

# Revisiting managerial “style”: The replicability and falsifiability of manager fixed effects for firm policies

Victor Esteban Jarosiewicz  | David Gaddis Ross 

Warrington College of Business Administration, University of Florida, Gainesville, Florida, USA

## Correspondence

David Gaddis Ross, Warrington College of Business Administration, University of Florida, Gainesville, FL 32611, USA.  
Email: [david.ross@warrington.ufl.edu](mailto:david.ross@warrington.ufl.edu)

## Abstract

**Research Summary:** We attempt to replicate Bertrand and Schoar (2003), who used manager fixed effects to provide evidence for managerial “style” as measured by metrics such as leverage, R&D spending, and Tobin’s Q. We find that the statistical and economic significance of manager fixed effects is generally lower in our results than in theirs. This discrepancy motivated us to conduct placebo tests in which we randomize managers’ spells at their firms. The results with the randomized data are generally as strong as those with the real data, suggesting (a) that the apparent explanatory power of manager fixed effects may be largely a statistical artifact and (b) that managerial “style” may manifest itself in more complex ways than can be ascertained from examining financial statement metrics.

**Managerial Summary:** An influential body of research uses large datasets of firm financial information to study whether individual top managers have a “style” as measured by metrics such as leverage and R&D spending. We attempt to replicate a leading paper in this body of research but obtain quite different, usually weaker results. We also try repeating the analysis using artificial datasets where managers are randomly assigned to firms other than those at which they really worked. These randomized datasets produce results that are generally as strong as those with the real data.

The implications are that the statistical methodology used in this body of research may be generating spurious results and that managerial “style” may be too complex to ascertain from examining financial statement metrics.

#### KEY WORDS

firm policy, governance, panel-data methods, replication, top management

## 1 | INTRODUCTION

The field of strategic management is fundamentally concerned with identifying the drivers of profitability (Makadok, 2011). In that vein, much progress has been made in theorizing about antecedents of firm policy and performance such as industry structure (Brandenburger & Nalebuff, 1996; Porter, 1980), resource acquisition (Barney, 1986), and first-mover advantage (Lieberman & Montgomery, 1988), as well as in quantifying the relative importance of competitive elements like industry, year, business unit, and corporate parentage (e.g., McGahan & Porter, 1997; Rumelt, 1991). Similarly, an extensive literature has identified the contribution of top management team members, such as the “CEO effect,” to various measures (Hambrick & Quigley, 2014; Lieberson & O’Connor, 1972; Quigley & Graffin, 2017; Withers & Fitzmaurice, 2017).

And yet, exactly how top executives influence their firms’ profits is not fully resolved, despite their apparent importance for firm policy and performance (Barnard, 1938; Castanias & Helfat, 1991, 2001; Hambrick & Mason, 1984; Penrose, 1959). Readily observable firm policies like gender diversity in top management (Dezső & Ross, 2012), diversification into new industries (Villalonga, 2004), and market entry mode (Shaver, 1998) are only effective given certain contingencies. Strategies that work well for some firms do not work well for others. One may infer that however much we may associate famous executives like Steve Jobs with innovation in consumer technology or Sam Walton with efficient discount retailing, imitating readily observable policies of theirs is insufficient to match their success.

It is perhaps for this reason that the work of economists Bertrand and Schoar (2003) (“B&S”) has attracted so much attention. B&S created a dataset of top managers who moved among firms, allowing them to simultaneously calculate manager and firm fixed effects with respect to firm policy and performance (the “MDV” methodology for “mover dummy variable”). Using this novel dataset, B&S provided strong statistical evidence that individual manager fixed effects had meaningful statistical and economic explanatory power for a wide range of firm policy and performance metrics. B&S also conducted a “placebo test” in which managers were, counterfactually, placed at firms 3 years before their actual tenure there. This placebo test generated substantially weaker results, suggesting B&S’s findings were not a statistical artifact.

On the one hand, B&S’s work, as well as work using a similar research design by Abowd, Kramarz, and Margolis (1999) (“AKM”), has spawned a rich and growing literature on measuring individual manager heterogeneity, including studies on individual auditors (Aobdia, Lin, & Petacchi, 2015; Gul, Wu, & Yang, 2013), firm accounting policies (DeJong & Ling, 2013), board members (Cavaco, Crifo, Rebérioux, & Roudaut, 2017), managerial compensation and incentives (Brookman & Thistle, 2013; Coles & Li, 2020; Graham, Li, & Qiu, 2012), and insider

trading (Hillier, Korczak, & Korczak, 2015). On the other hand, some have argued that manager fixed effects or similar variance decomposition methodologies may falsely attribute to the manager what is due to chance (Blettner, Chaddad, & Bettis, 2012; Fitza, 2014, 2017) or endogenous managerial changes (Fee, Hadlock, & Pierce, 2013).

This work raises the question of whether quantification of manager impact using fixed effects is truly measuring something of importance to strategic management, and, more broadly, whether a recipe for managerial success could indeed be found by imitating the financial statement policies associated with larger-than-life top managers such as James Pierpont Morgan, Henry Ford, and John D. Rockefeller from yesteryear, Steve Jobs and “Neutron” Jack Welch from recent history, and Tesla’s Elon Musk and Amazon’s Jeff Bezos today.

We revisit this question by attempting to replicate B&S. First, we construct a dataset that matches as closely as we are able the dataset in B&S. We then redo B&S’s main analysis with these data,<sup>1</sup> but we find that the statistical and economic significance of manager fixed effects is considerably lower in our results than in theirs, although sometimes we obtain stronger results. There is no obvious pattern to the discrepancies. The instability of these findings across what in principle should be very similar samples raises the possibility that the apparent impact of managers might be an artifact of the fixed effects methodology.<sup>2</sup>

So, following B&S itself and other prior work (e.g., Bertrand, Duflo, & Mullainathan, 2004), we conduct placebo tests. Our approach is to randomly assign manager spells to firms, thereby breaking the link between the firm and the manager. Such random assignment should make the explanatory power of manager fixed effects much lower, even nonexistent, because they are in principle just fitting random noise. And yet, our results with the randomized data are similar to—and sometimes stronger than—those with the real data, and in any case, too statistically and economically significant to arise by chance in an unbiased random process.

All told, our results suggest that the statistical and economic significance of manager fixed effects may not reflect actual differences among managers, but may instead be a statistical artifact of the methodology, perhaps of over-parameterizing the regressions with so many fixed effects. An implication is that truly understanding top managers’ styles may require a deeper dive into how these managers organize their firms’ business systems, interact with rivals, and acquire valuable resources. Another implication is that regressions with a large number of fixed effects have the potential to generate spurious results.

Reinforcing this point is our attempt to replicate B&S’s “parametric” analysis in which, for a given financial metric, they filter out firm fixed effects and other covariates and then regress the resulting average residual associated with a manager at one firm on the average residual associated with that manager at another firm. We fail to replicate this parametric analysis by an even wider margin than we did with B&S’s main analysis using manager fixed effects. Whereas B&S report very strong statistical significance for *every* financial metric they examine, in our attempt at replication, we obtain only two coefficients out of 11 with statistical significance of 10% or better, and the explanatory power ( $R^2$ ) of our regressions is generally much lower than in B&S.

<sup>1</sup>Specifically, we focus on tables III and IV from B&S, which examine the incremental explanatory power of executives for firm policy and performance variables. We also revisit some of the subsequent analyses from B&S, but we do not attempt to replicate their entire paper because our results differ so markedly from theirs.

<sup>2</sup>Shortly before submitting this paper to *Strategic Management Journal*, we contacted B&S, asking whether they would be willing to share some data to which we did not then have access, for an explanation of how they calculated some variables with respect to which their original paper is not clear, and for any thoughts they might have on our paper. B&S did not respond.

We also redo B&S's main analysis using data from a later time period. The results differ substantially from those we obtained with data from the earlier, original time period, and there is no obvious pattern to the discrepancies. Again, these results suggest that manager fixed effects are not capturing differences among managers.

## 2 | EMPIRICAL ANALYSIS

### 2.1 | Data sources

For the past couple decades, U.S. public firms have been required to disclose information on their chief executive officer (CEO) and other top four most highly paid managers, such as the chief financial officer (CFO), chief operating officer (COO), and the heads of divisions. B&S obtained information on such executives for the years 1992–1999 from ExecuComp, which covers the firms in the S&P 1,500, an index designed to reflect the overall U.S. equity market (Dezső, Ross, & Uribe, 2016). B&S report that they supplemented their data with the Forbes 800 files from 1969 to 1999.<sup>3</sup> The Forbes 800 is not a proper index or database, *per se*. Rather, it is a name used by researchers for the firms on the Forbes 500 lists. From 1969 to 2003, *Forbes* magazine released annual lists of the 500 largest public U.S. firms (using the previous year's data), according to different metrics (e.g., sales or assets). Because, in a given year, the firms comprising these different lists were not entirely the same, a total of approximately 800 firms would appear on at least one list. Starting in 1971 (which covered results for 1970), *Forbes* included the names of the CEOs of the firms appearing on one of the Forbes 500 lists. Thus, in principle, the B&S data would begin in 1970.

Like B&S, we obtain information on top managers from ExecuComp for 1992 to 1999, and supplement that with information about CEOs from the Forbes 800 files from 1970 to 1999. To ensure that we have the most accurate and complete compilation of the Forbes 800, as well as to merge the Forbes 800 as accurately as possible with ExecuComp, we obtained compilations of the Forbes 800 files up to the start of ExecuComp from DeJong and Ling (2013) and Murphy (2013),<sup>4</sup> and we also hand collected and assembled our own compilation of the Forbes 800 files directly from the source magazine issues. Then, using extensive hand-matching where necessary, we assembled a composite compilation of the Forbes 800 from these three sources and merged it into ExecuComp.

Following B&S, we obtain firm financial information from Compustat, stock trading information from CRSP, and merger and acquisition activity from Thompson SDC Platinum (which started its coverage in 1976). Inevitably, due to changes in the underlying data,<sup>5</sup> differences in the treatment of mergers, and so on, our dataset will not exactly correspond to the B&S dataset (see below). However, as we explain later, it does not seem plausible that such differences can explain the principal result of this paper, namely that we obtain results with randomized data that are so similar to what we obtain with real data and so strong in absolute terms relative to what one would expect from randomized data.

<sup>3</sup>From 1992 to 1999, the Forbes 800 files overlap with the S&P 1,500, which covers about 90% of the US stock market by value and includes the 500 largest firms that trade on the NYSE and Nasdaq and 1,000 “midcap” and “smallcap” firms that comprise companies with market capitalizations of at least \$3.6 billion and \$850 million, respectively (although these figures were considerably lower during the sample period).

<sup>4</sup>We are extremely grateful to these authors for sharing their data with us.

<sup>5</sup>Neither ExecuComp nor Compustat are “stable” databases with respect to historical data. ExecuComp reportedly backfills firm-years occasionally, for example, if a new firm is added to the S&P 1,500, and Compustat reportedly changes firm financial information in prior years, for example, if there is a restatement.

