dummy variables. Here, the number of variables is very high in relation to the sample sizes. Hence, adjusting $R^2$ will have a pronounced impact on the CEO effect.

In Table 2 and Figure 1, we present the CiC model findings for H&Q's original (2014) sample and our five sample extensions, using both unadjusted $R^2$s (Panel A) and adjusted $R^2$s (Panel B).[24] As shown in Table 1, we present point estimates and confidence intervals for the effects, using Fisher's $z$-transformation. Panel C of Table 2 presents the differences between the point estimates. Figure 1 graphically summarizes the main results from Table 2, presenting the confidence intervals for the CiC effect estimates based on unadjusted $R^2$s next to those for estimates based on adjusted $R^2$s. The four effects of the CiC model are presented in rows, with the unadjusted effect estimates on the left and the adjusted effect estimates on the right.

When we compare the results based on adjusted $R^2$s with those calculated with unadjusted $R^2$s, the differences are consistent with our reasoning above. We focus first on the point estimates. The year effects based on adjusted $R^2$s are somewhat lower than those based on unadjusted $R^2$s, but the differences are small, at least in absolute terms. As one would expect, the differences between the adjusted and unadjusted year effects get smaller as we expand the sample in the later steps by adding more observations for an (almost) constant number of years. The estimates for industry and firm effects are very similar for unadjusted and adjusted $R^2$s (mean differences across samples <0.1%).[25] In contrast, and predictably, the CEO effect based on adjusted $R^2$s is markedly lower than the estimates based on unadjusted $R^2$s—the estimates now range between 27.81% (Step 2) and 34.02% (Step 4), whereas those based on unadjusted $R^2$ are between 34.92% (Step 1) and 40.72% (Step 4). On average, across the H&Q (2014) sample and our sample extensions, the CEO effect estimates based on adjusted $R^2$s are 6.99 percentage points lower than those based on unadjusted $R^2$s.

An important insight, particularly apparent in Figure 1, is that the $R^2$ adjustment leaves the differences between effect estimates across sample extensions (i.e., our five steps) largely intact: the estimates are effectively shifted downwards in parallel. Hence, whenever confidence intervals do not overlap between subsamples with unadjusted $R^2$s, they also do not overlap for

[23]One could argue that *I-BENCH*, while technically only one variable, is similar to a set of industry benchmarks, that is, industry-year dummy variables or industry-year ROA means. This is because excluding the focal firm from the yearly ROA mean leads to values that are firm-year-specific, but nonetheless very similar within a given industry-year. This holds especially in large industries, where the removal of one observation does not materially affect the mean. The more the ex-focal-firm benchmark correlates with the benchmark including the firm, the more it is comparable to a set of industry-year indicators. To the best of our knowledge, there is no guidance in the literature on how to adjust $R^2$ in such a context, and we therefore use the "standard" approach of considering *I-BENCH* as a single regressor. However, if *I-BENCH* were interpreted as a set of indicators, introducing it into the regression equation would use up a higher number of degrees of freedom, and the adjustment of $R^2$ would have a greater impact on the industry effect.

[24]The effect estimates based on unadjusted $R^2$ in Table 2, Panel A, are identical to those presented in Table 1. We include them here again to ease direct comparisons with the results based on adjusted $R^2$s.

[25]As appears in Panel C of Table 2, most of the differences between the industry and firm estimates based on unadjusted $R^2$s and those based on adjusted $R^2$s are negative. However, these differences are all very close to zero in absolute terms. Furthermore, the estimation of the industry and the firm effects follows the estimation of the year effects, and the $R^2$-adjustment reduces the year effects more strongly, leaving more variance potentially to be explained in the subsequent stages of the CiC model. The resulting increase in the industry and firm effects appears to overcompensate for the very small adjustment of the partial $R^2$s.

