The Review of Financial Studies

Symmetry in Pay for Luck

Naveen D. Daniel
LeBow College of Business, Drexel University

Yuanzhi Li
Fox School of Business, Temple University

Lalitha Naveen
Fox School of Business, Temple University

In this study, we take a comprehensive look at asymmetry in pay for luck, which is the finding that CEOs are rewarded for good luck, but are not penalized to the same extent for bad luck. Our main takeaway, which is based on over 200 different specifications, is that there is no asymmetry in pay for luck. Our finding is important given that the literature widely accepts the idea of asymmetry in pay for luck and typically points to this as evidence of rent extraction. (*JEL* G32, G34)

Received 11 February, 2016; editorial decision 10 March, 2019 by Editor Francesca Cornelli. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

Executive compensation has attracted widespread scrutiny from academics, practitioners, and regulatory authorities. Indeed, as Murphy (2013) notes, “scrutinizing, criticizing, and regulating high levels of executive pay has been an American pastime for nearly a century.” Given the intense scrutiny and debate on compensation, it is critical for researchers to understand the extent to which contracts are set optimally.

In this study, we focus on understanding one important aspect of executive compensation: asymmetry in pay for luck. Prior literature finds that chief executive officers (CEOs) are rewarded for good luck but not penalized to the same extent for bad luck, and uses the term “asymmetry in pay for luck” to describe this phenomenon (see, e.g., Garvey and Milbourn 2006). Our central

The authors are grateful for helpful comments from three anonymous referees, Francesca Cornelli (editor), Viral Acharya, Rajesh Aggarwal, Jay Cai, Jeff Coles, Wayne Ferson, Eli Fich, Gerald Garvey, Radhakrishnan Gopalan, Swaminathan Kalpathy, Xi Li, Connie Mao, Todd Milbourn, Kevin Murphy, Federico Nardari, Oleg Rytchkov, Fenghua Song, Ralph Walkling, Jide Wintoki, and David Yermack; and participants at Financial Management Association Conference in Chicago, CFEA Conference at USC, Drexel University, Penn State University at Harrisburg, SUNY Binghamton, Temple University, and Villanova University. An earlier version of this paper was titled “Asymmetry in Pay for Luck: A Size Effect?” Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Lalitha Naveen, Fox School of Business, Temple University, 1801 Liacouras Walk, Philadelphia PA 19122; telephone: 215-204-6435. E-mail: lnaveen@temple.edu

The Review of Financial Studies 33 (2020) 3174–3204

doi:10.1093/rfs/hhz057 Advance Access publication September 21, 2019

finding is that there is no asymmetry in pay for luck. This finding is robust to using: several plausible decompositions of performance into luck and skill; several plausible luck factors; different measures of pay; various accounting and/or stock-based performance measures that are relevant for pay; different performance periods that are relevant for pay; nonlinear specifications; and different subsamples based on governance. Our finding is important given that the literature largely assumes that asymmetry in pay for luck exists, and that it is symptomatic of rent extraction (see, for instance, Frydman and Jenter 2010, p. 16).[1]

To help motivate why it is important to revisit the finding of asymmetry in pay for luck, we first present the two-stage methodology used in the literature to examine asymmetry.

First-stage regression: Firm Performance measure = Function {Luck Factors}

Second-stage regression: Pay measure = Function {Luck, Luck × Bad Luck, Controls}

That is, first, a firm's performance is regressed on luck factor(s) to decompose performance into luck and skill. Second, a measure of pay is regressed on luck estimated from the first step, luck interacted with "bad" luck (an indicator variable that equals one if luck is bad), and other control variables. A negative coefficient on the interaction term implies asymmetry in pay for luck. Finally, studies examine whether the asymmetry is greater in subsamples where we might expect rent extraction to dominate.

Overall, across the two regressions, researchers need to make at least 17 decisions. Each decision has several reasonable choices. For instance, consider the decision involving the methodology to decompose performance into luck and skill. Researchers may choose between several possible options, such as whether to estimate a pooled regression or estimate regressions separately for each CEO. Other decision variables include luck factors, performance measures, relevant performance windows, pay measures, etc. Thus, we need to ensure that any finding relating to asymmetry in pay for luck is robust across an entire range of reasonable choices.

In making our choices, we start with the literature on pay-performance sensitivity and, its extension, the literature on relative performance evaluation (RPE).[2] The central idea of RPE is that performance depends on CEO effort as well as luck; shareholders should pay the CEO for performance net of any observable luck. In other words, if we interpret CEO effort or actions as skill, the focus of RPE literature is on pay for skill rather than on pay for performance. The pay-for-luck literature, on the other hand, focuses more on the second aspect of performance: luck. The literature on asymmetry in pay-for-luck extends

1 Asymmetry may also be consistent with optimal contracting (e.g., Gopalan, Milbourn, and Song 2010; Bizjak, Lemmon, and Naveen 2008).

2 We are grateful to one of our referees for suggesting that we start from first principles by going back to the RPE literature.

this idea still further and examines whether pay for luck varies in a nonlinear fashion. The asymmetry literature, therefore, is simply an extension of RPE.

We start by describing our choices for our baseline regressions and providing the rationale for these choices. At every point, we discuss the key alternative choices we could make, and examine the robustness of our results to these choices. Our first regression has at least nine decisions. The purpose of this first stage is to decompose the relevant firm performance measure into luck and skill. The starting point is an assumption about how firm performance evolves; we borrow from Bertrand and Mullainathan (2001, p. 904):

$$\begin{aligned} \textit{Firm Performance} = &\textit{CEO's Actions} + \textit{Sensitivity to Luck Observable Luck} \\ &+ \textit{Unobservable Noise}. \end{aligned} \tag{1}$$

We measure performance using stock returns, as is typical in the literature. Our first decision is to choose the luck factors: we use the firm's industry returns and market returns. We use market returns to capture the impact of any factor orthogonal to the industry factor that could affect firms' returns. In this regard, we follow the RPE literature (Gibbons and Murphy 1990). Our results are robust to using just the industry returns or using the four factors employed by Carhart (1997).[3]

Second, we assume CEO skill within each industry is i.i.d. with mean zero. Therefore, we equal-weight firms' returns to arrive at industry and market returns. Our results are robust to using value-weighted returns.

Third, the assumption of i.i.d. skill with mean zero within each industry suggests that we should include the firm's returns in industry returns. Otherwise, average skill in the firm's industry return will not equal zero, and the industry return cannot be viewed as a luck factor (Section 3.1.3 discusses this in more detail).[4] Our results are robust to excluding the firm's returns from the industry returns.

Fourth, we use the sample of *Execucomp* firms as the firm's peer group to compute industry returns. This is because peers are typically firms of similar size in the same industry, and *Compustat* firms may be too small to be considered peers. Our results are robust to including all firms on *Compustat* to compute industry returns.

Fifth, we use firms within the same 2-digit SIC as peer firms. Our results are robust to using all firms in the same text-based network industry classification (Hoberg and Phillips 2010).

Sixth, we implement the firm-performance decomposition using time-series regressions for each CEO-firm observation. We use the word "executive"

3 When we say "results are robust" we mean that, like in the baseline, we find no negative asymmetry as documented in the literature. In some cases, we find the opposite result: the coefficient of the interaction term (*Luck* × *Bad Luck*) is significantly positive.

4 We confirm, empirically, that the mean skill across executives within an industry is insignificantly different from zero for all 58 industries (at the 2-digit SIC level) in our sample.

to refer to the unique CEO-firm combination identified by *co_per_rol* in *Execucomp* database. The time-series (as opposed to cross-sectional or pooled) regressions allows for the sensitivity of firms' returns to luck factors to vary across executives and thus provides a better estimate of skill. Moreover, the intercept has the natural interpretation as the average skill for each executive. The mutual fund literature interprets the intercept from such a time-series regression (of portfolio returns on risk factors) as the fund manager's alpha (Carhart 1997). We define skill for a given year as the intercept plus the residual for the year from the first-stage regression. In all cases, by construction, luck plus skill equals firm returns. Thus, luck is the firm returns net of skill. We base our choice of CEO-firm-specific regressions (rather than CEO-specific or firm-specific regressions) on Gabaix and Landier (2008). They argue that CEOs have different talents and are competitively matched to firms. Thus, a given executive's ability to generate value for a firm will depend on the firm that employs the executive.