## 2.2 | Data samples

We follow B&S by excluding all utilities and financial companies (apparently, as defined by SIC codes in the 4,900 or 6,000 range). Consistent with B&S, we identify executive type by their last known role (e.g., CEO, CFO, or Other) in a sample period.<sup>6</sup> We create three different kinds of dataset from these data, as we now describe.

### 2.2.1 | Actual spells

A “spell” is a period of years in which a manager works for a particular firm. Following B&S, we create an unbalanced panel of firm-years by dropping all manager-firm spells that are less than 3 years in duration, and also all managers who do not have at least two such spells at different firms.<sup>7</sup> The movement of managers among firms is what allows the firm and manager fixed effects to be estimated simultaneously. (AKM showed that it is actually possible to identify the fixed effects of some managers who do not change firms from the movement of managers who do move from those same firms; we tried AKM’s methodology and obtained similar results.) The resulting dataset has 466 firms and 338 individuals, compared with about 600 firms and 500 individuals in B&S. Following B&S (pp. 1175–1176), “for each firm satisfying these requirements, we keep all observations, i.e., including years where the firm has managers that we do not observe in multiple firms.” Including these years helps estimate the firm fixed effects. Our dataset contains more than 9,000 observations compared to 6,766 observations in the original B&S sample.<sup>8</sup> On average, we obtain more than 24 fiscal years for each firm in our sample.

### 2.2.2 | Scrambled data

In a departure from B&S, we also create five “scrambled” datasets. In a scrambled dataset, executives are reassigned from their true firms to alternate firms, chosen at random from the dataset as it exists up to this point. What is essential is to scramble the spells in order to break the manager-firm association, so that when firm effects and manager effects are estimated, they are being estimated for managers not at their original firms for the firm-years in the observation.

<sup>6</sup>B&S (p. 1175) indicate that they use *titleann* in ExecuComp to code manager titles, which is a non-standardized text field copied out of firm annual proxy statements and is frequently missing. B&S do not explain what criteria they use to determine an executive’s role from the text in *titleann*. Like B&S, we use *titleann* to search for CEO and CFO titles, but we also make use of *ceoann* and *cfoann*, two fields in ExecuComp that indicate whether an executive is CEO or CFO, respectively. For over 90% of the firm-years it covers in our data, ExecuComp reports one and only one CEO as indicated by *ceoann*. ExecuComp reports one and only one CFO as indicated by *cfoann* for about half the firm-years it covers in our data, and rarely more than one. This is consistent with research showing that the CFO position became popular during the 1990s (Zorn, 2004). As noted, the Forbes 800 files only include CEOs.

<sup>7</sup>B&S argue 3 years is enough time for a manager to imprint his or her employer with the manager’s style. More practically, a 3-year minimum aids in identifying the individual fixed effects.

<sup>8</sup>We do not know why we obtain more observations yet have fewer firms and individual managers than B&S. In the Forbes data, which do not include a stable firm identifier like *gvkey* in Compustat, we found many instances of the names of firms and CEOs changing subtly from 1 year to the next. It is possible that B&S miscoded these firms or CEOs as different when they were the same. We also found instances of father-son pairs with the same name who were CEO of the same firm at different times. Again, we do not think these discrepancies with the B&S data could reasonably explain why we get results with randomized data that are so similar to those we obtain with actual data.

The randomization process is implemented by assigning an index number to each manager-firm spell and then randomly reassigned managers' spells (i.e., those index numbers) among other firms. If a given manager has a style that he or she imprints at, say, two different firms at which the manager works in the sample, we might expect that style to be reflected in the manager fixed effect. But if a manager is randomly assigned to firm spells, there is no reason, in principle, why the manager fixed effect should be a meaningful predictor of firm policy for the alternative firm. As an example, a Steve Jobs fixed effect might meaningfully explain the performance of Pixar and Apple, but would not explain the performance of two random firms to which it was assigned (like, say, Wal-Mart and Citigroup), especially since the managers at these two latter firms were also assigned randomly to various firms other than their own.

### 2.2.3 | Placebo data

We conduct an additional analysis in an attempt to break the link between manager and firm. Similar in spirit to the analysis in B&S table V (which we also attempt to replicate—see below), we employ a placebo process in which managers are “placed” at firms 3 years before they join them. We take firm-year observations 3 years before the manager joins each of the two firms, not just the second firm. We then scramble that dataset, randomly assigning executives to firms other than those where they actually served. So, not only are the executives at the wrong firms, but they are also there 3 years prior to their initial placement in their actual firms. In doing so, we attempt to create datasets where there is no plausible actual or even residual link between managers and either of the firms where they actually served.

## 2.3 | Variable definitions

### 2.3.1 | Policy and performance-dependent variables

We use the same firm policy and performance variables as B&S and make every effort to calculate these variables the same way, based on the information B&S provide in their data appendix. All years refer to fiscal years.

*Investment* is capital expenditures over net property, plant, and equipment at the beginning of the year (i.e., prior year end). *Leverage* is long-term debt plus debt in current liabilities, all divided by the sum of long-term debt, debt in current liabilities, and the book value of common equity. *Cash holdings* are cash and short-term investments, scaled by net property, plant, and equipment at the beginning of the year. *Interest coverage* is operating income before depreciation over interest expense. *Dividends/earnings* is the ratio of the sum of cash dividends on common and preferred stock, divided by earnings before interest, taxes, depreciation, and amortization (EBITDA). *R&D* is the ratio of R&D expenditures to prior year book assets. *Advertising* is the ratio of advertising expenditures to prior year book assets. *SG&A* is the ratio of selling, general, and administrative expenses to sales.

*N of acquisitions* is the total number of acquisitions a firm undertakes in the year. *N of diversifying acquisitions* is the subset of these acquisitions that a firm undertakes of a target firm in a different industry from the focal firm, where industry is defined by two-digit SIC. *Return on assets* (ROA) is the ratio of EBITDA over assets from the prior year. *Operating return on assets* is the ratio of operating cash flow to prior year assets.

### 2.3.2 | Control variables

We follow B&S exactly in the selection of covariates and control variables, both in the variables included in each specification and in the definition and measurement of each variable. The variable definitions are included below, and regression specifications are described in the following section.

*Tobin's Q* is the market value of assets divided by the book value of assets, where the market value of assets equals the book value of assets plus the market value of common equity less the sum of the book value of common equity and balance sheet deferred taxes. *Cash flow* is income before extraordinary items, divided by net property, plant, and equipment at the beginning of the fiscal year. Following B&S, log transformations are used for some variables to reduce skewness: *lnassets* is the log of total assets at the end of the fiscal year, and *laglnassets* is the lagged value of *lnassets*.

### 2.3.3 | Independent variables of interest

The primary variables of interest are the fixed effects of the managers in the dataset, estimated using dummy variables for each manager. In the estimation of the sensitivity of *investment* to *lagged Tobin's Q* (*inv to Q sensitivity*) and of the sensitivity of *investment* to *cash flow* (*inv to CF sensitivity*), the primary variables of interest are again the manager fixed effects, estimated using dummy variables for each manager, and their interactions with *lagged Tobin's Q* and *cash flow*, respectively.

The final “actual” dataset consists of 9,074 firm-years, with each firm-year containing 16 financial policy variables and covariates. There are also 338 individual person dummy variables, 30-year dummy variables, 466 firm dummy variables, 466 firm dummy variables interacted with *lagged Tobin's Q*, and another 466 firm dummy variables interacted with *cash flow*. Finally, there are 338 dummy variables each for the interactions between *lagged Tobin's Q* and the manager dummies and for the interactions between *cash flow* and the manager dummies, which are necessary for the analyses of investment sensitivity to *lagged Tobin's Q* and investment sensitivity to *cash flow*. Table 1a reports summary statistics and correlations for the dataset.<sup>9</sup>

Table 1b presents a comparison of our dataset with that of B&S. The descriptive statistics of our sample are sometimes quite different from those of the B&S sample. Two notable differences are in *Tobin's Q* and *leverage*. B&S reported means of 2.4 and 0.35 for *Tobin's Q* and *leverage*, respectively, while our sample means, respectively, are 1.75 and 0.25. These differences in means generate a *t*-stat of 13.10 for *Tobin's Q* and 19.35 for *leverage*. We confirmed that we are using the same definitions as B&S and cannot explain the differences, other than as a result of variation between different samples. On the other hand, of the 14 variables of interest, 9 have sample means that are statistically similar to their corresponding means in the B&S sample.

<sup>9</sup>It is unclear in B&S how certain missing values were treated. For acquisitions, it is reasonable to presume that not reporting a value in a firm-year can be treated as a zero. However, we do not assume that interest, R&D or advertising are zero if a firm does not report them in a given year but treat any such case as a missing observation for the corresponding variable. It is also unclear why in B&S the sample size for the SG&A regression is less than half of the overall sample, whereas in our analysis it is almost the same as the overall sample; we do not make any adjustments to SG&A observations.

TABLE 1 (a) Correlations, replication sample and (b) comparison of replication sample and Bertrand &amp; Schoar sample

(a)	1	2	3	4	5	6	7	8	9	10	11	12	13
1. <i>Investment</i>													
2. <i>Tobins Q</i>	0.10												
3. <i>Cash flow</i>	-0.95	-0.02											
4. <i>N of acquisitions</i>	0.25	0.02	0.03										
5. <i>Leverage</i>	-0.02	-0.24	-0.03	0.02									
6. <i>Interest coverage</i>	0.01	0.12	0.02	0.00	-0.07								
7. <i>Cash holdings</i>	0.10	0.45	-0.02	-0.03	-0.38	0.10							
8. <i>Dividend earnings</i>	-0.01	-0.01	0.00	-0.03	0.01	-0.00	-0.02						
9. <i>N diversifying acq.</i>	0.05	-0.00	0.02	0.79	-0.01	-0.01	-0.03	0.01					
10. <i>R&amp;D</i>	0.15	0.28	-0.20	0.04	-0.15	0.02	0.35	-0.04	0.01				
11. <i>Advertising</i>	0.09	0.11	-0.70	0.03	-0.06	-0.00	0.08	0.01	-0.00	0.87			
12. <i>SG&amp;A</i>	0.03	0.14	-0.08	-0.01	-0.05	0.02	0.16	-0.01	-0.01	0.13	0.38		
13. <i>ROA</i>	-0.10	0.22	0.18	-0.02	-0.14	0.06	-0.05	0.02	-0.02	-0.30	-0.12	-0.18	
14. <i>Operating ROA</i>	-0.09	0.14	0.33	-0.04	-0.06	0.05	0.01	0.08	-0.04	-0.26	-0.21	-0.18	0.78

  

(b)	Replication sample			B&S sample <sup>a</sup>			Replication versus B&S		
	N	Mean	SD	N	Mean	SD	t-stat	p-value	
1. <i>Investment</i>	8,377	0.34	2.05	6,631	0.39	2.94	1.20	.231	
2. <i>Tobins Q</i>	8,082	1.75	1.38	6,631	2.4	3.85	13.10	<.001	
3. <i>Cash flow</i>	8,446	0.30	9.90	6,631	0.44	1.91	1.26	.206	
4. <i>N of acquisitions</i>	3,207	0.88	1.76	6,593	0.77	1.48	-2.93	.004	
5. <i>Leverage</i>	8,721	0.25	0.18	6,563	0.35	0.39	19.35	<.001	
6. <i>Interest coverage</i>	8,274	40.94	641.02	6,278	35	875.1	-0.45	.650	
7. <i>Cash holdings</i>	8,455	0.10	0.14	6,592	0.11	0.16	2.37	.018	
8. <i>Dividend earnings</i>	8,426	0.13	0.81	6,580	0.11	0.79	-1.39	.164	
9. <i>N diversifying acq.</i>	3,207	0.46	1.22	6,593	0.32	1.09	-5.37	<.001	
10. <i>R&amp;D</i>	5,280	0.06	0.17	4,283	0.05	0.07	-3.17	.002	
11. <i>Advertising</i>	3,606	0.05	0.15	2,584	0.05	0.06	-0.49	.623	

TABLE 1 (Continued)

(b)	Replication sample		B&S sample <sup>a</sup>		Replication versus B&S			
	N	Mean	N	Mean	SD	t-stat	p-value	
12. SG&A	7,838	0.26	1,06	2,397	0.26	0.98	0.10	.922
13. ROA	8,709	0.16	0.12	6,593	0.16	0.11	1.01	.314
14. Oper ROA	4,400	0.09	0.11	5,135	0.09	0.12	-1.03	.304

<sup>a</sup>Two-sample t-tests assume for each B&S statistic the same number of observations as in the regressions in tables III and IV of B&S.