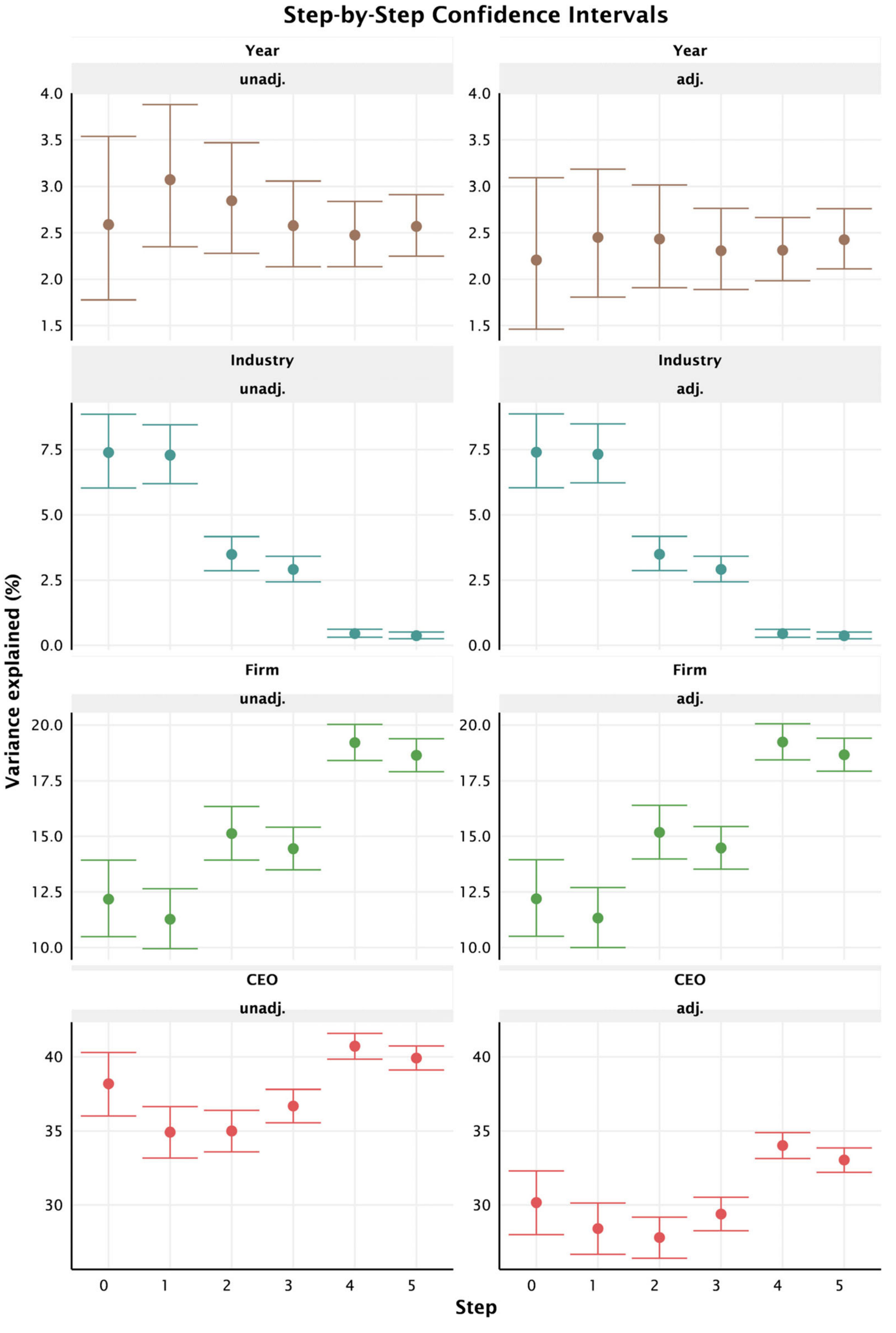


FIGURE 1 Legend on next page.

**FIGURE 1** Confidence intervals for year, industry, firm, and CEO effect estimates from the CiC model, based on OLS estimation and unadjusted $R^2$s (left column) and adjusted $R^2$s (right column), for the original H&Q (2014) sample and step-by-step sample extensions. The "handlebars" represent 95% confidence intervals, based on Fisher's $z$-transformations, for the effect estimates from the CiC model estimations, that is, the lower bar indicates the lower value of the confidence interval, and the upper bar indicates the upper value. The graphs use different scalings on the vertical axes. Descriptive statistics of the various samples are presented in Table 1

subsamples with adjusted $R^2$s. This is the case for the industry and firm effects in Steps 2 and 4, and for the CEO effect in Step 4.

To sum up, CEO effect estimates are inflated when they are based on unadjusted $R^2$s. Thus, in line with Quigley and Graffin (2017), the H&Q (2014) CiC model should be used only with adjusted $R^2$s. As our analysis shows, doing so markedly lowers the CEO effect, but does not strongly alter the year, industry, and firm effects, nor the comparisons of effects across samples.

## 5.3 | Part 3: Sensitivity of H&Q (2014) model findings to sample characteristics

In the third and last part of our analyses, using only adjusted $R^2$s, we conduct additional tests to examine in detail the sensitivity of the H&Q (2014) CiC method to two sample characteristics, firm size, and CEO tenure. This analysis also responds to a call, expressed several times in the CEO effect literature, to investigate in more detail under what conditions and circumstances CEOs matter more, or less (e.g., Fitza, 2017; Hambrick & Quigley, 2014; Quigley & Graffin, 2017). We examine firm size because H&Q (2014) focused on large U.S. corporations, and in the stepwise sample enlargements in the first part of our empirical analyses we saw that adding smaller firms to H&Q's sample had a marked impact on the CiC estimations. And we investigate CEO tenure because it is closely connected to the central object of our investigation, the CEO effect, and because CEO tenure plays a central role in the argumentation of Fitza (2014). We conduct the following tests on our total sample, the one comprising 33,996 firm-year observations.

### 5.3.1 | Firm size

To examine the impact of firm size more systematically, we divide our largest sample into quartiles, using sales as a size proxy.[26] Figure 2 presents CiC estimation findings for the four quartiles; we also present the estimation results and descriptive statistics for the quartiles in Table S2.

The graph documents that firm size has nontrivial implications for the CiC model effects. Most striking is the impact on the industry effect—the performance of small firms tends to be idiosyncratic, and consequently, in the first quartile there is practically no industry effect at all (and even the year effect is significantly smaller than in the other size classes). As firms become larger, their performance gets more similar to the performance of their respective industries; that is, the industry effect estimates increase monotonically from one size quartile to the next,

[26]The results are qualitatively very similar if we use total assets instead.