Our results are robust to using several alternative methods to decompose performance into luck and skill: (1) Pooled regression using all firms: in this approach (followed by most of the asymmetric-pay-for-luck literature), the implicit assumption is that industry and market sensitivities are the same for all firms. As with prior literature, we define skill as the residual (because there is only one intercept in the pooled regression). (2) Pooled regression but with executive fixed effects: here, skill is the executive fixed effect plus the residual. (3) Regression by executive-year (rather than by executive): this approach allows the industry and market sensitivities to vary not only in the cross-section but also over time. We use the estimated industry and market sensitivity from the prior year along with contemporaneous factor realizations to estimate the luck and skill for current year. (4) Executive-specific regressions (as in our baseline): but we define skill as the residual (and not the intercept plus residual). This assumes that the intercept does not reflect the skill of the CEO but, rather, the net impact of some omitted luck factor on returns. (5) Executive-specific regressions (as in our baseline): but we define skill as the intercept (and not the intercept plus residual). This assumes that CEO skill is fixed over the tenure of the CEO in that firm and the residuals are simply noise. (6) No regression: here, luck is the industry return and skill is the firm's return in excess of its industry return. This approach has been adopted in the early RPE literature (Gibbons and Murphy 1990).

Seventh, because we estimate the first-stage regressions for each executive, we include all firms regardless of fiscal year end. Previous studies limit their sample to firm-years with a December fiscal year end because they use pooled regressions and, therefore, need the performance periods to be comparable across firms. Our results are robust if we consider only firms with December fiscal year end.

Eighth, in our baseline regressions, we use monthly returns. This provides reasonably precise estimates of firms' sensitivity to industry and market factors. Our results are robust if, like in the prior literature, we use annual returns.

Ninth, in our baseline, we winsorize returns at the 1st and 99th percentile levels to reduce the impact of outliers and potential data errors.[5] Our results are robust to using unwinsorized returns.

Next, we describe the eight decisions for our baseline second-stage regression of CEO pay. First, we use change in pay as the dependent variable (Garvey and Milbourn 2006). Our results are robust if we consider (1) the log of the ratio of current year's pay to prior year's pay, which is the logarithm of the rate of change in pay (Gibbons and Murphy 1990), and (2) the logarithm of the level of pay (Bertrand and Mullainathan 2001).

Second, we use stock returns as our baseline measure of firm performance. Our results are robust if we consider accounting returns instead. Our results are also robust if we include accounting returns in addition to stock returns.

Third, we use luck and skill estimated as rates of return, consistent with much of the prior literature on RPE (Gibbons and Murphy 1990) and pay for luck (Bertrand and Mullainathan 2001). Our results are robust if we use dollar values instead (Garvey and Milbourn 2006), by multiplying our luck and skill measures by the firm's lagged market capitalization.

Fourth, we include firm size as an additional variable (relative to prior literature) in the pay regression. Our results are robust to excluding firm size.

Fifth, we consider luck and skill estimated over the fiscal year to match up with pay that is awarded over the fiscal year. Our results are robust if we consider alternatively: (1) performance over the 12 months prior to the largest grant date and (2) performance over a cumulative 3-year period (current year plus 2 prior years).

Sixth, we include the cdf of the variance of luck and skill and the interaction of these terms with luck and skill [*cdf (Variance of Luck)*, *cdf (Variance of Skill)*, *Luck* $\times$ *cdf (Variance of Skill), Skill* $\times$ *cdf (Variance of Skill)*]. This is based on Aggarwal and Samwick (1999), who show that pay-performance sensitivity is negatively related to variance of performance; in other words, pay is negatively related to performance interacted with variance of performance. Our results are similar if we drop the cdf terms and the interactions.

Seventh, our results are robust to a parsimonious specification that excludes terms associated with skill [*Skill*, *Bad Skill*, *cdf (Variance of Skill)*, *Skill* $\times$ *cdf (Variance of Skill)*].

Eighth, in our baseline, we explore asymmetry by considering bad luck to be realizations of luck below zero. For robustness, we consider extreme bad luck. Specifically, we redefine bad luck as luck below the 20th percentile (we also consider the 10th percentile). Alternatively, to allow for general nonlinearity

[5] Firms' winsorized returns are used to compute industry returns. We do not winsorize industry and market returns given that these are portfolio returns.

without specifying the exact kink in the distribution, we include the square of luck (in lieu of luck interacted with bad luck). Our finding of no asymmetry holds in these nonlinear specifications as well.

Our discussion to this point shows that our main finding of no asymmetry is robust to changing the baseline one dimension at a time. Some researchers may argue for an alternative baseline specification. We explore this by using five variations of the baseline specification: (1) we use pooled regression (instead of regressions by executive) in the first stage to effect the decomposition of stock returns to luck and skill, (2) we use value-weighted (instead of equal-weighted) industry and market returns in the first-stage regression, (3) we use *log(Pay)* instead of ΔPay as the dependent variable in the second-stage regression, (4) we use $\Delta log(Pay)$ instead of ΔPay as the dependent variable in the second-stage regression, and (5) we use dollar values of luck and skill (instead of rates of return) in the second-stage regression. For each of these alternative baselines, we perform all the robustness checks described earlier. Across all the resulting specifications, we find less than 3% of the coefficients on *Luck* × *Bad Luck* are negative and statistically significant at the 0.10 level. In less than one-third of the cases, the coefficient is negative, but not statistically significant.

We next examine asymmetry across different governance subsamples. The idea is that, even though there is no asymmetry on average, there might be asymmetry in particular subsamples of governance. Specifically, if asymmetry is consistent with optimal contracting, we should be more likely to observe it in firms that have good governance. On the other hand, if it is consistent with rent extraction, we should be more likely to observe it in firms that have poor governance. We use various dimensions of governance such as the governance index of Gompers, Ishii, and Metrick (2003), institutional ownership, the fraction of independent directors, the fraction of co-opted directors, CEO tenure, and CEO ownership. We find no evidence of asymmetry in pay for luck in any of the subsamples of governance. Similarly, we find no evidence of asymmetry in pay for luck when we examine subsamples by industry or subsamples by time period.

Finally, we take a closer look at the most-widely followed methodology to test asymmetry in pay for luck. We are able to replicate the finding from prior literature of asymmetry in pay for luck. However, this result does not hold out of sample (2002–2014), and even small changes to the specification result in the disappearance of asymmetry.

Overall, we conclude that there is no asymmetry in pay for luck of the form documented in the literature, whereby CEOs are rewarded for good luck more than they are penalized for bad luck. Or, at least, if there is such asymmetry, it is not immediately apparent in the data.[6] As such, our results pose a challenge to the literature on pay-for-luck asymmetry. Our findings indicate that researchers

[6] The data are available on the authors' Web sites.

need to rethink the idea—-one that appears to be firmly embedded in the compensation literature—-that managers extract rents through pay-for-luck asymmetry. Our findings, however, do not imply that executive compensation is optimally structured; rather, they indicate that if there is rent extraction, it is not through the channel of pay-for-luck asymmetry.

1. Data and Summary Statistics

We start with individuals identified as CEOs (*CEOANN* variable) in the Standard & Poor's *ExecuComp* database. *ExecuComp* indicates the dates when the CEO assumed office and when the CEO left office, but, in some cases, fails to identify an executive as the CEO even though he or she appears to be the CEO based on these dates. We classify these individuals also as CEOs. We include all data from 1992 to 2014 and our initial sample consists of 36,809 observations. We exclude CEO-years with less than 2 years of tenure so that we do not consider pay changes that include partial CEO-years. For the same reason, we exclude turnover years. With these restrictions, our sample size decreases to 28,878 observations, but the number of observations in the regressions vary depending on the specification. We winsorize all our variables at the 1st and 99th levels to minimize the effect of outliers.

Table 1 reports summary statistics for our key variables. The firms in our sample are large, with mean market capitalization of $5,825 million. The mean stock return is 18%. *Pay*, consistent with the literature, is the total compensation measure from *Execucomp* (= TDC1). ΔPay is change in pay given by current year's pay less prior year's pay. CEOs, on average, are paid $4.3 million, but their annual pay change is only $0.3 million. The change in logarithm of pay [= $\Delta log(Pay)$] is the logarithm of the ratio of current year's pay to prior year's pay, which is the same as the logarithm of rate of change in pay. This equals 9%, on average. In our sample, CEOs have an average tenure of 9.4 years.