TABLE 2 Executive transitions between positions and industries

<b>Panel A: B&amp;S data</b>			
To:	CEO	CFO	Other
<i>From:</i>			
CEO	117 63%	4 75%	52 69%
CFO	7 71%	58 71%	30 57%
Other	106 60%	0	145 42%

  

<b>Panel B: Replication sample</b>			
To:	CEO	CFO	Other
<i>From:</i>			
CEO	48 58%	0	19 47%
CFO	0	6 100%	1 100%
Other	85 59%	36 72%	143 61%

Note: This table summarizes executives' transitions across positions and industries in the manager-firm matched panel data set. All transitions are across firms. The first entry in each cell reports the number of transitions from the row position to the column position. The second line in each cell reports the fraction of the transitions in that cell that are between different two-digit industries.

Moreover, the means reported by B&S only show at most two decimal places, implying that the means of 4 of 14 variables of interest are the same or within rounding error of the means reported by B&S.

Following B&S, our Table 2, Panel B, tabulates the nature of the executive transitions in our overall sample, while Panel A presents the B&S data for comparison. We separate three major executive categories: CEOs, CFOs, and "Others." Of the set of managers identified in our sample, 48 are individuals who move from a CEO position in one firm to a CEO position in another firm; no CEOs move to CFO positions; and 19 are CEOs who move to other top positions. Among the set of executives starting as CFOs, we observe none becoming CEOs, six moving to another CFO position, and one moving to an "other" top position. Finally, among the managers who start in another top position, 85 become CEOs, 36 become CFOs, and 143 move to a different "other" position.

In the second row of each cell in Table 2, we report the fraction of moves that are between firms in different two-digit industries. Just as with the B&S sample, a large fraction of the executive moves in our sample are between industries, though the fractions differ in some of the cells. For example, 47% of our CEO-to-other moves are across different two-digit industries (compared to 69% in B&S), as are 61% of the other-to-other moves (compared to 42% for the B&S data). While the between-industry percentages are not entirely dissimilar to the B&S sample,

our sample partly disaffirms the main observation drawn by B&S (p. 1178) of an intuitive pattern that other-to-other between-industry movers are a relatively lower fraction than CEO-to-CEO or CFO-to-CFO movers because CEOs and CFOs need relatively less industry and firm-specific knowledge.

It is also noteworthy that CEOs and CFOs account for a much higher proportion of movers in the B&S data than in our data. Specifically, B&S find (117 + 4 + 52) 173 moving CEOs and (7 + 58 + 30) 95 moving CFOs. The comparable figures in our data are 67 and 7, respectively. Conversely, in our data, 36 executives move from Other to CFO versus 0 in the B&S data. We suspect the differences may in part reflect how B&S coded executive roles, but without any specific information on how they did so, we cannot be sure.

## 2.4 | Models

### 2.4.1 | Econometric specification

B&S (p. 1179) provide the basic specification that we follow in our analysis. The model examines the effect of individual heterogeneity by controlling for observable firm characteristics: “The nature of our identification strategy can be most easily explained with an example. Consider the dividend payout ratio as the corporate policy of interest. From a benchmark specification we derive residual dividend payouts at the firm-year level after controlling for any average differences across firms and years as well as for any firm-year specific shock, such as an earnings shock, that might affect the dividend payout of a firm. We then ask how much of the variance in these residual dividend payouts can be attributed to manager-specific effects.”

More specifically, for each dependent variable of interest, B&S and we estimate the following regression:

$$y_{it} = \alpha_t + \gamma_i + \beta X_{it} + \lambda_{\text{CEO}} + \lambda_{\text{CFO}} + \lambda_{\text{Others}} + \epsilon_{it}, \quad (1)$$

where  $y_{it}$  stands for one of the corporate policy variables,  $\alpha_t$  are year fixed effects,  $\gamma_i$  are firm fixed effects,  $X_{it}$  represents a vector of time-varying firm-level controls, and  $\epsilon_{it}$  is an error term. The remaining variables in Equation (1) are fixed effects for the managers that we observe in multiple firms. To separately study the effect of CEOs, CFOs, and other top executives on corporate policies, we follow B&S in creating three different groups of manager fixed effects:  $\lambda_{\text{CEO}}$  are fixed effects for the group of managers who are CEOs in the last position in which we observe them,  $\lambda_{\text{CFO}}$  are fixed effects for the group of managers who are CFOs in the last position in which we observe them, and  $\lambda_{\text{Others}}$  are fixed effects for the group of managers who are neither CEOs nor CFOs in the last position in which we observe them.<sup>10</sup>

The statistical significance of the sets of manager fixed effects is evaluated using  $F$ -tests. B&S do not indicate how they calculate  $F$ -stats. We presume they follow standard practice in using a formula based on comparing the fit of the restricted model (without a set of manager

<sup>10</sup>The proportion of executives with role/title changes in the B&S sample and our own raises an interesting question about the nature of the individual manager fixed effect when the title or role (and thus the responsibilities and impact on firm policies) is not fixed across firms, but that analysis is beyond the scope of B&S and this paper.

fixed effects) to that of the unrestricted model (with those manager fixed effects). We use such a formula throughout:

$$F(q, N-k) = \frac{[(R^2_{\text{unrestricted}} - R^2_{\text{restricted}})/q]}{[(1-R^2_{\text{unrestricted}})/(N-k)]}, \quad (2)$$

where  $N$  is the number of observations,  $k$  is the number of variables in the unrestricted model, and  $q$  is the number of restrictions or constraints (i.e., excluded manager fixed effects).<sup>11</sup>

We follow B&S in including the following firm-level controls: For *leverage*, *interest coverage*, *cash holdings*, and *dividends/earnings*, we include *ROA*, lagged *Tobin's Q*, and lagged natural log of assets (*laglnassets*). For *investment*, *inv to Q sensitivity*, *inv to CF sensitivity*, and *N of acquisitions*, we include *cash flow*, lagged *Tobin's Q* and *laglnassets*. For *N of diversifying acquisitions*, we include *ROA*, *cash flow*, natural log of assets (*lnassets*), and a dummy variable indicating whether any acquisitions took place at all during the year. For *R&D*, *advertising*, and *SG&A*, we include *cash flow*, *ROA*, and natural log of assets (*lnassets*).

Investment sensitivity to lagged *Tobin's Q* (*inv to Q sensitivity*) is estimated as an interaction between *investment* and *lagged Tobin's Q*. Investment sensitivity to cash flow (*inv to CF sensitivity*) is estimated as an interaction between *investment* and *cash flow*. Specifically, following B&S, for *inv to Q sensitivity*, we start by regressing *investment* on year fixed effects, *cash flow*, *lagged Tobin's Q*, the lagged logarithm of total assets (*laglnassets*), firm fixed effects, and firm fixed effects interacted with *lagged Tobin's Q*. We then add to this benchmark specification manager fixed effects as well as manager fixed effects interacted with *lagged Tobin's Q*. We analyze *inv to CF sensitivity* in a similar fashion.

#### 2.4.2 | Analysis

Our work requires the analysis of high-dimensional fixed effects models, which can be computationally challenging. Several econometric packages have been developed to aid in the analysis of such models, for example, Gormley and Matsa (2014), Guimaraes and Portugal (2010), and Correia (2015). Due to the size of our datasets and number of fixed effects involved, we rely mainly on the “REGHDFE” Stata ado command created by Sergio Correia.<sup>12</sup>

<sup>11</sup>B&S (p. 1179) say they “account for serial correlation by allowing for clustering of the error term at the firm level.” This comment is hard to interpret because it would normally apply to the calculation of *t*-stats for individual coefficients, but this is not part of B&S’s analysis using Equation (1). It is in principle possible to use a clustered covariance matrix to calculate an *F*-stat for a set of constraints (here, a set of excluded manager fixed effects). But as frequently happens in regressions with many dummy variables, when we tried clustering at the firm level, the rank of the resulting covariance matrix was consistently less than the number of manager fixed effects, making it impossible to test their joint significance. Moreover, using the clustered covariance matrix to test more than a few constraints at a time is known to generate type I errors in finite samples; for instance, Nichols and Schaffer (2007, p. 47) report they performed simulations in which the probability of rejecting the null would rise from 5% to 100% as the number of constraints grew.

<sup>12</sup>This Stata ado-file was published by Sergio Correia on or after 2015, and certainly well after the original B&S work was conducted. To address the concern that somehow different Stata regression commands might have affected our results, we redid a subset of the analyses using XTREG, AREG, and REGRESS, and the results did not materially change.

### 3 | RESULTS

Tables 3–6 present a side-by-side comparison of B&S's original results and our actual replication results, as well as the results of individual randomized (scrambled) datasets and randomized with placebo treatment (placebo) datasets. We present five randomizations (hereafter, “runs”) for each of the scrambled and placebo datasets.<sup>13</sup>

#### 3.1 | Statistical significance

Tables 3 and 4 show the results of 28 regressions (two for each policy variable, one with CEO fixed effects only and one with CEO, CFO, and Other fixed effects), yielding 56 sets of manager fixed effects in total. Using *F*-tests, B&S find 46 of these sets of fixed effects to be statistically significant at the 10% level or better, whereas the corresponding figure for us is 27. Thus, our results are generally somewhat weaker than those of B&S, but we often obtain manager fixed effects that are significant at the 10% level or better when B&S do not. There is no pattern to the discrepancies.