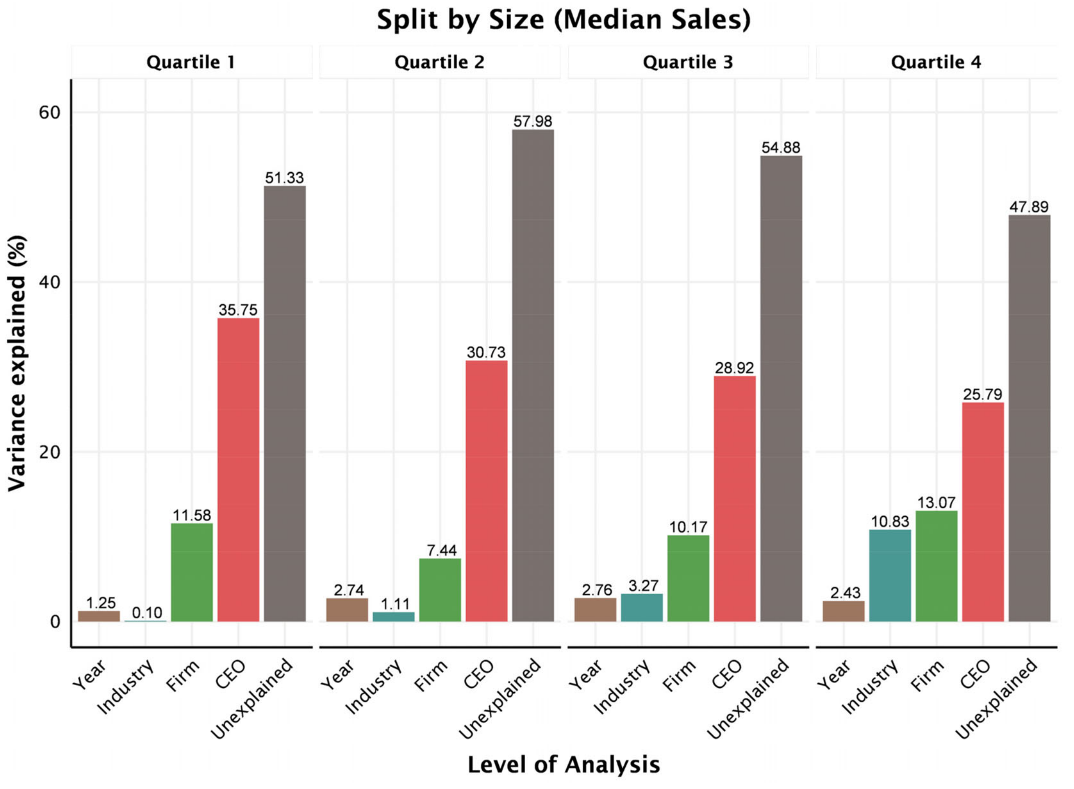


**FIGURE 2** CiC model estimations for quartile subsamples based on firm size (sales). The figure presents results from H&Q CiC model estimations for subsamples—quartiles based on firm size, as measured by sales revenues—of our largest sample ($n = 33{,}996$). The estimation findings for the smallest firms are presented on the left, and the estimation findings for the largest firms are on the right. The bars represent the estimated year, industry, firm, and CEO effects as well as the remaining unexplained variance from the CiC model. All effects are calculated using OLS estimation and adjusted $R^2$s. The estimation results and descriptive statistics for the various subsamples are summarized in Table S2

and in the fourth size quartile the industry effect explains a little more than 10 % of the firms' performance variance (10.83%).[27] A technical reason that helps to explain the pronounced impact of firm size on the industry effect is that H&Q use size-weighted ROAs as industry benchmarks. In this analysis, we split our sample according to size and then compare the firms in each quartile with the complete, "unsplit" industry benchmarks, that is, the entire population of firms in the same industry available in Compustat and Refinitiv Eikon. Obviously, the size-weighted benchmarks explain the performance of larger firms better than that of smaller firms. The quartile regressions presented in Figure 2 thus bring to light a specific aspect of the

[27]Here, and generally in the third part of our analyses, we focus on cross-quartile differences for which the confidence intervals derived from Fisher's z-transformations do not overlap. For example, the confidence intervals indicate that the industry effect estimates are clearly different from quartile to quartile. Analogously, the firm effects in quartiles one, two, and three are different from those in quartiles two, three, and four, respectively. And the CEO effect in the first quartile is different from that in the second. The CEO effects in the second and third quartiles are not different from each other, but the estimator in the third quartile is again different from that in the fourth. The confidence intervals for all effect estimates in all quartiles are reported in Table S2.

CiC method, because, in the "normal" application of the method in the total sample, smaller sample firms are also compared to size-weighted industry benchmarks. Figure 2 further reveals that the firm effect is more pronounced for very small (11.58%) and very large firms (13.07%) and less pronounced for medium-sized ones (7.44% and 10.17%). It seems that inherited performance and health persist longer in small and large firms than in medium-sized firms. Finally, the CEO effect decreases monotonically as firm size increases. It is 35.75% for the smallest size quartile and only 25.79% for the largest quartile—10% points smaller in the largest quartile. It is intuitively obvious, and in line with previous studies (Finkelstein & Hambrick, 1990), that CEOs have relatively less discretion and thus less impact on firm performance as firms get larger.[28]

### 5.3.2 | CEO tenure

We present two sets of analyses to examine the impact of CEO tenure on the H&Q (2014) CiC model findings. First, we again split our sample into quartiles, this time by the firms' maximum CEO tenures.[29] The findings are summarized in Figure 3.[30] Second, to provide yet deeper insights we run a series of 24 CiC model estimations in which we successively increase the minimum tenure required for including a CEO in the samples. The "effect lines" resulting from these estimations are presented in Figure 4.

The two figures document that CEO tenure profoundly influences the estimation results. Figure 3 shows that in firms with shorter CEO tenures, that is, firms in the first and the second quartile, the industry effect does not explain firm performance at all. That is, the performance of these firms is quite idiosyncratic, dissimilar to that of their industry peers.[31] At the same time, firms with shorter CEO tenures have a pronounced firm effect. In fact, the firm effect is high in the first three quartiles, at 19.57% in the first quartile, 17.45% in the second, and 19.84% in the third quartile, before dropping to only 9.11% in the fourth quartile. Figure 3 further indicates that the CEO effect declines monotonically with increasing CEO tenure. It is 34.68% in the first quartile and only 26.42% in the fourth quartile.[32]

[28]We also note that firm size is related to CEO tenure. The average tenure increases monotonically with size; mean (median) tenure is only 5.31 (4) years in the first size quartile and 7.17 (6) years in the fourth quartile (Table S2). The length of CEOs' tenures may itself be related to the strength of the CEO effect; we examine this relation below. More generally, we acknowledge that the univariate analysis in this part of our study does not allow us to isolate the "pure" impact of a given firm characteristic (e.g., size) on the CiC model effects, by holding all other factors constant.

[29]As in the previous investigation on firm size, we split the sample into quartiles based on a *firm* characteristic (here, maximum tenure) and assign all observations for a given firm to the same quartile. Thus, the number of firms is roughly the same in all four quartiles. However, in the present investigation, the numbers of firm-year observations are of course higher in the quartiles with longer maximum tenures.