Table 1
Descriptive statistics

	N	Mean	SD	Median	p25	p75
Market capitalization ($M)	28,869	5,825	14,565	1,305	502	4,032
Returns	28,878	0.18	0.51	0.12	−0.11	0.37
Pay ($1,000)	28,422	4,362	5,273	2,457	1,158	5,354
ΔPay ($1,000)	26,586	300	3,645	137	−358	943
Δlog(Pay)	26,541	0.09	0.64	0.08	−0.16	0.36
CEO tenure (years)	27,895	9.4	7.1	7.0	4.2	12.0

The table provides descriptive statistics on firm characteristics, chief executive officer (CEO) compensation, and CEO characteristics for the sample period 1992–2014. *Returns* is the annual stock return. *Pay* is the total compensation of the CEO (*ExecuComp*: TDC1). The variable includes CEOs' salary, bonus, the value of stock and option grants, long-term incentive payouts, other annual compensation, and all other compensation. ΔPay is CEOs' current-year pay minus their prior-year pay. $\Delta log(Pay)$ is the change in logarithm of pay. This variable is the logarithm of the ratio of CEOs' current-year pay to their prior-year pay, which is the same as the logarithm of rate of change in pay. *CEO tenure* is the number of years that the executive has been the CEO of the firm. We winsorize all variables at the 1st and 99th percentile levels.

2. Main Result: Symmetry in Pay for Luck

In this section, we develop our baseline specification to detect asymmetry in pay for luck.

2.1 First-stage regression: Decomposition of firm performance into luck and skill

In the first step, we estimate luck and skill using the following specification:

$$\begin{aligned} \textit{Firm Performance}_{i,t} &= \alpha_i + \beta_i \ \textit{Industry Performance}_{j,t} \\ &+ \delta_i \ \textit{Market Performance}_t + \varepsilon_{i,t}, \end{aligned} \tag{1}$$

where j corresponds to the industry of firm i. Although this first-stage specification includes only two luck factors, it nevertheless allows for a lot of flexibility in terms of plausible empirical choices. For our baseline, we (1) use stock returns as the measure of firm performance and hence the luck factors are industry returns and overall market returns; (2) use equal-weighted return for industry and market; (3) include the firm's own returns in industry return; (4) use the sample of *Execucomp* firms as the firm's peer group to compute industry return; (5) use firms within the same 2-digit SIC as peer firms to calculate industry returns; use all firms in the CRSP universe to calculate market returns (CRSP equal-weighted index); (6) use monthly returns of firm, industry, and market in the regressions; (7) include all firms regardless of fiscal year end; (8) winsorize firm returns at the 1st and 99th percentile levels; the industry return is computed with the winsorized firm returns; and (9) estimate the regression for each executive ("*co_per_rol*" in *Execucomp*). The rationale for these choices has already been discussed in the Introduction.

The intercept has the natural interpretation as average skill for each executive. Thus, skill for a given month is the intercept plus the residual. Luck for a given month is the firm's return less skill. Here, and in the rest of the paper, (1) luck plus skill equals the firm's returns, and (2) we annualize luck and skill by taking the average monthly estimates over the fiscal period and multiplying by 12.

Panel A of Table 2 reports the cross-sectional mean, median, standard deviation, and the 25th and 75th percentile values of the coefficients across all executives. Average α equals –0.002, average β = 0.999, and average δ = 0.024.[7] By construction, the correlation between monthly luck and skill is zero for each executive.

[7] In our baseline specification here, the regression is firm returns on equal-weighted industry returns and equal-weighted market returns. We include market returns to be comprehensive and capture anything that affects firms' returns that is orthogonal to industry returns. The industry includes the firm itself, and the industries are defined by the sample of included firms only, so we should expect that the average coefficient of industry returns across executive-firm observations is equal to 1 and the average for the market returns is equal to 0 (we find coefficients of 0.999 and 0.024 because of winsorizing). We do, however, consider several alternative ways of estimating the first-stage regression where this is not true (such as using value-weighted returns). We discuss this later in the paper.

Table 2
Main result

Panel A. First-stage regression decomposing firm's stock returns into luck and skill

Regression estimates: Firm returns$_{i,t} = \alpha_i + \beta_i$ Industry returns$_{i,t} + \delta_i$ Market returns$_{i,t} + \varepsilon_{i,t}$						
	# CEO-firm years	Mean	SD	Median	p25	p75
α_i	6,376	−0.002	0.022	0.0005	−0.009	0.008
β_i	6,376	0.999	0.756	0.967	0.618	1.328
δ_i	6,376	0.024	0.868	−0.013	−0.341	0.370
R^2	6,376	0.351	0.181	0.335	0.211	0.480

Panel B. Summary statistics of luck and skill

	N	Mean	SD	Median	p25	p75
Luck	24,714	0.153	0.267	0.143	0.009	0.295
Skill	24,714	0.002	0.316	0.008	(0.165)	0.176

Panel C. Second-stage regression testing asymmetry in pay for luck

		Dependent variable = ΔPay
Luck	γ_1	1,506.32***
		(3.6)
Luck × cdf (Variance of luck)	γ_2	−401.52
		(−0.8)
Luck × Bad luck	γ_3	**341.63**
		(0.7)
Skill		2,333.42***
		(6.8)
Skill × cdf (Variance of skill)		−1,350.25***
		(−3.4)
Skill × Bad skill		219.45
		(0.7)
cdf (Variance of luck)		−61.89
		(−0.4)
cdf (Variance of skill)		132.70
		(1.1)
Tenure		24.52
		(0.5)
Size		58.87
		(1.0)
Pay for good luck for median-risk firm	$\gamma_1 + 0.5\,\gamma_2$	**1,305.56*** **
		(5.6)
Observations		24,714
R^2		0.026
Intercept, executive FEs, year FEs		Yes

Panel A reports the first-stage regression results, where monthly firm returns are regressed on equal-weighted industry (firms in same 2-digit SIC in *Execucomp* sample) returns and equal-weighted (CRSP) market returns. The regression is estimated for each of the 6,376 CEO-firm observations. We use the word "executive" to refer to the unique CEO-firm combination identified by *co_per_rol* in *Execucomp* database. Panel A reports the cross-sectional mean of the coefficient estimates. t-statistics are in the spirit of Fama-MacBeth. They are computed as the average coefficient estimate across all executives divided by its standard error, which equals the standard deviation of the coefficient estimates divided by the square root of the number of executives. The intercept from the first-stage regression has the natural interpretation as average monthly skill for each executive. Skill for a given month is the intercept plus the residual. Luck for a given month is the firms' returns minus skill. *Luck* and *Skill* are the annualized estimates of monthly luck and skill (= monthly average over the fiscal period, multiplied by 12). Panel B reports the summary statistics of *Luck* and *Skill*. Panel C reports the second-stage regression results, where we test for asymmetry in pay for luck. The dependent variable is Δ*Pay*, which is the current year pay minus the prior year pay. *Pay* is the total annual compensation of the CEO (*Execucomp*: TDC1). We exclude CEO-years with less than 2 years of tenure and the final year of turnover so that we do not consider pay changes that include partial CEO-years. *Bad luck* is an indicator variable that equals 1 when $Luck \leq 0$ and equals 0 otherwise. *Bad skill* is an indicator variable that equals 1 when $Skill \leq 0$ and equals 0 otherwise. We estimate the variance of luck and skill using rolling 60-month windows that ends with the fiscal end. *cdf (Variance of luck)* and *cdf (Variance of skill)* denote the cumulative distribution functions of luck and skill. *Size* is the natural log of lagged market capitalization. t-statistics in panel C are based on standard errors adjusted for heteroscedasticity and firm-level clustering. * $p < 0.10$; ** $p < .05$; *** $p < .01$.

Panel B reports the summary statistics for annualized luck and skill. The mean skill is virtually zero (= 0.002), but there is significant variation across firm-years (SD = 0.316). Mean luck is 0.153; again, there is significant variation (SD = 0.267).