Specifically, our actual results reproduce the original B&S findings for only 1 of 14 dependent variables (*ROA*) in the sense that our findings and B&S's original findings both indicate the same executive types are statistically significant at the 10% level or higher. For another nine dependent variables (*investment*, *investment-to-Q*, *investment-to-cash flow*, *N acquisitions*, *cash holdings*, *dividends-to-earnings*, *N diversifying acquisitions*, *R&D*, *advertising*), our regressions generated fewer sets of executive types whose fixed effects were statistically significant at the 10% level or higher than did B&S; for *investment-to-cash flow*, for instance, B&S found all executive types to be statistically significant at the 10% level, whereas none were for us. For three other dependent variables (*leverage*, *SG&A*, *operating ROA*), we found *more* executive types to be statistically significant at the 10% level than did B&S. For example, B&S found only CEOs to be statistically significant at the 10% level for *operating ROA*, whereas we found *all* executive types (CEOs, CFOs, and Others) to be that significant. Finally, for *interest coverage*, B&S and we both obtained two sets of statistically significant executive types, CEOs in our regressions (twice), and CFOs and Others in B&S's original findings.

More concerningly, when we compare our results using actual data with our results using scrambled and placebo data, we find that the scrambled and placebo data produce results that do not appear to be the result of a well-identified random process. Recall that we have five scrambled runs and five placebo runs for each regression in B&S that we are reproducing. In theory, only 1 in 10 of the sets of manager fixed effects should be statistically significant at the 10% level (and only 1 in 100 at the 1% level, etc.). However, the manager fixed effects in our scrambled and placebo runs are statistically significant far more often than that, often when the run with the actual data does not produce statistically significant manager fixed effects at a comparable level.

In particular, the CEO-only regressions test 14 relationships, one for each dependent variable. Using the actual data, the CEO fixed effects are statistically significant at the 1% level 6 times (42.9%), far more than random. These 14 relationships are tested 5 times for each of the regressions with scrambled data and with the placebo data, that is,  $5 \times 14 = 70$  relationships for

<sup>13</sup>We present the statistics of the individual randomizations rather than averaging them to give readers a sense of how much results varied—or did not vary—for individual “runs.”

TABLE 3 F-tests comparison (p-values and N constraints): B&amp;S results versus replication versus scrambled

	B&S	Actual	Scrambled (random seed #100x)					Placebo (random seed #100x)				
			1	2	3	4	5	1	2	3	4	5
Investment (N obs)	6,631	7,727	7,756	7,756	7,756	7,756	7,756	7,733	7,733	7,733	7,733	7,733
Only CEO (p-values and N constraints)	<.0001 (198)	.942 (124)	.684 (122)	1,000 (121)	.914 (122)	.011 (126)	.904 (121)	.408 (124)	.073 (117)	.007 (124)	.068 (128)	.408 (124)
CEO	<.0001 (192)	.937 (124)	.351 (122)	1,000 (121)	.643 (121)	.002 (126)	.354 (121)	.399 (124)	.071 (117)	.006 (124)	.061 (128)	.152 (124)
CFO	<.0001 (55)	.345 (35)	<.0001 (39)	.516 (39)	.459 (38)	.004 (37)	.044 (35)	.006 (38)	.939 (38)	.440 (38)	<.0001 (37)	.879 (35)
Other	<.0001 (200)	.020 (117)	.571 (131)	.088 (141)	.001 (145)	.473 (140)	.927 (137)	.595 (132)	.231 (141)	<.0001 (143)	.546 (135)	.014 (139)
Inv-to-Q (N obs)	6,631	7,472	7,756	7,756	7,756	7,756	7,756	7,733	7,733	7,733	7,733	7,733
Only CEO	<.0001 (223)	1,000 (120)	.963 (119)	.157 (115)	.999 (116)	<.0001 (124)	<.0001 (116)	.269 (122)	.935 (116)	.223 (119)	.059 (124)	.180 (117)
CEO	<.0001 (221)	1,000 (123)	.955 (122)	.539 (119)	.967 (124)	<.0001 (122)	<.0001 (117)	.231 (121)	.999 (116)	.156 (119)	.004 (124)	.161 (117)
CFO	<.0001 (58)	.168 (30)	<.0001 (35)	.011 (36)	.442 (38)	.009 (35)	.462 (38)	.009 (34)	.996 (38)	.315 (35)	.012 (35)	.156 (30)
Other	<.0001 (208)	<.0001 (105)	.100 (133)	<.0001 (134)	<.0001 (137)	1,000 (130)	.795 (128)	1,000 (125)	.001 (131)	<.0001 (131)	.701 (122)	<.0001 (130)
Inv-to-CF (N obs)	6,631	7,472	7,756	7,756	7,756	7,756	7,756	7,733	7,733	7,733	7,733	7,733
Only CEO	<.0001 (223)	1,000 (119)	.630 (119)	.013 (115)	.999 (118)	<.0001 (122)	.117 (116)	<.0001 (122)	<.0001 (116)	<.0001 (119)	<.0001 (124)	.013 (117)
CEO	<.0001 (221)	1,000 (124)	.701 (124)	.912 (118)	.936 (121)	<.0001 (124)	.267 (124)	<.0001 (122)	<.0001 (116)	<.0001 (119)	<.0001 (124)	<.0001 (117)
CFO	<.0001 (58)	.239 (33)	<.0001 (37)	.293 (35)	.977 (34)	.306 (36)	.011 (35)	.176 (34)	<.0001 (38)	.968 (35)	<.0001 (35)	.112 (31)
Other	<.0001 (208)	.990 (108)	.381 (130)	.076 (132)	<.0001 (136)	.296 (131)	.472 (133)	.494 (125)	.027 (131)	<.0001 (131)	.701 (122)	<.0001 (130)
N acquisitions (N obs)	6,593	2,685	2,713	2,713	2,713	2,713	2,713	2,684	2,684	2,684	2,684	2,684
Only CEO	<.0001 (204)	<.0001 (97)	<.0001 (85)	.847 (105)	.043 (101)	<.0001 (105)	.072 (94)	<.0001 (94)	.948 (103)	.548 (108)	<.0001 (113)	.720 (102)
CEO	<.0001 (199)	<.0001 (95)	<.0001 (81)	.881 (99)	.012 (97)	<.0001 (96)	.324 (91)	<.0001 (92)	.012 (100)	<.0001 (107)	<.0001 (110)	<.0001 (100)
CFO	.0006 (55)	.430 (28)	.010 (33)	<.0001 (29)	.861 (28)	.835 (29)	<.0001 (27)	.003 (33)	<.0001 (32)	.896 (37)	.632 (31)	.588 (29)
Other	<.0001 (203)	.631 (87)	.031 (102)	.544 (97)	.047 (106)	<.0001 (100)	<.0001 (105)	.024 (112)	<.0001 (131)	<.0001 (113)	.013 (107)	<.0001 (114)
Leverage (N obs)	6,563	8,415	8,450	8,450	8,450	8,450	8,450	8,450	8,450	8,450	8,450	8,450
Only CEO	.5294 (203)	<.0001 (131)	<.0001 (127)	<.0001 (127)	<.0001 (123)	<.0001 (129)	<.0001 (129)	<.0001 (127)	<.0001 (125)	<.0001 (127)	<.0001 (130)	<.0001 (129)
CEO	.9190 (199)	<.0001 (131)	<.0001 (127)	<.0001 (127)	<.0001 (123)	<.0001 (129)	<.0001 (129)	<.0001 (127)	<.0001 (125)	<.0001 (127)	<.0001 (130)	<.0001 (129)
CFO	.0225 (54)	>.0001 (36)	>.0001 (42)	<.0001 (39)	<.0001 (40)	>.0001 (38)	<.0001 (37)	<.0001 (41)	<.0001 (39)	<.0001 (40)	<.0001 (38)	<.0001 (37)
Other	.0230 (203)	<.0001 (118)	<.0001 (142)	<.0001 (146)	<.0001 (152)	<.0001 (145)	<.0001 (141)	<.0001 (140)	<.0001 (146)	<.0001 (150)	<.0001 (147)	<.0001 (147)

TABLE 3 (Continued)

	B&S	Actual	Scrambled (random seed #100×)					Placebo (random seed #100×)					
			1	2	3	4	5	1	2	3	4	5	
Int. coverage (N obs)	6,278	8,032	8,064	8,064	8,064	8,064	8,064	8,082	8,082	8,082	8,082	8,082	
Only CEO	.99 (193)	<.0001 (129)	<.0001 (126)	<.0001 (125)	1.000 (121)	<.0001 (128)	1.000 (127)	<.0001 (125)	<.0001 (122)	1.000 (126)	<.0001 (128)	1.000 (128)	
CEO	.99 (192)	<.0001 (129)	<.0001 (126)	<.0001 (125)	1.000 (120)	<.0001 (128)	1.000 (127)	<.0001 (125)	<.0001 (122)	1.000 (126)	<.0001 (128)	1.000 (128)	
CFO	<.0001 (50)	.790 (34)	.690 (41)	.697 (38)	<.0001 (39)	.022 (36)	<.0001 (35)	.696 (38)	.698 (39)	.698 (38)	<.0001 (39)	.423 (36)	<.0001 (34)
Other	<.0001 (192)	1.000 (116)	1.000 (139)	<.0001 (144)	1.000 (148)	1.000 (141)	<.0001 (136)	1.000 (137)	1.000 (145)	.018 (145)	1.000 (141)	<.0001 (143)	
Cash holdings (N obs)	6,592	8,433	8,468	8,468	8,468	8,468	8,468	8,486	8,486	8,486	8,486	8,486	
Only CEO	<.0001 (204)	1.000 (131)	<.0001 (127)	1.000 (127)	<.0001 (123)	<.0001 (129)	<.0001 (129)	.903 (127)	1.000 (125)	<.0001 (127)	<.0001 (130)	<.0001 (129)	
CEO	<.0001 (201)	1.000 (131)	<.0001 (127)	1.000 (127)	<.0001 (123)	<.0001 (129)	<.0001 (129)	.405 (127)	1.000 (125)	<.0001 (127)	<.0001 (130)	<.0001 (129)	
CFO	<.0001 (54)	<.0001 (36)	1.000 (42)	<.0001 (39)	1.000 (40)	<.0001 (38)	<.0001 (37)	<.0001 (41)	<.0001 (39)	1.000 (40)	.708 (38)	<.0001 (37)	
Other	<.0001 (202)	<.0001 (118)	1.000 (40)	<.0001 (145)	<.0001 (145)	<.0001 (145)	<.0001 (145)	<.0001 (140)	<.0001 (146)	<.0001 (150)	.818 (147)	1.000 (147)	
Div/earnings (N obs)	6,580	8,161	8,195	8,195	8,195	8,195	8,195	8,195	8,195	8,195	8,195	8,195	
Only CEO	<.0001 (199)	1.000 (131)	1.000 (131)	1.000 (126)	1.000 (126)	1.000 (123)	1.000 (129)	<.0001 (127)	1.000 (125)	<.0001 (127)	<.0001 (130)	<.0001 (129)	
CEO	3368 (54)	999 (36)	1.000 (42)	1.000 (39)	1.000 (40)	1.000 (38)	1.000 (37)	1.000 (41)	<.0001 (39)	1.000 (40)	1.000 (38)	1.000 (37)	
Other	<.0001 (203)	1.000 (118)	1.000 (146)	1.000 (151)	1.000 (144)	1.000 (141)	1.000 (141)	1.000 (140)	<.0001 (146)	<.0001 (150)	.002 (129)	1.000 (147)	

Note: B&S as reported in original. Replication is from reconstructed sample 1969–1999. Scrambled runs #1–#5 based on randomized executives' association with each firm scrambled datasets. Placebo runs #1–#5 based on randomized datasets where executives are placed at firms 3 years prior to actual start years.