[30]For an overview of estimation results and descriptive statistics, see Table S3.

[31]Accordingly, the standard deviation of ROA is negatively related to CEO tenure; it is 46.9% in the first quartile and only 12.9% in the fourth quartile; see Table S3 for details; also see Brookman and Thistle (2009) on the relation between CEO tenure and firm characteristics.

[32]These numbers gauge the importance of the CEO relative to the variance of ROA in the respective subsamples. In absolute terms, CEOs "matter" much more in the first quartile because the performance of firms in this quartile is much more heterogeneous. Among the firms in the first quartile, CEOs explain 36.36% of a standard deviation of ROA of 46.9% age points, whereas in the fourth quartile CEOs explain 26.34% of a standard deviation of ROA of only 12.9% points.

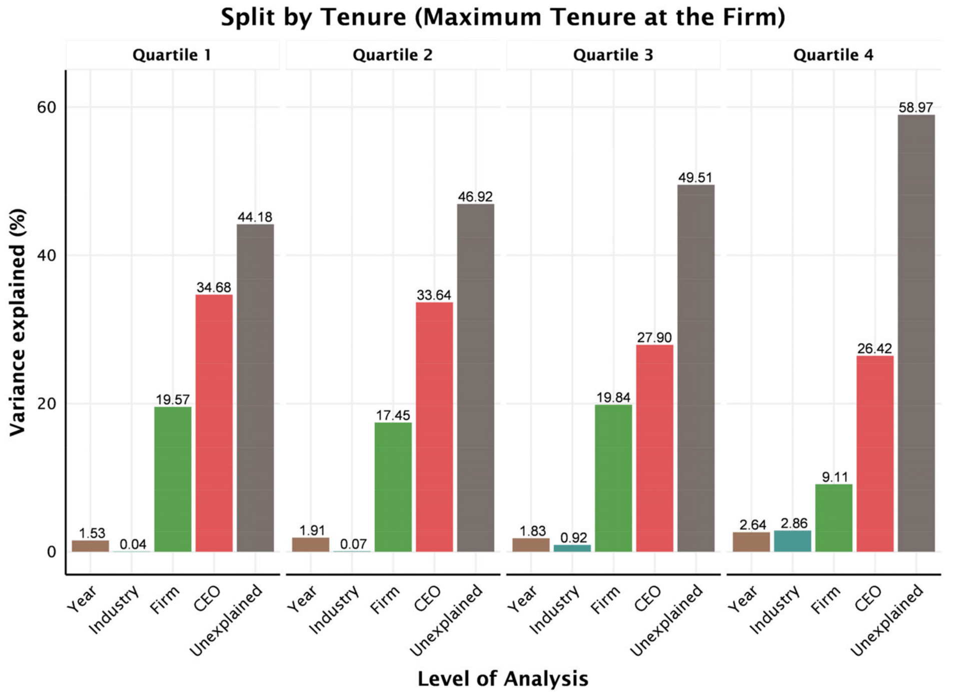


**FIGURE 3** CiC model estimations for quartile subsamples based on length of CEO tenure (maximum CEO tenures). The figure presents results from H&Q CiC model estimations for subsamples—quartiles based on the length of the CEOs' tenures, as measured by firms' maximum CEO tenures—of our largest sample ($n$ = 33,996). The estimation findings for firms with the shortest CEO tenures are presented on the left, and the findings for firms with the longest CEO tenures are on the right. The bars represent the estimated year, industry, firm, and CEO effects as well as the remaining unexplained variance from the CiC model. All effects are calculated using OLS estimation and unadjusted $R^2$s. The estimation results and descriptive statistics for the various subsamples are summarized in Table S3

The effect lines in Figure 4 show in greater detail that year (brown line) and industry effect (turquoise line) estimates increase steadily with longer CEO tenures, albeit on low absolute levels. The firm effect (green line) and the CEO effect (red line) decrease as we increase the tenure requirement, until minimum tenures of 13 and 11 years, respectively. The negative association between CEO tenure and the firm effect may be at least partly due to a technical reason. The CiC model proxies for the firm effect, inherited profitability ($INPROF_{ko}$), and inherited health ($INHEALTH_{ko}$), are CEO-specific and get "updated" with each new CEO. While these proxies yield greater insight than the simple firm dummies used in traditional variance decomposition studies, relying on them implies that the firm itself affects performance over the entire course of the CEO's tenure *solely* through inheritance. Put differently, changes on the firm level that affect performance are attributed entirely to the CEO; the rest of the organization is assumed not to matter at all. It should also be noted that, in general, with increasing tenures, it becomes increasingly difficult to disentangle statistically the firm effect and CEO effect.

Turning to the CEO effect, one might expect that CEOs with long tenures influence their firms' performance more profoundly. Figure 4 indicates the opposite: the CiC model estimates of the CEO effect become smaller and smaller as we impose higher and higher minimum tenure restrictions. There are several possible explanations for this phenomenon. First, over time CEOs