2.2 Second-stage regression: Testing for asymmetry in pay for luck

In the second step, we use luck and skill estimated in the first step:

$$\begin{aligned}\Delta Pay = {} & \gamma_0 + \gamma_1 Luck + \gamma_2 Luck \times cdf\ (Variance\ of\ Luck) \\ & + \gamma_3 Luck \times Bad\ \ Luck + \gamma_4\ Skill\ + \gamma_5\ Skill\ \times\ cdf\,(Variance\ of\ Skill) \\ & + \gamma_6\ Skill\ \times\ Bad\ Skill\ + \gamma_7\ cdf\,(Variance\ of\ Luck) \\ & + \gamma_8\ cdf\ \ (Variance\ of\ Skill)\ + \gamma_9\ Tenure \\ & + \gamma_{10}\ Size\ + Executive\ and\ Year\ FE + \varepsilon. \end{aligned} \tag{2}$$

For our baseline, we use (1) Δ*Pay* as the dependent variable; (2) stock returns as the measure of firm performance; hence, *Luck* and *Skill* come from the first-stage decomposition of stock returns; (3) use *Luck* and *Skill* as rate of return (rather than dollar values); and (4) include *Size* (firm's market capitalization as of the end of the prior year) as a control. Bertrand and Mullainathan (2001) discuss using firm size as a control variable (see pp. 905 and 928 of their paper).

Bad luck and *Bad skill* are indicator variables that equal one when *Luck* and *Skill* are negative. A negative coefficient on *Luck* $\times$ *Bad luck* implies asymmetry in pay for luck. Aggarwal and Samwick (1999) show that pay-performance-sensitivity is negatively related to variance of stock returns. In other words, pay change is negatively related to variance of stock returns interacted with stock returns. They estimate volatility using rolling 5 years of data, and use the transformation of volatility using the cumulative distribution function (cdf). We follow their methodology and extrapolate their logic to our setting. Instead of cdf of variance of stock returns, we include *cdf (Variance of Luck)* and *cdf (Variance of Skill)*; similarly, instead of stock returns interacted with variance of stock returns, we include *Luck* $\times$ *cdf (Variance of Luck)* and *Skill* $\times$ *cdf (Variance of Skill)*.

Panel C of Table 2 presents the results of our baseline specification above. We find that pay change is positively related to *Luck* (γ_1 = 1,506.32). Consistent with the logic in Aggarwal and Samwick (1999), we find the coefficient on *Luck* interacted with *cdf (Variance of luck)* is negative (γ_2 = –401.52), though not statistically significant. For the median-risk firm, the *cdf (Variance of luck)* = 0.5. Thus, "pay for good luck" is given by $\gamma_1 + 0.5\gamma_2$ = 1,506.32 + 0.5 $\times$ (–401.52) = 1,305.56. Thus, when a firm's return increases by 0.01 because of luck, the average pay increases by \$13,056 (= 0.01 $\times$ 1,305.56 $\times$ 1,000; because pay is in \$000s).[8] The bottom of Table 2 reports this economic significance. This is statistically significant at the 0.01 level (t-statistic = 5.6).

The "pay for bad luck" is given by $\gamma_1 + 0.5\gamma_2 + \gamma_3$ and equals 1,305.56 + 341.63 = 1,647.19. The coefficient on *Luck* × *Bad luck* (= γ_3) captures the incremental pay for luck when luck is bad. This is not statistically significant (t-statistic = 0.7), indicating that there is no asymmetry in pay for luck. Moreover, the coefficient is not negative, ruling out the possibility that there is asymmetry, but we are unable to detect it due to lack of power.

In all our tables going forward, we report the pay for good luck (= $\gamma_1 + 0.5\gamma_2$) along with incremental pay for bad luck (= γ_3), making it easier to gauge the economic significance of pay for bad luck when it does exist.

3. Robustness

We perform broadly four types of robustness: (1) robustness to various choices we make in estimating the baseline first-stage regression; (2) robustness to various choices we make in estimating the baseline second-stage regression; (3) comprehensive robustness to various alternative baselines; and (4) robustness across different subsamples.

3.1 Robustness to choices for the first-stage regression

While our baseline regression provides no evidence of asymmetry in pay for luck, it is possible that the results are sensitive to the specification. We explore several alternative specifications to the first-stage decomposition of firm performance (as described in the Introduction). In each case, we hold the baseline constant and change only one dimension at a time. Table 3 reports the results of the various tests in a concise manner. We report the pay for good luck for the median-risk firm (= $\gamma_1 + 0.5\gamma_2$) and the coefficient (= γ_3) that indicates the asymmetry in pay for luck. Row 0 reproduces the baseline results from panel C of Table 2 for ease of comparison.

3.1.1 Model to decompose performance into luck and skill. First, we explore several ways of decomposing returns into luck and skill. For our baseline, we use time-series regressions for each executive. This approach allows skill to vary with each executive, with the intercept having the natural interpretation as average skill for each executive. Moreover, this approach allows for the correlation between the firm's returns and industry returns (industry sensitivity) and the correlation between the firm's returns and market returns (market sensitivity) to vary across firms. Both assumptions seem reasonable.

For robustness, we use the same specification as the baseline, but instead of estimating the regressions separately for each executive, we estimate

[8] A 0.01 increase in the median market capitalization of \$1,305 million corresponds to \$13.05 million. Thus, the coefficient above implies that when a firm's return due to luck increases by \$13.05 million, CEO pay increases by \$13,056. That is, for a \$1,000 increase in market capitalization due to luck, the pay increase is \$1. The corresponding number in Garvey and Milbourn (2006) is \$0.79.

Table 3
Robustness to choices for the first-stage regression

Row	Robustness to the first-stage regression	Pay for good luck	Incremental pay for bad luck	N
(0)	Baseline specification (panel C, Table 2)	1,305.56***	**341.63**	24,714
		(5.6)	**(0.7)**	
(1)	Pooled (instead of by executive) regression	1,101.63***	**977.51***	24,714
		(4.5)	**(1.8)**	
(2)	Pooled regressions with executive FEs (instead of by executive)	1,100.12***	**978.87***	24,714
		(4.5)	**(1.8)**	
(3)	Luck sensitivity estimated each year using prior year data	1,426.12***	**84.47**	23,106
		(9.0)	**(1.3)**	
(4)	Skill = Residual (not Residual + Intercept)	1,263.29***	**746.94***	24,714
		(5.5)	**(1.7)**	
(5)	Skill = Intercept (not Residual + Intercept)	1,510.40***	**205.92**	24,714
		(10.3)	**(0.7)**	
(6)	No regression; luck = industry returns (similar to RPE)	1,101.63***	**977.51***	24,714
		(4.5)	**(1.8)**	
(7)	Luck factor: Only industry returns (no market returns)	1,195.09***	**1050.24****	24,714
		(4.6)	**(2.0)**	
(8)	Luck factors: Fama-French + Momentum factors	792.25***	**361.15**	24,714
		(2.7)	**(0.6)**	
(9)	Industry = *Compustat* (instead of *Execucomp*) firms in same 2-digit SIC	1,007.53***	**787.78****	25,154
		(3.7)	**(2.0)**	
(10)	Industry = Firms with same TNIC (instead of same 2-digit SIC)	1,084.03***	**576.89**	16,011
		(3.9)	**(1.2)**	
(11)	Excluding (instead of including) firm's return in industry return	1,198.77***	**383.51**	24,714
		(4.8)	**(0.7)**	
(12)	Value-weighted (instead of equal-weighted) industry and market returns	843.86***	**931.96**	24,714
		(2.9)	**(1.5)**	
(13)	Only December fiscal year end firms (instead of all firms)	777.24***	**692.21**	17,127
		(2.7)	**(1.2)**	
(14)	Annual returns (instead of monthly returns)	958.31***	**1,039.88*****	23,483
		(5.3)	**(2.7)**	
(15)	Unwinsorized (instead of winsorized) firms returns	1,288.60***	**153.69**	24,714
		(5.7)	**(0.3)**	

This table explores robustness to asymmetry in pay for luck to various choices to the first-stage regressions.

$$\textit{First stage: Firm Performance}_{i,t} = \alpha_i + \beta_i \textit{Industry Performance}_{j,t} + \delta_i \textit{ Market Performance}_t + \varepsilon_{i,t} \textit{[j=firm i's industry]}$$

$$\textit{Second stage:} \Delta Pay = \gamma_0 + \gamma_1 Luck + \gamma_2 \textit{ Luck} \times \textit{cdf (Variance of Luck)} + \gamma_3 \textit{ Luck} \times \textit{Bad Luck} + \textit{Controls}$$