TABLE 4 F-tests comparison (p-values and N constraints): B&amp;S results versus replication versus scrambled datasets

	B&S	Actual	Scrambled (random seed #100x)					Placebo (random seed #100x)				
			1	2	3	4	5	1	2	3	4	5
N diversif. acq. (N obs)	6,593	2,986	3,019	3,019	3,019	3,019	3,019	2,958	2,958	2,958	2,958	2,958
Only CEO (p-values and N constraints)	<.0001 (204)	<.0001 (99)	<.0001 (96)	1,000 (109)	.754 (107)	<.0001 (111)	.755 (103)	<.0001 (101)	1,000 (110)	.850 (113)	<.0001 (120)	.990 (110)
CEO	.016 (202)	<.0001 (93)	<.0001 (89)	1,000 (102)	.996 (103)	<.0001 (102)	.605 (100)	<.0001 (98)	.847 (108)	<.0001 (112)	<.0001 (116)	<.0001 (109)
CFO	<.0001 (53)	.982 (29)	.514 (34)	<.0001 (31)	.999 (31)	.999 (31)	.002 (30)	.662 (35)	<.0001 (34)	.995 (39)	.987 (32)	.816 (32)
Other	<.0001 (202)	1,000 (92)	.994 (108)	.998 (104)	.885 (113)	.720 (108)	<.0001 (111)	.534 (117)	<.0001 (120)	<.0001 (120)	.967 (120)	<.0001 (124)
R&D (N obs)	4,283	5,256	5,275	5,275	5,275	5,275	5,275	5,277	5,277	5,277	5,277	5,277
Only CEO	<.0001 (145)	1,000 (98)	.136 (101)	1,000 (107)	<.0001 (103)	.718 (104)	.805 (108)	.667 (102)	1,000 (107)	<.0001 (108)	<.0001 (103)	<.0001 (108)
CEO	<.0001 (143)	1,000 (96)	.131 (100)	1,000 (107)	<.0001 (103)	.738 (104)	.801 (107)	<.0001 (102)	1,000 (107)	<.0001 (107)	<.0001 (103)	<.0001 (107)
CFO	<.0001 (45)	.927 (34)	<.0001 (32)	.348 (29)	.943 (33)	.860 (29)	<.0001 (25)	<.0001 (31)	<.0001 (29)	<.0001 (32)	<.0001 (28)	<.0001 (27)
Other	<.0001 (143)	.915 (86)	1,000 (104)	.481 (115)	.993 (107)	1,000 (105)	1,000 (106)	1,000 (106)	<.0001 (113)	.298 (107)	.937 (110)	.503 (112)
Advertising (N obs)	2,584	3,583	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600
Only CEO	<.0001 (95)	.991 (71)	<.0001 (78)	.865 (75)	<.0001 (69)	.259 (75)	1,000 (75)	.041 (80)	.024 (77)	.989 (77)	<.0001 (85)	<.0001 (78)
CEO	<.0001 (95)	.976 (70)	<.0001 (76)	.856 (75)	<.0001 (66)	.124 (74)	.998 (74)	<.0001 (80)	.021 (77)	.961 (76)	.024 (85)	<.0001 (78)
CFO	.667 (21)	.074 (22)	.945 (20)	.961 (20)	.276 (19)	.454 (22)	<.0001 (18)	<.0001 (19)	<.0001 (23)	<.0001 (24)	<.0001 (21)	<.0001 (20)
Other	<.0001 (80)	<.0001 (61)	<.0001 (73)	.006 (78)	<.0001 (80)	<.0001 (72)	.037 (84)	<.0001 (81)	<.0001 (88)	.027 (84)	.003 (75)	.044 (87)
SG&A (N obs)	2,397	7,579	7,608	7,608	7,608	7,608	7,608	7,621	7,621	7,621	7,621	7,621
Only CEO	<.0001 (123)	.057 (117)	<.0001 (121)	1,000 (125)	<.0001 (118)	<.0001 (127)	.149 (126)	.989 (120)	1,000 (123)	1,000 (122)	1,000 (125)	1,000 (122)
CEO	<.0001 (118)	.058 (117)	<.0001 (121)	1,000 (125)	<.0001 (118)	<.0001 (126)	<.0001 (125)	.946 (120)	1,000 (123)	1,000 (122)	1,000 (125)	1,000 (122)
CFO	.793 (42)	1,000 (36)	<.0001 (40)	<.0001 (38)	1,000 (38)	.989 (35)	<.0001 (35)	1,000 (38)	1,000 (38)	1,000 (39)	1,000 (36)	1,000 (34)
Other	.978 (146)	<.0001 (113)	1,000 (132)	.351 (137)	1,000 (144)	<.0001 (139)	1,000 (132)	1,000 (133)	1,000 (137)	1,000 (144)	<.0001 (139)	1,000 (141)
ROA (N obs)	6,593	8,705	8,742	8,742	8,742	8,742	8,742	8,760	8,760	8,760	8,760	8,760
Only CEO	<.0001 (217)	<.0001 (132)	<.0001 (130)	<.0001 (129)	<.0001 (129)	<.0001 (129)	<.0001 (132)	<.0001 (130)	<.0001 (130)	<.0001 (131)	<.0001 (131)	<.0001 (132)
CEO	<.0001 (201)	<.0001 (132)	<.0001 (130)	<.0001 (129)	<.0001 (129)	<.0001 (130)	<.0001 (132)	<.0001 (130)	<.0001 (130)	<.0001 (131)	<.0001 (131)	<.0001 (132)
CFO	<.0001 (54)	<.0001 (36)	<.0001 (42)	<.0001 (39)	<.0001 (41)	<.0001 (38)	<.0001 (39)	<.0001 (42)	<.0001 (42)	<.0001 (41)	<.0001 (38)	<.0001 (39)
Other	<.0001 (202)	<.0001 (118)	<.0001 (144)	1,000 (>151)	<.0001 (156)	<.0001 (150)	<.0001 (145)	<.0001 (144)	<.0001 (156)	<.0001 (155)	<.0001 (151)	<.0001 (149)
Op. ROA (N obs)	5,135	4,392	4,429	4,429	4,429	4,429	4,429	4,429	4,429	4,429	4,429	4,429

TABLE 4 (Continued)

	R&S	Actual	Scrambled (random seed #100×)					Placebo (random seed #100×)				
			1	2	3	4	5	1	2	3	4	5
Only CEO	<.0001 (217)	.008 (128)	<.0001 (130)	.256 (127)	<.0001 (129)	.002 (131)	<.0001 (130)	.827 (129)	.014 (130)	.086 (130)	.001 (129)	
CEO	<.0001 (216)	.003 (128)	.001 (130)	.065 (127)	<.0001 (128)	<.0001 (129)	<.0001 (130)	.612 (129)	.004 (130)	.035 (130)	.001 (129)	
CFD	.979 (.58)	<.0001 (36)	>.0001 (41)	<.0001 (39)	<.0001 (40)	<.0001 (37)	<.0001 (39)	.025 (42)	.012 (39)	<.0001 (39)	.001 (37)	.679 (.39)
Other	.454 (217)	<.0001 (118)	.387 (136)	>.0001 (145)	<.0001 (154)	<.0001 (145)	<.0001 (137)	.793 (138)	.001 (143)	.855 (150)	.794 (145)	.217 (141)

Note: R&S as reported in original. Replication is from reconstructed sample 1969–1999. Scrambled runs #1–#5 based on randomized datasets where executives are placed at firms 3 years prior to actual start years.

TABLE 5 Adjusted  $R^2$  comparison: B&S results versus replication versus scrambled datasets

	<b>B&amp;S</b>	<b>Actual</b>	<b>Scrambled (random seed #100x)</b>					<b>Placebo (random seed #100x)</b>				
			<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Investment (N obs)	6,631	7,727	7,756	7,756	7,756	7,756	7,756	7,773	7,773	7,773	7,773	7,773
No executive FE	0.91	0.38	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39
CEO FE	0.94	0.38	0.39	0.38	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39
CEO, CFO, and other FE	0.96	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.40	0.39	0.39
Inv-to-Q (N obs)	6,631	7,472	7,756	7,756	7,756	7,756	7,756	7,773	7,773	7,773	7,773	7,773
No executive FE	0.95	0.59	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58
CEO FE	0.97	0.59	0.58	0.58	0.58	0.60	0.59	0.59	0.59	0.58	0.59	0.58
CEO, CFO, and other FE	0.98	0.60	0.60	0.60	0.60	0.60	0.60	0.59	0.59	0.60	0.59	0.59
Inv-to-CF (N obs)	6,631	7,472	7,756	7,756	7,756	7,756	7,756	7,773	7,773	7,773	7,773	7,773
No executive FE	0.97	0.60	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59
CEO FE	0.98	0.60	0.59	0.59	0.59	0.61	0.60	0.59	0.60	0.60	0.62	0.60
CEO, CFO, and other FE	0.98	0.61	0.60	0.60	0.61	0.61	0.60	0.60	0.62	0.63	0.64	0.62
N acquisitions (N obs)	6,593	2,685	2,713	2,713	2,713	2,713	2,713	2,684	2,684	2,684	2,684	2,684
No executive FE	0.25	0.31	0.31	0.31	0.31	0.31	0.31	0.34	0.34	0.34	0.34	0.34
CEO FE	0.28	0.35	0.34	0.31	0.32	0.35	0.32	0.41	0.33	0.34	0.41	0.34
CEO, CFO, and other FE	0.36	0.34	0.36	0.34	0.32	0.37	0.47	0.42	0.41	0.39	0.42	0.40
Leverage (N obs)	6,563	8,415	8,450	8,450	8,450	8,450	8,450	8,468	8,468	8,468	8,468	8,468
No executive FE	0.39	0.54	0.54	0.54	0.54	0.54	0.54	0.57	0.57	0.57	0.57	0.57
CEO FE	0.39	0.56	0.57	0.56	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55
CEO, CFO, and other FE	0.41	0.58	0.58	0.59	0.59	0.58	0.57	0.57	0.57	0.57	0.57	0.57
Int. coverage (N obs)	6,278	8,302	8,064	8,064	8,064	8,064	8,064	8,082	8,082	8,082	8,082	8,082
No executive FE	0.31	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
CEO FE	0.31	0.12	0.10	0.10	0.09	0.13	0.08	0.11	0.12	0.08	0.14	0.08

TABLE 5 (Continued)

	B&S	Actual	Scrambled (random seed #1000x)					Placebo (random seed #100x)				
			1	2	3	4	5	1	2	3	4	5
CEO, CFO, and other FE	0.41	0.11	0.10	0.11	0.12	0.10	0.10	0.10	0.12	0.13	0.10	0.10
Cash holdings (N obs)	6,592	8,433	8,468	8,468	8,468	8,468	8,468	8,486	8,486	8,486	8,486	8,486
No executive FE	0.77	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68
CEO FE	0.78	0.68	0.69	0.68	0.72	0.72	0.69	0.68	0.68	0.69	0.74	0.69
CEO, CFO, and other FE	0.80	0.72	0.73	0.71	0.72	0.73	0.74	0.74	0.70	0.70	0.74	0.74
Div/earnings (N obs)	6,580	8,161	8,195	8,195	8,195	8,195	8,195	8,195	8,195	8,195	8,195	8,195
No executive FE	0.65	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
CEO FE	0.71	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	0.00	0.00
CEO, CFO, and other FE	0.72	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	0.00	-0.01	-0.02	-0.01	-0.02

Note: B&S as reported in original. Replication is from reconstructed sample 1969–1999. Scrambled runs #1–#5 based on randomized (executives' association with each firm scrambled) datasets. Placebo runs #1–#5 based on randomized datasets where executives are placed at firms 3 years prior to actual start years.