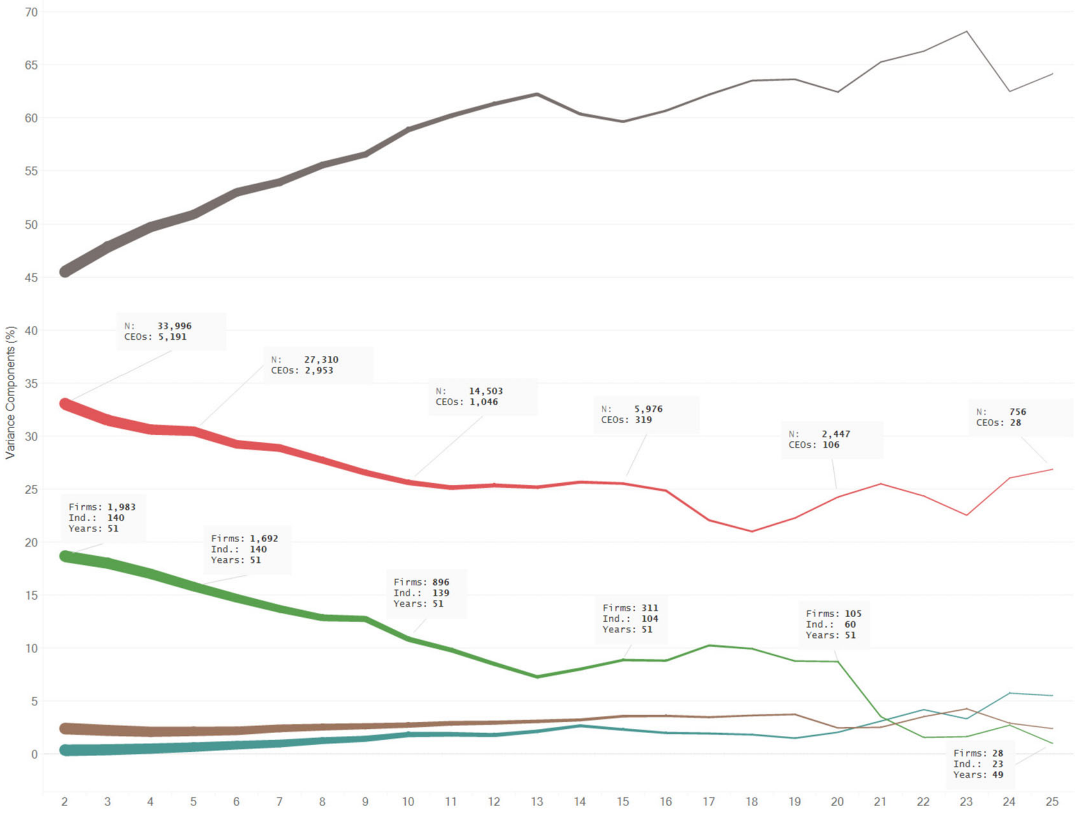


**FIGURE 4** CiC model effect lines, conditional on CEO minimum tenure. The figure presents estimation results from a series of 24 CiC model estimations in which we successively increase the minimum tenure that we require for the inclusion of the CEOs in the respective samples. The brown line represents the results for the year effects, the turquoise line the industry effect, the green line the firm effect, and the red line the CEO effect. The gray line at the top of the figure shows the remaining unexplained variance. All of the effect estimates are based on OLS estimations and adjusted $R^2$s. The points on the very left of the lines are based on a minimum tenure of 2 years, that is, the same requirement as in our main analysis. As we move along the lines to the right, we increase the minimum tenure to 3, 4, and more years, and eliminate all CEOs with shorter tenures, until we reach a minimum tenure of 25 years. The thickness of the effect line indicates the size of the sample

may lose some of their energy and become less adaptable to changing environments (e.g., Miller, 1991; on the dynamics of managerial capabilities, also see Adner & Helfat, 2003; Helfat & Martin, 2015).[33] Second, firms with longer tenures may operate in stable markets, in which firms may need (and allow) minimal CEO discretion (Hambrick & Abrahamson, 1995).[34] Third, the association between CEO tenure and the strength of the CEO effect may also be driven by other variables that are related to both tenure and the CEO effect—for example, firm size, which increases, and firm performance, which improves, over the CEO tenure quartiles

[33]Also see Simsek (2007) on the relation between CEO tenure and top management's risk-taking propensity.

[34]Accordingly, we find that industry discretion scores from Hambrick and Abrahamson (1995) and Finkelstein, Hambrick, and Cannella (2009) decrease monotonically over the four quartiles.

(Table S3).[35] Fourth, the use of CEO dummies assumes a constant CEO effect over the whole tenure. More precisely, the dummies measure a time-invariant difference in ROA for a given CEO compared to other, earlier, and later, CEOs at the helm of the same firm. But the ROA differential might not be constant. For example, during a crisis a CEO may incur write-offs and restructuring losses[36]; if the turnaround is successful, earnings recover in later periods. Thus, a strong CEO action may lead, in the CiC model (as well as in other variance decomposition models) to a rather weak CEO effect.

Given that the two most potent effects in the CiC model, the firm effect and the CEO effect, lose explanatory power with higher minimum tenures, it is clear that the unexplained residual variance must increase. In Figure 4, the residual variance (represented by the gray line) is 45.49% at the very left, that is, in our full sample with the two-year minimum tenure, and it increases to more than 60% with minimum tenures of 11 or more years.

Finally, Figure 4 reveals that the relations between CEO tenure and year, industry, firm, and CEO effects become less stable on the right-hand side of the graph, with sudden kinks and reversals. One reason is that the ever-stricter tenure requirements lead to changes in the industry composition of the sample, which in turn can provoke sudden changes in effect estimates.[37] Another reason is that the size of the sample decreases towards the right-hand side, and removing further firms from the sample thus can have a stronger impact on the overall result.

## 6 | DISCUSSION

In the following, we summarize and discuss the main findings from our empirical analyses.

(1) Our replication generally confirms H&Q's main finding of a high CEO effect on firm performance. We confirm the high CEO effect in H&Q's original sample and in our stepwise sample extensions. We again document strong CEO effects, albeit shifted downwards to a somewhat lower absolute level, when we reestimate the CiC model using adjusted $R^2$s (33.04% for the largest sample, see Table 2). And we also generally find pronounced CEO effects when we split our sample by firm size or maximum CEO tenure, still using adjusted $R^2$s. The last results in our analyses suggest that a lower-bound estimate for the CEO effect, among U.S. CEOs with long tenures, may be 25% (Figures 3 and 4). While lower than the 38.5% originally reported by H&Q, our estimates are corrected for the number of variables used in the estimation equations, and they are still larger than the CEO effects reported in most previous studies that used traditional ANOVA or MLM variance partitioning methods in most previous studies that used traditional ANOVA or MLM variance partitioning methods (see Table S1 in the online appendix).