We report the pay for good luck [= $\gamma_1 + 0.5\gamma_2$] and the incremental pay for bad luck [= γ_3] from the second-stage estimates for the median-risk firm. $\gamma_3 < 0$ indicates asymmetry in pay for luck. Row 0 reports the baseline (panel C of Table 2). In each case, we maintain the baseline, but change only one dimension at a time. In all cases, luck + skill = returns. *Row 1*: pooled (instead of executive-specific) regression. *Skill* = residual. *Row 2*: pooled regression with executive fixed effects included (instead of regression by executive). *Skill* = executive fixed effect + residual. *Row 3*: regressions by executive-year (instead of by executive). *Luck* = sensitivity estimated using prior year's beta .× current year's realization of industry and market returns. *Row 4*: same as baseline except *Skill* = residual (instead of intercept + residual). *Row 5*: same as baseline except *Skill* = intercept (instead of intercept + residual). *Row 6*: no regression; *Luck* = industry return. *Row 7*: only industry returns is included (and not industry and market returns). *Skill* = intercept + residual. *Row 8*: Luck factors are the four factors of the Carhart model. *Row 9*: industry returns are computed using all firms in *Compustat* (instead of *Execucomp*). *Row 10*: industry returns are computed using all firms with the same TNIC code (Hoberg and Phillips 2010), instead of SIC code. *Row 11*: we exclude the firm's own return in computing industry returns. *Row 12*: we use industry and market returns that are value-weighted (instead of equal-weighted). *Row 13*: we limit the sample to firm-years with a December fiscal year end. *Row 14*: regression is based on annual (rather than monthly) returns. *Row 15*: we use firms' unwinsorized returns (as opposed to winsorized values). For Rows 8–15, *Skill* = intercept + residual. t-statistics are based on standard errors adjusted for heteroscedasticity and firm-level clustering. $^{*}p < 0.1$; $^{**}p < .05$; $^{***}p < .01$.

a pooled regression, like in the prior literature. Pooling the observations implicitly assumes that the industry and market sensitivities are the same for all firms. Moreover, there is only one estimated intercept for all executive-month observations. Thus, as done in prior literature, we use the predicted value from the first-stage regression as our measure of luck. The residual is skill, which implicitly assumes that average skill across all executive-month observations is zero. Row 1 reports the results. The coefficient on *Luck* × *Bad luck* is now positive and significant, suggesting that, if anything, CEOs get penalized more for bad luck. As our later results will show, the positive asymmetry is not a robust result, but the lack of negative asymmetry is generally very robust.

We believe the assumptions behind the pooled-regression approach are restrictive. Thus, we modify the pooled regression by including executive fixed effects. This effectively allows each executive to have different skill and the fixed effect is the average skill across time for a given executive. We know that in a regression with fixed effects, the residuals sum to zero across time for each cross-sectional unit. Therefore, we estimate skill for each executive-month as the executive fixed effect plus the residual for the month; luck is the difference between the firm's returns and skill. Row 2 reports the results. As before, we find no asymmetry.[9]

Our baseline does not allow for the sensitivity to vary over time, even though they vary in the cross-section. To allow for time-varying sensitivity to industry and market returns, we use our baseline specification, but estimate the first-stage regressions for each executive-year. We require at least nine (monthly) observations per executive-year for the regression. We then predict luck for each month as the product of the sensitivity estimated using prior year's data and the current month's realization of industry and market returns. Skill is the difference between the firm's returns and luck. Row 3 presents the results. Again, we find no asymmetry in pay for luck.

Next, we maintain the baseline but change our definition of skill and luck. In our baseline specification, the intercept is the CEO's average skill over her tenure in our sample period and we compute the skill each month as the intercept plus the residual. We perform two robustness checks. First, we define skill as the residual (and not intercept + residual) from the baseline regression. This effectively assumes that the intercept does not reflect the skill of the CEO but, rather, the net impact of some omitted luck factor on returns. Row 4 reports the results; we find no asymmetry. Second, we define skill as the intercept only. This assumes that deviations from the CEO's long run mean skill is just noise. Row 5 presents the results;[10] as before, there is no asymmetry.

[9] As mentioned earlier, when we say we find no asymmetry, we mean the coefficient on *Luck* × *Bad Luck* is not significantly negative. In a few of our robustness checks, including this one, we find the opposite.

[10] We cannot use executive fixed effects in the second-stage regression, like in our baseline, because skill is the same for each executive. Therefore, we estimate this with firm fixed effects.

Next, we fall back on the early literature on RPE (Gibbons and Murphy 1990). We do not estimate a first-stage regression: luck is the equal-weighted industry return and skill is the firm's return net of luck. This is equivalent to assuming, in the pooled regression, an intercept of zero, industry sensitivity of one, and market sensitivity of zero. Row 6 presents the results.[11] Again, we find no asymmetry.[12]

3.1.2 Luck factors. In this subsection, we stick to the baseline, but consider alternative luck factors. First, we drop the market returns and consider industry returns as the only luck factor and reestimate the first-stage regression. Row 7 reports the results. We find no asymmetry.

Next, instead of using industry and market returns, we rely on the mutual fund literature, and use the four factors of Carhart (1997).[13] Row 8 shows that there is no asymmetry.

3.1.3 Other choices. Next, we focus on the construction of the industry return as one of the luck factors. In our baseline regression, we limit our sample to firms in the *Execucomp* universe to compute industry returns. We now consider, instead, all firms in the broader *Compustat* universe. Row 9 reports the results.[14] We do not find negative asymmetry; in fact, the opposite is true.

In our baseline, we define industry as firms within the same 2-digit SIC code. For robustness, we use TNIC (text-based network industry classification). Hoberg and Phillips (2010, 2016) argue that TNIC, which uses information from firms' 10K statements, is better at identifying competitors. Row 10 reports the results. Significantly fewer observations are available relative to the baseline, because TNIC data are only available through 2010. We find no asymmetry.[15]

Next, in our baseline, we include the firm's own returns in computing the industry returns because we assume skill is i.i.d. with mean zero. Excluding the

11 We replicate Gibbons and Murphy (1990) by regressing $\Delta log(Pay)$ on the firm's returns and its industry returns. We find that the coefficient on the firm's returns = 0.024 (compared with 0.016 in Gibbons and Murphy), and the coefficient on the industry returns is –0.075 (compared with –0.050 in Gibbons and Murphy). In Table 3 (Row 6), however, we slightly deviate from Gibbons and Murphy, to be consistent with our baseline. Thus, we use ΔPay (rather than $\Delta log(Pay)$) to estimate monthly (rather than annual) regressions, include (rather than exclude) the firm's returns in industry returns, and use equal-weighted (rather than value-weighted) industry returns.

12 The results are exactly the same as those obtained with the pooled regression (Row 1). This is because when we estimate pooled regressions with equal-weighted industry returns and market returns, the coefficient on the former is one, and the coefficients on the intercept and the market return are both equal to zero.

13 We thank Ken French for kindly providing the data on his Web site. Also, we use raw (rather than excess) firm and market returns to be consistent with our baseline.

14 We have more observations than does the baseline, because we drop industry-years with fewer than five firms. Doing this results in fewer observations getting dropped when we use the *Compustat* rather than the *Execucomp* sample.

15 We consider two additional variations. First, we include no industry at all (only equal-weighted or value-weighted market returns are included in the first stage). Second, we include compensation peer group returns, instead of industry returns. We calculate compensation peer groups, like in Bizjak, Lemmon, and Naveen (2008). We continue to find no asymmetry. Appendix Table A11 reports the results.

firm's own return will result in the average industry return being contaminated by skill.[16] We now verify that this choice does not affect our key inferences by repeating our estimation using industry returns that exclude the firm's own return. Row 11 shows that there is no negative asymmetry.

Next, we use value-weighted industry and market returns instead of equal-weighted returns. Row 12 reports the results. The coefficient of *Luck* × *Bad Luck* is statistically insignificant, indicating that there is no asymmetry in pay for luck.