TABLE 6 Adjusted  $R^2$  comparison: B&S results versus replication versus scrambled datasets

	B&S	Actual	Scrambled (random seed #100x)					Placebo (random seed #100x)				
			1	2	3	4	5	1	2	3	4	5
N diversif. acq. (N obs)	6,593	2,986	3,019	3,019	3,019	3,019	3,019	2,958	2,958	2,958	2,958	2,958
No executive FE	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.21	0.21	0.21	0.21	0.21
CEO FE	0.25	0.24	0.28	0.20	0.22	0.28	0.22	0.33	0.20	0.21	0.31	0.20
CEO, CFO, and other FE	0.33	0.22	0.27	0.22	0.31	0.27	0.37	0.33	0.30	0.25	0.30	0.28
R&D (N obs)	4,283	5,256	5,275	5,275	5,275	5,275	5,275	5,277	5,277	5,277	5,277	5,277
No executive FE	0.78	0.55	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
CEO FE	0.79	0.54	0.50	0.50	0.51	0.50	0.50	0.50	0.49	0.54	0.53	0.54
CEO, CFO, and other FE	0.83	0.54	0.50	0.50	0.50	0.50	0.50	0.79	0.77	0.63	0.58	0.57
Advertising (N obs)	2,584	3,583	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600	3,600
No executive FE	0.79	0.77	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
CEO FE	0.81	0.77	0.66	0.65	0.67	0.65	0.65	0.65	0.65	0.65	0.67	0.67
CEO, CFO, and other FE	0.84	0.78	0.67	0.65	0.68	0.67	0.67	0.72	0.69	0.94	0.94	0.94
SG&A (N obs)	2,397	7,579	7,608	7,608	7,608	7,608	7,608	7,621	7,621	7,621	7,621	7,621
No executive FE	0.46	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
CEO FE	0.83	0.31	0.31	0.29	0.36	0.56	0.31	0.30	0.29	0.30	0.29	0.30
CEO, CFO, and other FE	0.83	0.36	0.57	0.36	0.35	0.59	0.57	0.29	0.28	0.28	0.30	0.28
ROA (N obs)	6,593	8,705	8,742	8,742	8,742	8,742	8,742	8,760	8,760	8,760	8,760	8,760
No executive FE	0.72	0.50	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49
CEO FE	0.74	0.51	0.50	0.50	0.50	0.51	0.50	0.51	0.50	0.50	0.50	0.50
CEO, CFO, and other FE	0.77	0.53	0.53	0.52	0.53	0.53	0.53	0.51	0.52	0.52	0.51	0.51
Op. ROA (N obs)	5,135	4,392	4,429	4,429	4,429	4,429	4,429	4,429	4,429	4,429	4,429	4,429
No executive FE	0.34	0.43	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
CEO FE	0.39	0.43	0.42	0.44	0.45	0.43	0.43	0.42	0.42	0.42	0.43	0.43
CEO, CFO, and other FE	0.39	0.45	0.46	0.45	0.46	0.46	0.46	0.43	0.43	0.43	0.42	0.43

Note: B&S as reported in original. Replication is from reconstructed sample 1969–1999. Scrambled runs #1–#5 based on randomized (executives' association with each firm scrambled) datasets. Placebo runs #1–#5 based on randomized datasets where executives are placed at firms 3 years prior to actual start years.

each set of runs, or  $70 + 70 = 140$  in total. In expectation, only 1.4 or 1–2 of the sets of CEO fixed effects across these 140 regressions should be statistically significant at the 1% level. In fact, 34 of the regressions with scrambled data (48.6%) and 36 with the placebo data (51.4%) are significant at the 1% level. Thus, the actual data perform somewhat worse than the randomized data, implying that most of the apparent statistical significance of the CEO fixed effects with the actual data is spurious.

A similar pattern arises with the regressions that include fixed effects for CEOs, CFOs, and Other executives. These 14 regressions test three relationships each, one for each executive role, or  $14 \times 3 = 42$  in total. With the actual data, 17 of these 42 sets of fixed effects (40.5%) are statistically significant at the 1% level. The comparable figures for the 210 sets of fixed effects using the scrambled data and the placebo data are 104 (49.5%) and 112 (53.3%), respectively. In other words, in terms of statistical significance, the randomized data again perform a bit better than the actual data. Indeed, for none of the 56 sets of fixed effects (CEOs alone and CEOs along with other executives) do the actual data yield statistically significant manager fixed effects and all five scrambled runs fail to produce a corresponding set of significant executive fixed effects.

It is interesting that the placebo data slightly outperform the scrambled data in terms of statistical significance. Regardless, we are not aware of an explanation for why scrambled data should produce results that are so similar to those with actual data in terms of statistical significance, other than that the results with actual data are largely a statistical artifact and do not reflect actual differences among managers. A possible source of this statistical artifact may be that the specification over-parameterizes the regressions with “too many” fixed effects, thereby producing spurious results.

### 3.2 | Adjusted $R^2$ (explanatory power)

Tables 5 and 6 repeat the exercise above but for adjusted  $R^2$ —that is, the amount of variance explained by manager fixed effects ( $R^2$  with a penalty, or adjustment, for number of variables). The comparison with B&S mirrors that for statistical significance in that our results are generally weaker but sometimes stronger, with no obvious pattern to the discrepancies. Among 28 regressions with manager fixed effects, B&S obtain a higher adjusted  $R^2$  than we do in 23 of them, and in 18 the discrepancy is large, usually much greater than 0.1; likewise, among the 5 regressions where we obtain a higher adjusted  $R^2$ , the difference is greater than 0.1 in 2 of them.

Notably, the adjusted  $R^2$ 's in B&S are often quite high by any standard, even reaching 0.98 in some cases, a level well in excess of what is typical of research using these data sources and these variables. The consequence is that sometimes B&S's adjusted  $R^2$  is *much* larger than ours; for example, B&S obtain an adjusted  $R^2$  of 0.72 for *Dividends/earnings* when all the executive fixed effects are included, whereas we obtain –0.04.

Again, we turn to a comparison between our results with actual data and our results with scrambled and placebo data. Leaving aside the benchmark regression without manager fixed effects, we have, as before, two regressions for each dependent variable, implying  $2 \times 14 \times 5 = 140$  runs each for the scrambled and placebo data. For the scrambled runs, vis-à-vis the actual data, we obtain the same adjusted  $R^2$  with the scrambled data (to the second decimal place) 47 times, a larger adjusted  $R^2$  38 times, and a lower adjusted  $R^2$  55 times. Thus, the actual data perform slightly better in terms of adjusted  $R^2$  than the scrambled data, but this should not be overstated. Of the 55 times the adjusted  $R^2$  is higher with the actual than with the

scrambled data, only in 10 cases is the discrepancy more than 0.05, all for *advertising*. Even in that case, the adjusted  $R^2$  with the scrambled data is high in absolute terms, ranging from 0.65 to 0.68. Moreover, for the placebo data, the adjusted  $R^2$  is the same as with the actual data 20 times, larger 52 times, and smaller 68 times; of the 68 times the adjusted  $R^2$  is higher with the actual than with the placebo data, the discrepancy is only more than 0.05 in 11 cases (now split between *advertising* and *SG&A*). Again, we find that the placebo data only slightly underperform the actual data in terms of adjusted  $R^2$ . As a practical matter, the explanatory power of manager fixed effects is qualitatively similar, regardless of whether managers actually worked at the firms with which they are associated or not.

We also note that, in principle, adjusted  $R^2$  should decline when “meaningless” variables are added to a regression. Yet, when we add CEOs to the baseline regression without manager fixed effects, we find that adjusted  $R^2$  increases 36 times with the scrambled data and 38 times with the placebo data and only declines 9 times and 12 times, respectively. This again suggests that the apparent explanatory power of manager fixed effects is largely spurious.

### 3.3 | Additional analysis

B&S undertook their own “parametric” robustness test of the persistence of manager effects. First, B&S (p. 1187) “estimate firm-year residuals by regressing the policy variables of interest on firm fixed effects, year fixed effects, and the time-varying firm controls,” and “then collapse [i.e., average] these annual residuals by manager-firm spell.” Second, B&S regress a manager’s average residual at their second firm on their average residual at their first firm, and tabulate the coefficients, standard errors, and  $R^2$ ’s from these bivariate regressions. Third, B&S conduct their own placebo test by following the first two steps just described but replace the manager’s actual average residual at their second firm with the average residual of the 3 years prior to the manager joining their second firm (i.e., when they were not yet there). The idea is that, if managers truly have a style, the placebo bivariate regressions should not generate results that are as statistically or economically significant as the bivariate regressions with actual data, and that is what B&S find.<sup>14</sup>

We replicate this exercise exactly with our actual data but obtain quite different results, as reported in Appendix Table A1. B&S (p. 1187) report that with their actual data, they obtain  $t$ -stats “varying between 4 and 16,” that is, of high statistical significance. By contrast, we do not obtain a single  $t$ -stat that high and only two (*cash holdings* and *R&D*) are statistically significant at the 10% level or higher. B&S generally obtain much higher  $R^2$  than we do, as well. Specifically, B&S obtain a higher  $R^2$  than we do in all 11 cases. The highest  $R^2$  B&S obtain is 0.51 (*dividends/earnings*) versus a high of 0.05 for us (*cash holdings*). Thus, our results are not only statistically and economically much weaker than those of B&S with this alternative analysis but also appear to be close to random.

B&S also note that their placebo results are much weaker, both statistically and economically, than their results with actual data. In our case, the placebo results are statistically similar in that only *cash holdings* and *SG&A* are significant (at the 5% level), and economically actually much stronger for *SG&A*; of the 11 cases, the  $R^2$  is higher with the actual data than with the

<sup>14</sup>B&S (p. 1187) note that this analysis cannot be performed for *investment-to-Tobin's Q* and *investment-to-cash flow* sensitivities. It is not clear, however, why B&S also omit *interest coverage* from this analysis, but we follow them in omitting it.

placebo data seven times, lower three times, and the same (to the third decimal place) one time. These results are similar to what we found earlier: most of the statistical and economic significance of manager effects appears to be noise; specifically, the actual data perform slightly worse than the randomized or placebo data in terms of statistical significance, and slightly better in terms of  $R^2$ .