[35]CEO tenure may also vary across industries. Hence, the firms in the four quartiles may differ not only in CEO tenure, but also in industry composition and perhaps other related characteristics.

[36]See, for example, Strong and Meyer (1987). Guan, Wright, and Leikam (2005) find evidence that suggests that incoming CEOs often actively depress earnings by engaging in "big bath" earnings management.

[37]Our full sample comprises firms from 140 industries (SIC 3). When we impose minimum CEO tenure requirements, up to a minimum of 6 years the number of industries within the sample remains constant. With minimum tenures between 7 and 10 years, the sample still comprises 139 industries, that is, only one industry drops out. However, at minimum tenures of 12 (15) [18] years, only 124 (104) [81] industries are left, and the drop-out of several industries in each of these steps can have a marked impact on the estimates. It should be noted that the sample sizes, in terms of firm-year observations, remain relatively large even at higher minimum tenure levels. For example, at a minimum CEO tenure of 16 years the sample is made up of 5,001 firm-year observations, more than in H&Q's original (2014) study.

While the CiC method consistently generates high CEO estimates, the industry effect estimates in our sample extensions and in most of the sample splits are lower (and firm effect estimates are higher) than those found by H&Q (2014).[38] There may be several reasons for this. First, our total sample comprises not only large firms, as did the original H&Q (2014) sample, but also medium-sized and smaller firms, whose performance tends to be more idiosyncratic. With the addition of the smaller firms to the sample, the within-industry variance becomes larger relative to the between-industry variance, rendering the industry effects less useful. Second, the H&Q CiC model uses size-weighted industry means as benchmarks. This was justified in H&Q's original study, with its focus on the largest U.S. firms, but may not be appropriate for studies targeting samples that also include medium-sized and smaller firms. Third, unlike the simple industry indicator variables in traditional variance partitioning studies, the CiC model's industry benchmarks (*I-BENCH*) by construction exclude the focal company and, in this sense, generate "out-of-sample" estimates, for which it is much harder to achieve high $R^2$ scores than for pure within-sample explanation. The weak industry effect leaves most of the performance variance unexplained in Equation (2) of the CiC model, allowing stronger explanatory power for the firm-specific variables *INPROF* and *INHEALTH* in Equation (3).

(2) The CEO effect is markedly lower with adjusted $R^2$s. The reason is that Equation (4) of the CiC model includes a large number of CEO dummy variables—in H&Q's original study, 829. Given the sample size of 4,866 firm-year observations, there are only 5.87 observations for each additional regression parameter to be estimated. Similarly, in our largest sample, with 5,191 CEOs and a sample size of 33,996 firm-year observations, there are 6.55 observations for each additional parameter estimate. Rules of thumb in the econometrics literature typically require 20, 15, or at least 10 observations per predictor in linear regression estimations.[39] In other words, while both H&Q's sample and our own may appear large in absolute terms, they are actually not, in relation to the number of model parameters to be estimated. In technical terms, the number of degrees of freedom in the estimation of Equation (4) of the CiC model is quite low (the same holds for the estimation of the CEO effect in other variance decomposition studies), and as a result, the CEO effect is likely to be overestimated because the model artificially describes part of the random error in the data. Because of this "overfitting" the unadjusted $R^2$s from the estimation do not reflect the model's "true" predictive ability in the general population.[40] To correct for the overfitting, in line with Quigley and Graffin (2017) the H&Q (2014) CiC model should be used with adjusted $R^2$s.[41]

---

[38]In H&Q's (2014) sample, the industry effect was 6.9%. In our largest sample it is only 0.37%. Earlier CEO studies, as well as studies focusing on industry, parent company, and business unit effects (e.g., Bowman & Helfat, 2001; McGahan & Porter, 2002; Short, Ketchen, Palmer, & Hult, 2007), mostly arrive at industry effects between 10% and 20%.

[39]See, for example, Harrell (2015, pp. 72–74). In a more detailed analysis, Green (1991) suggests that regression analysis should be based, at a minimum, on a base sample of 50 observations plus eight additional observations per explanatory variable. Also see Austin and Steyerberg (2015) and Jenkins and Quintana-Ascencio (2020).

[40]At first glance, one could perhaps argue that our dataset comes close enough to the general population of CEOs of U.S. stock-listed companies during our sample period so that there might be no need to make inferences; instead, it might suffice to interpret the observed effects. However, such an argument would miss the point that, even if we could indeed observe all CEOs of all U.S. firms during a given sample period, the unadjusted $R^2$s from estimations of the CiC model, while reflecting *statistical associations* within this sample, will always be inflated by random noise and thus cannot be interpreted as measuring the *effects of CEOs* on firm performance. Furthermore, the observed unadjusted $R^2$s are not predictive for new data generated by firms in years after the sample period, or for firms and CEOs in countries outside the U.S. Thus, it may be useful to interpret our sample as being drawn from a hypothetical infinite "superpopulation" (see Lavrakas, 2008).

[41]Alternatively, one could assess the model fit in the population also by using out-of-sample prediction or cross validation; see, for example, Harrell (2015).