Next, in our baseline, we estimate regressions separately by executive. Therefore, we can include firms with all fiscal year ends. In contrast, in the literature, luck and skill estimates are based on a single pooled regression of annual firms' returns on annual industry returns. This requires returns to be measured over the same horizon for all firms, which in turn limits the sample of firms to those with a December fiscal year end. To ensure that our choice to include firms with all fiscal year ends is not affecting the asymmetry result, we now reestimate our baseline, but keep only firm-years with a December fiscal year end. Row 13 reports the results. We restrict the sample to include only firm-years with a December fiscal end; doing so leads to significantly fewer observations relative to the baseline. We find no asymmetry.

Next, instead of estimating the first stage regression using monthly returns, we use annual returns to be consistent with the prior literature on pay for luck. Fewer observations are available here relative to our baseline, because we require at least 4 years of data per executive. Row 14 reports the results. The lack of negative asymmetry is a robust result; indeed, we find positive asymmetry.

Finally, in our baseline, we winsorize firms' returns at the 1st and 99th percentile levels before we estimate luck and skill. As a robustness check, we now reestimate our baseline using unwinsorized returns. Row 15 shows that there is no asymmetry.

Overall, the results in Table 3 reinforce our key result of no asymmetry in pay for luck.[17]

3.2 Robustness to choices for the second-stage regression

In this section, we consider robustness to the choices we make to the second-stage regression. Again, in each case, we hold the baseline constant and change only one dimension at a time. Table 4 reports the results. Row 0 reports the baseline specification from panel C of Table 2 for comparison purposes. In our baseline specification, we use Δ*Pay* as the dependent variable. This is because

[16] For example, consider an industry with three firms that have returns r_1, r_2, and r_3. Each firm's returns equal luck plus a skill term. Thus, r_1=L+s_1, r_2=L+s_2, and r_3=L+s_3. Further, if skill within each industry is assumed i.i.d. with mean zero, then $s_1+s_2+s_3=0$. If we exclude returns of firm 1 in its industry benchmark, the equal-weighted industry return will be $(r_2+r_3)/2$ = L + $(s_2+s_3)/2$ = L + $(-s_1)/2$. That is, the negative of the firm's own skill is included in its peer return.

[17] Online Appendix Table A1 provides the summary statistics for the luck and skill measures used in these 15 specifications.

Table 4
Robustness to choices for the second-stage regression

Row	Robustness to the second-stage regression	Pay for good luck	Incremental pay for bad luck	N
(0)	Baseline specification (panel C, Table 2)	1,305.56*** (5.6)	**341.63 (0.7)**	24,714
(1)	Δlog(Pay) instead of ΔPay	0.29*** (6.9)	**0.08 (0.9)**	24,676
(2)	log(Pay) instead of ΔPay	0.39*** (12.0)	**−0.03 (−0.4)**	25,132
(3)	Luck and skill in $ terms (instead of rates of return)	0.52*** (4.1)	**−0.01 (−0.3)**	24,714
(4)	ΔOI/assets (instead of stock returns)	7,083.91*** (3.1)	**5,580.02**** **(2.4)**	23,587
(5a)	ΔOI/assets and stock returns (instead of just stock returns): Coefficients relating to stock returns	1,286.56*** (5.4)	**323.80 (0.7)**	23,141
(5b)	ΔOI/assets and stock returns (instead of just stock returns): Coefficients relating to ΔOI/assets	3,464.31 (1.5)	**5,162.38**** **(2.2)**	23,141
(6)	Performance period: 12 months prior to date of largest grant	1,717.58*** (7.3)	**212.90 (0.4)**	23,064
(7)	Performance period: current fiscal year and the prior 2 fiscal years	3,058.21*** (7.6)	**2,284.56**** **(1.9)**	23,035
(8)	Median (instead of OLS) regression	533.19*** (10.4)	**59.43 (0.5)**	24,714
(9)	Exclude size as control variable	1,273.51*** (5.5)	**355.28 (0.7)**	24,714
(10)	Exclude skill terms	1,034.96*** (4.4)	**395.80 (0.8)**	24,714
(11)	Exclude cdf (Variance) terms	1,155.87*** (5.6)	**440.09 (0.9)**	25,178
(12)	Bad luck = bottom 20th percentile; good luck = top 20th percentile	1,568.43*** (6.7)	**318.62 (0.7)**	24,714
(13)	Bad luck = bottom 10th percentile; good luck = top 10th percentile	1,671.25*** (6.1)	**367.09 (0.7)**	24,714

This table explores robustness to asymmetry in pay for luck to various choices to the second-stage regressions.
First stage: $Firm\ Performance_{i,t} = \alpha_i + \beta_i\ Industry\ Performance_{j,t} + \delta_i\ Market\ Performance_t + \varepsilon_{i,t}$
[j = firm i's industry]
Second stage: $\Delta Pay = \gamma_0 + \gamma_1\ Luck + \gamma_2\ Luck \times cdf\,(Variance\ of\ Luck) + \gamma_3\ Luck \times Bad\ Luck + Controls$
We report the pay for good luck [= $\gamma_1 + 0.5\gamma_2$] and the incremental pay for bad luck [= γ_3] from the second-stage estimates for the median-risk firm. $\gamma_3 < 0$ indicates asymmetry in pay for luck. Row 0 reports the baseline (panel C of Table 2). In each case, we maintain the baseline, but change only one dimension at a time. *Row 1*: *Δlog(Pay)* is the dependent variable (instead of *ΔPay*). *Row 2*: *log(Pay)* is the dependent variable (instead of *ΔPay*). *Row 3*: luck and skill as dollar returns (instead of as rates of returns). Dollar returns are obtained by multiplying luck and skill estimated as rate of return by lagged firm's market capitalization. *Row 4*: luck and skill are based on Δ(OI/Assets) (instead of returns, where *OI*=EBIT). *Rows 5a, 5b*: luck and skill are based on both ΔOI/Assets and Returns. *Row 5a* reports coefficients relating to stock returns while *Row 5b* reports coefficients relating to accounting returns. *Row 6*: performance evaluation period equals the 12-month period prior to the largest equity grant date (rather than current fiscal year). *Row 7*: performance evaluation period equals the current fiscal year and the prior 2 fiscal years (rather than current fiscal year). *Row 8*: median (rather than OLS) regressions. *Row 9*: exclude the size control. *Row 10*: exclude the four skill terms [*Skill*, *Skill×Bad skill*, *Skill×cdf (Variance of skill)*, and *cdf (Variance of skill)*]. *Row 11*: exclude the four cdf variables [*cdf (Variance of luck)*, *Luck×cdf (Variance of luck)*, *cdf (Variance of skill)*, and *Skill×cdf (Variance of Skill)*]. *Row 12*: we define *Bad luck* = 1 if *Luck* < the 20th percentile (instead of < 0) and 0 otherwise. We also include *Luck×Good luck*, where *Good luck* = 1 if *Luck* > the 80th percentile. In keeping with the format, we report the pay for good luck for a firm with median risk [= $\gamma_1 + 0.5\gamma_2 + \delta$, where δ = coeff. on *Luck×Good luck*] and the incremental pay for bad luck [= $\gamma_3 - \delta$]. *Row 13*: define *Bad luck* = 1 if *Luck* < the 10th percentile (instead of < 0) and 0 otherwise. We also include *Luck×Good luck*, where *Good luck* = 1 if *Luck* > the 90^{th} percentile. t-statistics in panel C are based on standard errors adjusted for heteroscedasticity and firm-level clustering. $^{*}p < 0.1$; $^{**}p < .05$; $^{***}p < .01$.

most papers in the literature document asymmetry in pay for luck using change in pay as the dependent variable. For robustness, we consider two alternative dependent variables. First, we use $\Delta log(Pay)$. This is the same as the logarithm of the ratio of current year's pay to prior year's pay, which is the same as the logarithm of rate of change in pay. This measure has been used in various papers including the early ones on RPE (e.g., Gibbons and Murphy 1990). Row 1 reports the results. Second, we use the logarithm of the level of pay. This also has been used in the prior literature (e.g., Gopalan, Milbourn, and Song 2010). Row 2 report the results. In both cases, we find no asymmetry.

Next, we test whether our results are robust to using dollar values of luck and skill. In our baseline, we use luck and skill estimated as rates of return. In this, we are consistent with much of the prior literature that discusses pay for performance (starting with Coughlan and Schmidt 1985) and relative performance evaluation (Gibbons and Murphy 1990). As a robustness check, we now multiply our baseline measures of luck and skill estimated as rates of return by the firm's lagged market capitalization to obtain the dollar values of luck and skill (Garvey and Milbourn 2006). Row 3 reports the results. Again, we find no asymmetry.