We also undertook several other analyses. In recent years, for instance, a number of scholars have drawn attention to the fact that different, albeit potentially well-motivated, empirical modeling assumptions may yield quite different results about important social science questions. In consequence, well-known results may not be robust to changing these modeling assumptions (Durlauf, Navarro, & Rivers, 2016; Goldfarb & Yan, 2021; King, Goldfarb, & Simcoe, 2021). Given that our attempt to replicate B&S using their exact modeling assumptions has largely failed, a full “epistemic mapping” of the explanatory power of manager fixed effects across all reasonable specifications would go well beyond the scope of this work.

Nonetheless, in the spirit of epistemic mapping, it is worth asking whether results differ qualitatively using other common definitions or operationalizations of some of the variables. In that regard, it is common in research using Compustat data to impute a value of zero to SG&A and *R&D* if those variables are missing because their absence usually indicates that these expenses are not material at a particular firm (e.g., Dezső & Ross, 2012). Re-estimating manager fixed effects with missing values zeroed out for these two policy variables—besides significantly increasing the number of observations—reduced the adjusted  $R^2$  for SG&A by almost half and resulted in CEOs no longer being statistically significant for SG&A. *Tobin's Q* is a ratio which typically has outlying values, making it reasonable to winsorize them. However, there was no qualitative change from winsorizing *Tobin's Q*. B&S normalize *cash holdings* by dividing by beginning-year net property plant, and equipment, whereas some other authors have divided by total assets or total assets minus cash (Almeida, Campello, & Weisbach, 2004; Opler, Pinkowitz, Stulz, & Williamson, 1999). Using this alternative definition only resulted in reducing the adjusted  $R^2$  of the regressions by nearly half. In summary, using different variable definitions may change the results in some details but does not change the overall implications of our replication analysis.

Compustat, one of the primary sources of data for this paper and B&S, is not stable over time, because it is updated for accounting restatements and corrections. Likewise, ExecuComp reportedly backfills some firm-years when new firms are added to the S&P 1,500. We obtained Compustat data on ExecuComp firms that had been downloaded in 2007, about a decade earlier than the data we used in our replication of B&S and presumably more similar to the data B&S used. We recalculated the exact same variables, using the same variable definitions, for the same firm-years as in the main datasets, but using the 2007 Compustat data as our source instead of the later version of Compustat data. We then re-estimated the regressions that use variables that are constructed entirely from Compustat data (i.e., everything but the regressions related to acquisitions and investment).

For *Leverage*, *Interest Coverage*, *Cash Holdings*, and *Dividends-to-Earnings*, the results were indistinguishable using the 2007 version of Compustat. By contrast, the results for *R&D*, *Advertising*, *SG&A*, *ROA*, and *Operating ROA* did materially change using the older Compustat data. However, there is no pattern to the discrepancies. The older data generate more sets of statistically significant manager fixed effects (at the 10% level) for *R&D* and *Advertising*, whereas the newer data generate more sets for *SG&A*, *Operating ROA*, and *ROA*. Adjusted  $R^2$  is not consistent either. For example, the newer data with all managers generate an adjusted  $R^2$  of 0.54 for *R&D*, whereas the adjusted  $R^2$  is 0.84 with the older data; but, for *ROA*, the newer data

with all managers generate an adjusted  $R^2$  of 0.53, which drops to 0.10 with the older data. In general, the variables did not differ across the samples in terms of basic statistics (e.g., means), but for *ROA* in particular, we did find a large number of mostly small discrepancies between the older and newer data, as well as firm-years that were in the newer data but not the old (perhaps from backfilling firm-years by ExecuComp). All told, this instability in the results again suggests that the large number of fixed effects in these regressions is overfitting the data, producing results that are more the result of noise than of actual differences among managers.

We also experimented with a different time period, 2000–2013. These results are presented in Appendix Tables A2 (statistical significance) and A3 (adjusted  $R^2$ ). It is notable how much the results from the later period differ in detail from our results with the 1969–1999 data. Recall that we examine the statistical significance of 52 sets of manager fixed effects, or four with respect to each dependent variable, comprising 2 sets of CEO fixed effects (one in CEO-only regressions and one with other managers) and fixed effects for CFOs and for other managers. With the 1969–1999 data, 27 of these sets of manager fixed effects were significant at the 10% level or better; of these, only 15 of the corresponding fixed effects are significant with the 2000–2013 data. In total, only 19 sets of manager fixed effects are significant at the 10% level with the 2000–2013 data, perhaps because the later time period was shorter. Even so, for 4 of these 19 sets of fixed effects, the corresponding fixed effects were not significant with the 1969–1999 data. In general, there is little consistency across the two time periods.

For 3 of the 28 regressions with manager fixed effects, the 1969–1999 data yield an adjusted  $R^2$  that is at least 0.15 higher than do the 2000–2013 data, whereas for 11 of the 28 regressions, the 2000–2013 data yield an adjusted  $R^2$  that is at least 0.15 higher than do the 1969–1999 data. In other words, the explanatory power of manager fixed effects is not consistent across the two time periods either, nor is there an obvious pattern to the discrepancies. Some of the differences appear to be driven by the adjusted  $R^2$  of the baseline model without manager fixed effects, but frequently that is not so. As one example, consider *interest coverage*. With the 1969–1999 data, we obtain an adjusted  $R^2$  of 0.09 without manager fixed effects, 0.12 with CEOs only, and 0.11 with all managers, whereas with the 2000–2013 data, the corresponding figures are 0.13, 0.09, and 0.02. In other words, adding manager fixed effects has mixed implications for explanatory power in the first time period, but *decreases* it in the later time period.

## 4 | DISCUSSION

We attempted to reproduce the main analyses of B&S using the same data. In general, our results were quite different, usually less significant both statistically and economically, but sometimes more so. We then undertook falsification tests in which the match between firm and manager was effectively random. We did so two different ways to allay any concerns over even a residual connection between firms and managers. The results with the randomized data were statistically and economically quite similar to those with the real data. In general, in terms of both statistical significance and explanatory power (adjusted  $R^2$ ), our results with randomized data were too strong to be the product of random chance.

We do not know why we failed utterly to replicate B&S's findings. One possibility is that the inevitable small differences between our dataset and theirs interacted with a propensity for regressions with a large number of fixed effects to generate spurious results. The disparate results we obtained in our analysis of a later time period and with an older draw of Compustat suggest this may be the case. Yet, if random differences in datasets were the entire story, we

would expect our results to differ from, but be approximately as strong as, those of B&S. In fact, our results were generally weaker.

Of course, one possibility is that we made a mistake in the compilation of our data or in the implementation of B&S's MDV methodology. Although human error on our part could certainly explain our failure to replicate B&S's results, it does not obviously explain why our results with randomized data were so strong. One would normally expect that a failure on our part to implement MDV or some other aspect of B&S's methodology would introduce even more randomness into our regressions with randomized data, making the results weaker, not stronger. The implication is again that the economic and statistical significance of manager fixed effects may be an artifact of the estimation procedure rather than a reflection of managerial style.

As noted previously, a growing number of authors are using MDV or similar methodologies like AKM either to perform analyses like those of B&S or to use the pool of fixed effects generated thereby as a dependent variable in subsequent analyses. (B&S also regress the manager fixed effects they derive on corporate practices, managerial compensation, and governance measures.) If these individual fixed effects are largely noise, analyses that use them as dependent variables may be studying noise rather than the constructs of interest.

More broadly, fixed effects regression has become a popular estimation procedure—even the default—in many social sciences and related fields over the past couple of decades. A principal benefit of fixed effects with panel data is that they control for time-invariant unobservable heterogeneity. However, a growing number of scholars have started to argue that, vis-à-vis less parameterized specifications, fixed effects may be vulnerable to measurement error and reverse causality and rely on unrepresentative, small subsamples that exhibit within-panel variation (Choi, 2013; Collischon & Eberl, 2020; Hill, Davis, Roos, & French, 2020). Our results suggest another problem: regressions with a large number of fixed effects may overfit the data. Exploring this issue is a ripe topic for future research on methods in strategy and related fields.

All that said, we wish to emphasize that our results should not be construed as showing that managers do not matter. First, we acknowledge “a single replication cannot overturn prior evidence any more than a single empirical study can establish that we ‘know’ something with certainty” (Ethiraj, Gambardella, & Helfat, 2016). Second, a rich and exhaustive literature dating at least from Barnard (1938) and using all manner of empirical methodologies, including direct observation, has documented the importance of managers for firm policy and performance. Indeed, one of the most influential theories in the field of strategy is the Upper Echelons Perspective, which holds that a firm is a “reflection of its top managers” (Hambrick & Mason, 1984, p. 193); the empirical support for this proposition is legion (Westphal & Zajac, 2013).

Perhaps, we may find a hint in the observation that strategies that work well for some firms, do not work well for others. For example, Indian Motorcycle tried to compete with Harley Davidson by offering an everyday motorcycle for non-motorcycle enthusiasts, failed, and went bankrupt; yet, Honda succeeded spectacularly with the same strategy almost immediately afterwards (Makadok & Ross, 2018).

A related potential problem is that the same policy may have different implications depending on the other policies that a firm (or top manager) pursues, and failing to take that into consideration can lead to false inferences. An example may be the argument that business groups primarily serve to expropriate wealth, which was given empirical support in a well-known paper by Bertrand, Mehta, and Mullainathan (2002) using Indian data. This empirical context and research question were revisited by Siegel and Choudhury (2012), who showed that after correcting some data and statistical issues in the original work and, importantly,

accounting for differences in firm strategies, the original finding was reversed: rather than serving to expropriate wealth, Indian business groups generate scale and scope efficiencies.

One may infer that imitating readily observable policies of successful managers, especially in isolation, is insufficient to match their success. Something deeper and more tacit must be going on in the way that managers add value to their firms, and this may be hard to distill to a small set of readily imitable firm policy variables. Indeed, if it could be so distilled, other firms would readily imitate the underlying strategy.

That said, it is possible a more refined analysis of manager fixed effects might yield a stronger contrast between real and randomized data. Fee et al. (2013) report that endogeneity can drive the finding (or non-finding) of explanatory power for managerial fixed effects. B&S (p. 1176) themselves conjecture, “one could argue that [our] required focus on larger firms may in fact bias ... results against finding systematic effects of managers on firm policies. Indeed, a specific individual might be more influential in a smaller organization that requires more personal involvement of the top managers in day-to-day activities.” Relatedly, a recent work has focused directly on this “star power” effect in a theater setting, finding that “only the most talented ... significantly affect financial success” (Han & Ravid, 2020).