(3) The CiC model findings are sensitive to the sample characteristics firm size and CEO tenure. Some of the associations between these characteristics and the model results are conceptually founded and intuitively plausible. For example, CEO effects are larger, and firm effects are smaller, in smaller (and more volatile) companies. In other cases, our analyses have revealed associations that are not immediately obvious. One example is the link between firm size and the strength of the industry effect, a link that, inter alia, likely reflects the use of size-weighted industry means as benchmarks. Another example is the negative relation between CEO tenure and the strength of the CEO effect. One could perhaps argue that longer CEO tenures allow for a more precise estimation and that the lower CEO effect estimates in samples with longer CEO tenures are therefore more accurate. On the other hand, the median CEO tenure in our most comprehensive sample is 5 years. Restricting the sample to long tenures would thus ignore a large part of the population and would potentially bias the findings, to the extent that short CEO tenures indeed differ systematically from longer ones. As we explain above, focusing on long CEO tenures also raises another issue, the assumed constancy of the estimated effects. The CiC method generates average, "fixed" effects; in this sense, like most other variance partitioning models, it is "cross-sectional in design and static in its logic" (Henderson, Miller, & Hambrick, 2006, p. 447). However, research suggests that CEOs arrive in their positions well equipped and learn even more in the early phases of their tenures, but later become "tired, enshrined and stale" (Miller, 1991, p. 41). Moreover, the speed of this obsolescence depends on the dynamism of the industry (Henderson et al., 2006). The CiC model and most other variance decomposition models mask these temporal effects, as well as possible industry differences.[42]

## 7 | CONCLUSION

Using a refined and arguably superior method to partition firm performance variance and to isolate the CEO's impact on firm performance, and using data for large U.S. firms for the years 1992–2011, Hambrick and Quigley (2014) estimated the CEO effect at 38.5%, much higher than most previous estimates based on traditional ANOVA or multilevel modeling. In the present paper, we replicate H&Q's study, apply their CiC technique to a much more comprehensive U. S. sample, and assess the sensitivity and robustness of their findings to variations in the method and the data.

We generally confirm H&Q's finding of a high CEO effect, but find a smaller industry effect and a larger firm effect in our larger sample. We show that applying the CiC technique using adjusted $R^2$s changes year, industry, and firm effects only moderately, but markedly reduces the CEO effect. And we document that the CiC technique is sensitive to sample characteristics, namely firm size and CEO tenure.

While our replication study speaks directly to the validity and generalizability of H&Q's (2014) argument, we also more generally add to the discourse on CEO effects and the importance of top managers, and we hope that our findings will provide impulses for future research. For example, further research could revisit the question "when and where CEOs matter most

[42]Guo (2017) employs longitudinal multilevel modeling to analyze stable and dynamic parts of performance variance, albeit for business unit, corporation, industry, and year effects, not for CEO effects.

(and least)" (Hambrick & Quigley, 2014, p. 488) by examining moderators other than firm size and CEO tenure. Moreover, like most research in this area,[43] our study is based on data for U.S. firms, and future research could examine whether CEOs are equally important in countries other than the U.S. Finally, future research could also attempt to improve the external validity of CEO effect studies by going beyond mere in-sample estimations and assessing the models' predictive power also with out-of-sample tests or cross-validation.[44]

## ACKNOWLEDGEMENTS

We are very grateful to Tim Quigley for having made the H&Q (2014) dataset and Stata code available to us. We are also very grateful to Constance Helfat, the editor, for her excellent guidance throughout the review process. We also thank the two anonymous reviewers for their critical and helpful comments and Michael Massmann for helpful comments on the GEE method and the adjustment of $R^2$. Open Access funding enabled and organized by Projekt DEAL.

## ORCID

*Tobias Keller* https://orcid.org/0000-0002-4923-6837
*Martin Glaum* https://orcid.org/0000-0002-9300-1410
*Andreas Bausch* https://orcid.org/0000-0002-5473-9324
*Thorsten Bunz* https://orcid.org/0000-0002-3956-8117

## REFERENCES

Adner, R., & Helfat, C. E. (2003). Corporate effects and dynamic managerial capabilities. *Strategic Management Journal*, *24*(10), 1011–1025.

Austin, P. C., & Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology*, *68*(6), 627–636.

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*, *7*(2), 127–150.

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics*, *118*(4), 1169–1208.

Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchell, W. (2016). Creating repeatable cumulative knowledge in strategic management: A call for a broad and deep conversation among authors, referees, and editors. *Strategic Management Journal*, *37*(2), 257–261.

Bettis, R. A., Helfat, C., & Shaver, J. M. (2016). The necessity, logic, and forms of replication. *Strategic Management Journal*, *37*(11), 2193–2203.

Bowman, E. H., & Helfat, C. E. (2001). Does corporate strategy matter? *Strategic Management Journal*, *22*(1), 1–23.

Brookman, J., & Thistle, P. D. (2009). CEO tenure, the risk of termination and firm value. *Journal of Corporate Finance*, *15*(3), 331–344.

Chang, S.-J., & Sing, H. (2000). Corporate and industry effects on business unit competitive position. *Strategic Management Journal*, *21*(7), 739–752.

Child, J. (1972). Organizational structure, environment and performance: The role of strategic choice. *Sociology*, *6*(1), 1–22.

[43]An exception is the work of Crossland and Hambrick (2007, 2011), who investigate differences in CEO discretion across countries.

[44]In out-of-sample tests, models are estimated with only a portion of the complete samples (training data); the estimated models are then applied to the remaining data (holdout sample) to make predictions for the dependent variables. See, for example, Harrell (2015) for further details. Developing an out-of-sample test for CEO effect studies would be nontrivial, because CEOs, the main focus of these studies, work for companies only for some time. Ideally, the cross-validation design should ensure that the same CEOs exist in both the training and the hold-out set.

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*(2), 358–372.

Cox, N. J. (2008). Speaking Stata: Correlation with confidence, or Fisher's z revisited. *The Stata Journal*, *8*(3), 413–439.

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*(8), 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*(8), 797–819.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147–160.

Finkelstein, S., & Hambrick, D. C. (1990). Top-management-team tenure and organizational outcomes: The moderating role of managerial discretion. *Administrative Science Quarterly*, *35*(3), 484–503.

Finkelstein, S., Hambrick, D. C., & Cannella, A. A. (2009). *Theory and research on executives, top management teams, and boards*. Oxford: Oxford University Press.

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*(12), 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*(3), 802–811.

Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research*, *26*(3), 499–510.

Guan, L., Wright, C. J., & Leikam, S. L. (2005). Earnings management and forced CEO dismissal. *Advances in Accounting*, *21*, 61–81.