Next, we consider accounting performance (rather than stock returns) as our baseline measure of firm performance that explains variation in pay. Specifically, like in Bertrand and Mullainathan (2001), we consider the change in the ratio of operating income to assets (= ΔOI/Assets), where operating income = EBIT.[18] This necessitates reestimating the first-stage regression to decompose the firm's accounting performance into luck and skill. We compute the equal-weighted average of ΔOI/Assets for firms in the same industry in *Execucomp* (Industry ΔOI/Assets) and the equal-weighted average of ΔOI/Assets across all *Compustat* firms (Market ΔOI/Assets). We require at least five firms in the industry to compute the industry ratio. Of course, unlike stock returns, we do not have monthly accounting performance. Thus, we use annual accounting performance measures. We estimate the luck and skill for each executive. Row 4 reports the results when we use luck and skill based on ΔOI/Assets. We find no asymmetry.

Next, we consider the possibility that firms compensate executives based on both stock returns and accounting returns because both provide signals of executives' performance (Core, Guay, and Verrecchia 2003). Thus, we include luck and skill of accounting returns (estimated as described above) to the baseline. Row 5a reports the results relating to the stock variables, and Row 5b reports the results relating to the accounting variables. We find no asymmetry with either the stock or accounting variables.

Next, we consider different performance evaluation periods that might determine CEO pay. In our baseline, we consider luck and skill estimated

[18] Our results are qualitatively similar if we use EBITDA or NI instead of EBIT. Our results are also similar if use levels of these ratios rather than changes. Online Appendix Table A10 provides the results.

over the fiscal year to match up with pay that is paid over the fiscal year; therefore, we calculate luck and skill based on contemporaneous annual returns. Determining the exact performance for which executives are rewarded is, however, challenging. First, firms typically pay their executives the salary throughout the year, the bonus at the end of the year, and stock and option awards at different points in the year. Moreover, while salary level may be determined based on prior year performance, bonus depends on the current year performance, and the level of stock and option awards may be partly a reward for current and past performance and partly an incentive for future performance. Indeed, in the case of fixed-value option plans, it is not obvious that the level of awards changes in response to current year's performance, at least within the current plan cycle. Further, annual pay includes a long-term incentive component that typically considers performance over three-year windows. Thus, it is important to consider whether the results change when the performance window is changed. We consider two alternatives: (1) performance over the 12 months prior to the largest equity grant date[19] and (2) performance over a cumulative 3-year period (current year plus 2 prior years).

To implement the first alternative, we calculate the grant date value of all the firm's option grants (prior to 2006) and all the firm's stock as well as option grants (post-2006).[20] We then consider the performance window as the 12 months prior to the largest equity grant made in a given fiscal year. The idea is that at the time the board awards the CEO the equity grant, it is reacting to the performance in the immediate preceding 12-month period. We annualize the monthly skill and luck obtained using the baseline over this 12-month period. Row 6 reports the results. We find no asymmetry. To implement the second alternative, we annualize the monthly skill and luck obtained using the baseline over the current fiscal year and the 2 fiscal years prior to the current fiscal year. Row 7 presents the results. We do not find negative asymmetry; in fact, the asymmetry is positive and significant.

Our baseline second-stage regression is estimated as an ordinary least squares (OLS) regression. We winsorize all our variables at the 1st and 99th percentile levels to eliminate outliers. We now estimate a median regression as an additional way to control for the influence of outliers.[21] Row 8 reports the results. The lack of asymmetry remains.

In our baseline, we include firm size as a control variable in the second stage as prior papers show that firm size is an important determinant of change in pay (Murphy 1999). It seems reasonable to include lagged size to control for

[19] We are grateful to the referee for this suggestion.

[20] The terms "prior to 2006" and "post-2006" are used rather loosely here. We base the definitions on the firm's reporting before and after the firm adopts FAS 123R, which became effective in 2006. Prior to FAS 123R, firms did not have to provide grant dates for either stock or option awards, but grant dates for the latter can be inferred using expiration dates.

[21] We do not control for firm FEs here.

the fact that the dollar change in pay received by the CEO could be higher for larger firms (after controlling for stock returns). Some papers on pay for luck, however, do not include size as a control in the second stage, presumably because they use dollar measures of luck and skill. Therefore, as a robustness check, we reestimate our baseline but exclude size as a control variable. Row 9 shows that there is no asymmetry.

Next, we consider robustness to excluding skill terms. Theoretically, testing for asymmetry in pay for luck does not require us to include skill in the second-stage regressions as long as luck and skill are not correlated. Our baseline annualized luck and skill measures have a correlation of –0.097. Row 10 reports the results when we drop the four terms that include skill [*Skill*, *Skill* × *Bad Skill*, *Skill* × *cdf (Variance of Skill)*, and *cdf (Variance of Skill)*]. The results do not change.

Next, we consider robustness to excluding the (cdf of the) variance of luck and skill and its interaction terms. Row 11 reports the results. We find no asymmetry.

Finally, in our baseline, we consider bad luck to be realizations of luck < 0. Our finding of no asymmetry in pay for luck could result if the true (unobserved) kink is at a different point. We explore different functional forms for nonlinearity. We consider kinks at more extreme points. Specifically, we redefine the *Bad luck* indicator variable as follows: = 1 if *Luck* < 20th percentile; = 0 otherwise. Because this is more extreme bad luck, we also form another indicator variable, *Good luck*. This equals 1 if *Luck* > 80th percentile; and equals 0 otherwise. That is, we estimate the following:

$$\Delta Pay = \gamma_0 + \gamma_1 Luck + \gamma_2\, Luck \times cdf\,(Variance\ of\ Luck) + \gamma_3\, Luck \times Bad\ Luck$$
$$+\ \delta Luck \times Good\ Luck + Controls.$$

In keeping with the format in Table 4, we report the pay for good luck for a firm with median risk ($= \gamma_1 + 0.5\, \gamma_2 + \delta$) and the incremental pay for bad luck ($= \gamma_3 - \delta$). Row 12 shows the results. We find no asymmetry. Alternatively, we define *Bad luck* = 1 if *Luck* < 10th percentile; = 0 otherwise. We also define *Good luck* = 1 if Luck > 90th percentile and = 0 otherwise. Row 13 reports the results. Again, we find no asymmetry.[22]

Overall, the takeaway from this table is that our finding of lack of asymmetry in pay for luck is robust to numerous alternative specifications of the second-stage regression.

3.3 Comprehensive robustness to alternative baselines

Thus far, in our robustness tests, we keep all else constant in our baseline and examine the sensitivity of our results to one change at a time. Although

[22] As a final check for nonlinearity, we include the square of luck (in lieu of *Luck* × *Bad Luck*); all other variables are as per the baseline. This allows for general nonlinearity without specifying the exact kink in the distribution. A positive coefficient on the square term would indicate negative asymmetry. In untabulated results, we find the coefficient of the squared luck term to be –8.11 (t-statistic of –0.0), which implies that pay for luck is linear.

we have a strong rationale for our baseline specification, some readers may wish to see the same comprehensive robustness to alternative baseline specifications.

Tables 3 and 4 verify the robustness to 28 alternative specifications. These 28 specifications, in turn, are based on 17 plausible decision variables facing researchers. Our first decision variable is the method to decompose performance into luck and skill. We explore seven possible choices (Baseline + Rows 1–6 of Table 3). For the following five decision variables, we explore three specifications each. These five decision variables are (1) the luck factors (Baseline + Rows 7 and 8 of Table 3); (2) the dependent variable in the second-stage regression (Baseline + Rows 1 and 2 of Table 4); (3) the performance measure (Baseline + Rows 4–5 of Table 4); (4) the performance evaluation period (Baseline + Rows 6 and 7 of Table 4); and (5) the specification for bad luck (Baseline + Rows 12 and 13 of Table 4). For the remaining 11 decision variables, we consider two specifications each (e.g., value-weighted versus equal-weighted returns, winsorized versus unwinsorized firms' returns). Thus, we have nearly 3.5 million specifications ($= 7^1 \times 3^5 \times 2^{11}$) to explore. Clearly, examining this large number is not feasible here. Thus, we limit our exploration to five plausible alternative baselines. We repeat all our robustness checks for these different baselines and essentially estimate $5 \times 28 = 140$ specifications. Specifically, we replicate Tables 3 and 4 for each baseline. The Online Appendix provides the full set of tables. Online Appendix Tables A2–A6 report the robustness to using each of the five alternative baselines.