Yet another possibility is that managerial style changes over time. As the fresh ideas of youth give way to the wisdom of experience, managers may change their style in systematic ways, for example, in response to parenthood (Dahl, Dezső, & Ross, 2012) or to buying a large home (Liu & Yermack, 2012). Or managers may have a distinct style but only in a distinct role, that is, a given manager may perform the CFO job in one way and the CEO job in another. The B&S methodology would miss this subtlety because it uses a manager’s last job in the data to classify the manager throughout. A final alternative is that systematic variation in managerial style could exist at the cohort level because of common imprinting events in the past (Marquis & Tilcsik, 2013). Research has already demonstrated that communist education may reduce entrepreneurs’ propensity to internationalize (Marquis & Qiao, 2020) and that experience of low returns from investing reduces willingness to take financial risk (Malmendier & Nagel, 2011). Could experience with a major event like the Great Depression systematically affect managers? This area would appear to be fertile ground for research that is both theoretically grounded and relevant to practice, in a way that mechanical calculation of the statistical and economic significance of manager fixed effects by itself is not.

In summary, we attempted to replicate B&S, but we obtained quite different (usually weaker) results. We also obtained results with randomized data that were quite similar to the results we obtained with actual data and, in any case, too statistically and economically strong to have plausibly arisen by chance from an unbiased random process. Although our analysis only “... alters the balance of the evidence” (Ethiraj et al., 2016), it does suggest that MDV and similar methodologies that use a large number of fixed effects to estimate the impact of a population (e.g., of firms or individuals) on a variable of interest may generate spurious results. We hope our work will inspire future work to identify the best methods for analyzing managerial impact.

## ACKNOWLEDGEMENTS

This paper has benefitted from the comments of many colleagues. We would especially like to thank Sendil Ethiraj (editor), two anonymous reviewers, J. P. Eggers, Orie Shelef, and Todd Zenger for their helpful comments. We thank Peter Z. McKay (business librarian) and Ayoung

Jang (research assistant) for their excellent help in compiling the Forbes 800 files. The usual disclaimer applies.

## DATA AVAILABILITY STATEMENT

We are happy to share our data with other researchers.

## ORCID

Victor Esteban Jarosiewicz  <https://orcid.org/0000-0001-6275-3217>

David Gaddis Ross  <https://orcid.org/0000-0001-6561-6234>

## REFERENCES

- Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2), 251–333.
- Almeida, H., Campello, M., & Weisbach, M. S. (2004). The cash flow sensitivity of cash. *Journal of Finance*, 59(4), 1777–1804.
- Aobdia, D., Lin, C.-J., & Petacchi, R. (2015). Capital market consequences of audit partner quality. *The Accounting Review*, 90(6), 2143–2176.
- Barnard, C. (1938). *The functions of the executive*. Cambridge: Harvard University Press.
- Barney, J. B. (1986). Strategic factor markets: Expectations, luck, and business strategy. *Management Science*, 32, 1231–1241.
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1), 249–275.
- Bertrand, M., Mehta, P., & Mullainathan, S. (2002). Ferreting out tunneling: An application to Indian business groups. *Quarterly Journal of Economics*, 117(1), 121–148.
- Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics*, 118(4), 1169–1208.
- Blettner, D. P., Chaddad, F. R., & Bettis, R. A. (2012). The CEO performance effect: Statistical issues and a complex fit problem. *Strategic Management Journal*, 33, 986–999.
- Brandenburger, A., & Nalebuff, B. (1996). *Co-opetition*. New York: Doubleday.
- Brookman, J. T., & Thistle, P. D. (2013). Managerial compensation: Luck, skill or labor markets? *Journal of Corporate Finance*, 21, 252–268.
- Castanias, R. P., & Helfat, C. E. (1991). Managerial resources and rents. *Journal of Management*, 17(1), 155–171.
- Castanias, R. P., & Helfat, C. E. (2001). The managerial rents model: Theory and empirical analysis. *Journal of Management*, 27(6), 661–678.
- Cavaco, S., Crifo, P., Rebérioux, A., & Roudaut, G. (2017). Independent directors: Less informed but better selected than affiliated board members? *Journal of Corporate Finance*, 43, 106–121.
- Choi, I. (2013). Spurious fixed effects regression. *Oxford Bulletin of Economics and Statistics*, 75(2), 305–9049.
- Coles, J. L., & Li, Z. F. (2020). Managerial attributes, incentives, and performance. *Review of Corporate Finance Studies*, 9, 256–301.
- Collischon, M., & Eberl, A. (2020). Let's talk about fixed effects: Let's talk about all the good things and the bad things. *Köln Z Soziol*, 72, 289–299.
- Correia, S. (2015). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. *Statistical Software Components*.
- Dahl, M. S., Dezső, C. L., & Ross, D. G. (2012). Fatherhood and managerial style: How a male CEO's children affect the wages of his employees. *Administrative Science Quarterly*, 57(4), 669–693.
- DeJong, D., & Ling, Z. (2013). Managers: Their effects on accruals and firm policies. *Journal of Business Finance & Accounting*, 40(1–2), 82–114.
- Dezső, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, 33(9), 1072–1089.
- Dezső, C. L., Ross, D. G., & Uribe, J. N. (2016). Is there an implicit quota on women in top management? A large-sample empirical analysis. *Strategic Management Journal*, 37(1), 98–115.

- Durlauf, S. N., Navarro, S., & Rivers, D. A. (2016). Model uncertainty and the effect of shall-issue right-to-carry laws on crime. *European Economic Review*, 81, 32–67.
- Ethiraj, S. K., Gambardella, A., & Helfat, C. E. (2016). Replication in strategic management. *Strategic Management Journal*, 37(11), 2191–2192.
- Fee, C. E., Hadlock, C. J., & Pierce, J. R. (2013). Managers with and without style: Evidence using exogenous variation. *Review of Financial Studies*, 26(3), 567–601.
- Fitza, M. A. (2014). The use of variance decomposition in the investigation of CEO effects: How large must the CEO effect be to rule out chance? *Strategic Management Journal*, 35, 1839–1852.
- Fitza, M. A. (2017). How much do CEOs really matter? Reaffirming that the CEO effect is mostly due to chance. *Strategic Management Journal*, 38, 802–811.
- Goldfarb, B., & Yan, L. (2021). Revisiting Zuckerman's (1999) categorical imperative: An application of epistemic maps for replication. *Strategic Management Journal*, 42, 1963–1992.
- Gormley, T. A., & Matsa, D. A. (2014). Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies*, 27(2), 617–661.
- Graham, J. R., Li, S., & Qiu, J. (2012). Managerial attributes and executive compensation. *Review of Financial Studies*, 25(1), 144–186.
- Guimaraes, P., & Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. *Stata Journal*, 10(4), 628–649.
- Gul, F. A., Wu, D., & Yang, Z. (2013). Do individual auditors affect audit quality? Evidence from archival data. *The Accounting Review*, 88(6), 1993–2023.
- 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.
- Han, S., & Ravid, S. A. (2020). Star turnover and the value of human capital: Evidence from Broadway shows. *Management Science*, 66(2), 958–978.
- Hill, T. J., Davis, A. P., Roos, J. M., & French, M. T. (2020). Limitations of fixed-effects models of panel data. *Sociological Perspectives*, 63(3), 357–369.
- Hillier, D., Korczak, A., & Korczak, P. (2015). The impact of personal attributes on corporate insider trading. *Journal of Corporate Finance*, 30, 150–167.
- King, A., Goldfarb, B., & Simcoe, T. (2021). Learning from testimony on quantitative research in management. *Academy of Management Review*, 46(3), 465–488.
- Lieberman, M. B., & Montgomery, D. B. (1988). First-mover advantages. *Strategic Management Journal*, 9(1), 41–58.
- Lieberson, S., & O'Connor, J. F. (1972). Leadership and organizational performance: A study of large corporations. *American Sociological Review*, 37, 117–130.
- Liu, C., & Yermack, D. (2012). Where are the shareholders' mansions? CEOs' home purchases, stock sales, and subsequent company performance. In S. Boubaker, B. D. Nguyen, & D. K. Nguyen (Eds.), *Corporate governance: Recent developments and new trends*. Berlin, Heidelberg: Springer Science & Business Media.
- Makadok, R. (2011). The four theories of profit and their joint effects. *Journal of Management*, 37(5), 1316–1334.
- Makadok, R., & Ross, D. G. (2018). Losing by winning: The danger zone of adverse competitor replacement. *Strategic Management Journal*, 39(7), 1990–2013.
- Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*, 126(1), 373–416.
- Marquis, C., & Qiao, K. (2020). Waking from Mao's dream: Communist ideological imprinting and the internationalization of entrepreneurial ventures in China. *Administrative Science Quarterly*, 65(3), 795–830.
- Marquis, C., & Tilcsik, A. (2013). Imprinting: Toward a multilevel theory. *Academy of Management Annals*, 7(1), 195–245.
- McGahan, A. M., & Porter, M. E. (1997). How much does industry matter, really? *Strategic Management Journal*, 18, 15–30.
- Murphy, K. J. (2013). Executive compensation: Where we are, and how we got there. In G. M. Constantinidis, M. Harris, & R. Stulz (Eds.), *Handbook of the economics of finance*. Amsterdam, Netherlands: Elsevier.

- Nichols, A., & Schaffer, M. (2007). The cluster-robust variance-covariance estimator: A (stata) practitioner's guide.
- Opler, T., Pinkowitz, L., Stulz, R., & Williamson, R. (1999). The determinants and implications of corporate cash holdings. *Journal of Financial Economics*, 52, 3–46.
- Penrose, E. T. (1959). *The theory of the growth of the firm*. Oxford: Basil Blackwell.
- Porter, M. E. (1980). *Competitive strategy: Techniques for analyzing industries and competitors*. New York, NY: Free Press.
- Quigley, T. J., & Graffin, S. D. (2017). Reaffirming the CEO effect is significant and much larger than chance: A comment on Fitzmaurice (2014). *Strategic Management Journal*, 38, 793–801.
- Rumelt, R. P. (1991). How much does industry matter? *Strategic Management Journal*, 12(3), 167–185.
- Shaver, J. M. (1998). Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival? *Management Science*, 44(4), 571–585.
- Siegel, J., & Choudhury, P. (2012). A reexamination of tunneling and business groups: New data and new methods. *Review of Financial Studies*, 25(6), 1763–1798.
- Villalonga, B. (2004). Diversification discount or premium? New evidence from the *Business Information Tracking Series*. *Journal of Finance*, 59, 479–506.
- Westphal, J. D., & Zajac, E. J. (2013). A behavioral theory of corporate governance: Explicating the mechanisms of socially situated and socially constituted agency. *Academy of Management Annals*, 7(1), 607–661.
- Withers, M. C., & Fitzmaurice, M. A. (2017). Do board chairs matter? The influence of board chairs on firm performance. *Strategic Management Journal*, 38(6), 1343–1355.
- Zorn, D. M. (2004). Here a chief, there a chief: The rise of the CFO in the American firm. *American Sociological Review*, 69, 345–364.

## 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:** Jarosiewicz, V. E., & Ross, D. G. (2023). Revisiting managerial “style”: The replicability and falsifiability of manager fixed effects for firm policies. *Strategic Management Journal*, 44(3), 858–886. <https://doi.org/10.1002/smj.3439>