Guo, G. (2017). Demystifying variance in performance: A longitudinal multilevel perspective. *Strategic Management Journal*, *38*(6), 1327–1342.

Hambrick, D. C., & Abrahamson, E. (1995). Assessing managerial discretion across industries: A multimethod approach. *Academy of Management Journal*, *38*(5), 1427–1441.

Hambrick, D. C., & Finkelstein, S. (1987). Managerial discretion: A bridge between polar views of organizational outcomes. *Research in Organizational Behavior*, *9*, 369–406.

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review*, *9*(2), 193–206.

Hambrick, D. C., & Quigley, T. J. (2014). Toward more accurate contextualization of the CEO effect on firm performance. *Strategic Management Journal*, *35*(4), 473–491.

Hannan, M. T., & Freeman, J. (1977). The population ecology of organizations. *American Journal of Sociology*, *82*(5), 929–964.

Hardin, J. W., & Hilbe, J. M. (2013). *Generalized estimating equations* (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC Press.

Harrell, F. E. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis* (2nd ed.). Heidelberg: Springer.

Helfat, C. E., & Martin, J. A. (2015). Dynamic managerial capabilities: Review and assessment of managerial impact on strategic change. *Journal of Management*, *41*(5), 1281–1312.

Henderson, A. D., Miller, D., & Hambrick, D. C. (2006). How quickly do CEOs become obsolete? Industry dynamism, CEO tenure, and company performance. *Strategic Management Journal*, *27*(5), 447–460.

Hubbard, R., Vetter, D. E., & Little, E. L. (1998). Replication in strategic management: Scientific testing for validity, generalizability, and usefulness. *Strategic Management Journal*, *19*(3), 243–254.

Jacobs, G., & O'Neill, C. (2003). On the reliability (or otherwise) of SIC codes. *European Business Review*, *15*(3), 164–169.

Jenkins, D. G., & Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. *PLoS One*, *15*(2), e0229345.

Kahle, K. M., & Walkling, R. A. (1996). The impact of industry classifications on financial research. *Journal of Financial and Quantitative Analysis*, *31*(3), 309–335.

Kennedy, P. (2008). *A guide to econometrics* (6th ed.). Cambridge, MA: MIT Press.

Lavrakas, P. J. (2008). Superpopulation. In *Encyclopedia of survey research methods*. Thousand Oaks, CA: Sage.

Lieberson, S., & O'Connor, J. F. (1972). Leadership and organizational performance: A study of large corporations. *American Sociological Review*, *37*(2), 117–130.

Mackey, A. (2008). The effect of CEOs on firm performance. *Strategic Management Journal*, *29*(12), 1357–1367.

McGahan, A. M., & Porter, M. E. (2002). What do we know about variance in accounting profitability? *Management Science*, *48*(7), 834–851.

Miller, D. (1991). Stale in the saddle: CEO tenure and the match between organization and environment. *Management Science*, *37*(1), 34–52.

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*(6), 571–590.

Olkin, I., & Finn, J. D. (1995). Correlations redux. *Psychological Bulletin*, *118*(1), 155–164.

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*(3), 793–780.

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*(6), 821–830.

Rumelt, R. (1991). How much does industry matter? *Strategic Management Journal*, *12*(3), 167–185.

Schenker, N., & Gentleman, J. F. (2001). On judging the significance of differences by examining the overlap between confidence intervals. *The American Statistician*, *55*(3), 182–186.

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

Short, J. C., Ketchen, D. J., Jr., Palmer, T. B., & Hult, G. T. M. (2007). Firm, strategic group, and industry influences on performance. *Strategic Management Journal*, *28*(2), 147–167.

Simsek, Z. (2007). CEO tenure and organizational performance: An intervening model. *Strategic Management Journal*, *28*(6), 653–662.

Six, B., Normann, M., Stock, R. M., & Schiereck, D. (2013). Strategic leaders' impact on corporate policies and firm performance: Insights from CEOs and CFOs of large listed firms in Germany. *Schmalenbach Business Review*, *65*(2), 82–111.

Strong, J. S., & Meyer, J. R. (1987). Asset writedowns: Managerial incentives and security returns. *The Journal of Finance*, *42*(3), 643–661.

Thomas, A. B. (1988). Does leadership make a difference to organizational performance? *Administrative Science Quarterly*, *33*(3), 388–400.

Tsang, E. W., & Kwan, K. M. (1999). Replication and theory development in organizational science: A critical realist perspective. *Academy of Management Review*, *24*(4), 759–780.

Wang, W., Zhang, X., & Paap, R. (2019). To pool or not to pool: What is a good strategy for parameter estimation and forecasting in panel regressions? *Journal of Applied Econometrics*, *34*(5), 724–745.

Wasserman, N., Anand, B. N., & Nohria, N. (2010). When does leadership matter? A contingent opportunities view of CEO leadership. In N. Nohria & R. Khurana (Eds.), *Handbook of leadership theory and practice* (pp. 27–63). Boston, MA: Harvard Business Press.

Weiner, C. (2005). *The impact of industry classification schemes on financial research (discussion paper 2005–062)*. Berlin: School of Business and Economics, Humboldt-Universität Berlin.

Weiner, N. (1978). Situational and leadership influences on organization performance. In *Proceedings of the Annual Meeting of the Academy of Management Proceedings* (pp. 230–234).

Weiner, N., & Mahoney, T. (1981). A model of corporate performance as a function of environmental, organizational, and leadership influences. *Academy of Management Journal*, *24*(3), 453–470.

Withers, M. C., & Fitza, M. A. (2017). Do board chairs matter? The influence of board chairs on firm performance. *Strategic Management Journal*, *38*(6), 1343–1355.

Zhang, D. (2017). A coefficient of determination for generalized linear models. *American Statistician*, *71*(4), 310–316.

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, *12*(4), 399–413.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** 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*(4), 1111–1138. https://doi.org/10.1002/smj.3453