First, we change the baseline to use Δ*log(Pay)*, instead of Δ*Pay*, as the dependent variable in the second-stage regression. Second, we change the baseline to use *log(Pay),* instead of Δ*Pay*, as the dependent variable in the second-stage regression. Third, we use the dollar values of luck and skill (rather than rates of return) in the second-stage regression. Fourth, we use pooled regression (instead of executive-specific regressions) in the first-stage regression to effect the decomposition of stock returns to luck and skill. Fifth, we use value-weighted (instead of equal-weighted) industry and market returns in the first-stage regression.

Overall, we consider $5 \times 28 = 140$ specifications, of which 129 are unique. We find less than 3% have significantly negative coefficients on *Luck* $\times$ *Bad Luck* at the 0.10 level; in less than one-third of the cases, the coefficient is negative, but not statistically significant.

3.4 Governance subsamples

In this subsection, we consider whether asymmetry in pay for luck differs across different subsamples of governance. The idea here is that if there is asymmetry in pay for luck *and* if such asymmetry is suboptimal, then we should observe asymmetry in subsamples of weak corporate governance where agency problems are more pronounced. On the other hand, if there is asymmetry in pay for luck *and* such asymmetry is optimal, we should observe it in subsamples of strong corporate governance. We examine this issue by estimating our second-stage regressions separately for subsamples of weak and strong governance. We classify the following as having weak governance: firms with high corporate governance index ("GIM index"), low institutional ownership, low fraction of independent directors, high fraction of co-opted directors, high CEO ownership, and high tenure of CEO.[23] We use the annual median values of these variables to classify firms into "high" or "low" groups. Table 5 reports the results separately for the weak governance subsample (panel A) and the strong governance subsample (panel B). We find no evidence of asymmetry in pay for luck in any of the subsamples.

3.5 Other subsamples

We examine whether asymmetry in pay for luck varies across different periods. Both external and internal governance measures have improved over time, so it is possible that we observe a decline in asymmetry in pay for luck if it is the result of weak governance. Further, compensation structure has significantly changed in the last decade (Bettis et al. 2010), and it is not clear how this affected the sensitivity of pay for luck or the asymmetry in pay for luck. We, therefore, estimate the baseline regression separately for two periods: 1992–2005 and 2006–2014. This roughly divides our sample into two halves and, importantly, coincides with the introduction of FAS 123R (compensation reporting requirements and the structure of compensation changed following FAS 123R, and the new rules were effective fiscal 2006 for most firms). We find that there is no asymmetry in either period (see Online Appendix Table A7).

Next, we consider pay for luck separately across different industries. It is possible that differences in the level of product market competition, regulation, external takeover threats etc. across industries will result in differences in external corporate governance across industries. This, in turn, may lead to differences in asymmetry of pay for luck. To examine this, we estimate the second-stage regression separately for seven industry groups.[24] We find that there is no asymmetry (see Online Appendix Table A8). In 2 of the 7 industries, the coefficient is negative, but not statistically significant even at 0.10 level. In the *Manufacturing* industry, the asymmetry is close to statistically significant (t-statistic = –1.5) and economically significant: for a firm with median risk, pay for good luck is 1,912.26 and incremental pay for bad luck is –953.48.

---

[23] For more discussion of these governance measures, see Gompers, Ishii, and Metrick (2003), Bertrand and Mullainathan (2001), Weisbach (1988), Coles, Daniel, and Naveen (2014), and McConnell and Servaes (1990).

[24] We map 2-digit SIC codes to industry groups based on the classification provided at https://www.osha.gov/pls/imis/sic_manual.html. Firms with 2-digit SIC codes between 15 and 17 are classified as *Construction*, between 20 and 39 as *Manufacturing*, between 40 and 49 as *Regulation*, between 50 and 59 as *Trade*, between 60 and 69 as *Finance*, and between 70 and 89 as *Service*. We exclude the *Agriculture* sector, which includes firms with 2-digit SIC codes below 10, because only 23 such firm-years are in our sample.

Table 5
Is there asymmetry in different governance subsamples?

Panel A. Subsamples of weak governance

Row	Governance subsamples	Pay for good luck	Incremental pay for bad luck	N
(0)	Baseline specification (panel C, Table 2)	1,305.56***	**341.63**	24,714
		(5.6)	**(0.7)**	
(1)	High GIM	1,548.43***	**551.07**	8,521
		(3.5)	**(0.6)**	
(2)	Low institutional ownership	1,209.00***	**105.13**	9,193
		(3.3)	**(0.1)**	
(3)	Low fraction independence	1,922.85***	**295.40**	8,979
		(4.7)	**(0.4)**	
(4)	High fraction co-opted	833.73*	**326.08**	9,640
		(1.8)	**(0.3)**	
(5)	High CEO ownership	1,238.93***	**391.59**	12,899
		(4.1)	**(0.6)**	
(6)	High CEO tenure	869.04***	**557.63**	14,253
		(2.8)	**(0.8)**	

Panel B. Subsamples of strong governance

Row	Governance subsamples	Pay for good luck	Incremental pay for bad luck	N
(0)	Baseline specification (panel C, Table 2)	1,305.56***	**341.63**	24,714
		(5.6)	**(0.7)**	
(1)	Low GIM	1,259.63***	**117.69**	11,314
		(4.0)	**(0.2)**	
(2)	High institutional ownership	1,433.71***	**793.00**	9,802
		(3.5)	**(0.9)**	
(3)	High fraction independence	1,049.45***	**272.74**	8,503
		(2.1)	**(0.3)**	
(4)	Low fraction co-opted	2142.70*	**123.96**	7,696
		(5.0)	**(0.2)**	
(5)	Low CEO ownership	1,344.38***	**148.41**	11,245
		(3.3)	**(0.2)**	
(6)	Low CEO tenure	1,529.82***	**81.12**	10,461
		(3.7)	**(0.1)**	

The table replicates the baseline results (panel C of Table 2) for various subsamples of weak and strong governance. Specifically, we estimate the second-stage regression, $\Delta\ Pay = \gamma_0 + \gamma_1\ Luck + \gamma_2\ Luck \times cdf(Variance\ of\ luck) + \gamma_3 Luck \times Bad\ luck + Controls$, for governance subsamples. We report the pay for good luck [= $\gamma_1 + 0.5\gamma_2$] and the incremental pay for bad luck [= γ_3] from the second-stage estimates for the median-risk firm. $\gamma_3 < 0$ indicates asymmetry in pay for luck. Row 0 reports the baseline. We classify the following as having weak governance: firms with high corporate governance index ("GIM index"), low institutional ownership, low fraction of independent directors, high fraction of co-opted directors, high CEO ownership, and high tenure of CEO. We use the annual median values of these variables to classify firms into "high" or "low" groups. The table reports the results separately for the weak governance subsample (panel A) and the strong governance subsample (panel B). t-statistics are based on standard errors adjusted for heteroscedasticity and firm-level clustering. $^{*}p < 0.1$; $^{**}p < .05$; $^{***}p < .01$.

We also examine whether asymmetry in pay for luck varies across different compensation subsamples. We consider three compensation measures: total pay (*TDC1*), value of option grants, and the ratio of the value of option grants to total pay. In each case, we classify firms as belonging to the "high" or "low" group based on the annual median level of the corresponding compensation measure. We find no asymmetry in pay for luck in any subsample (Online Appendix Table A9).

In sum, across the main tables and the Online Appendix tables, we find significant negative asymmetry (at the 0.10 level) in less than 2% of the

205 specifications we have explored; in less than one-fourth of the cases, the coefficient is negative, but not statistically significant. These findings suggest that our failure to detect negative asymmetry is not due to lack of statistical power.

4. Reconciling Our Results with the Prior Literature

Thus far, our results indicate the complete absence of asymmetry. These results are in contrast to the findings of Garvey and Milbourn (2006; henceforth GM), who do the first detailed analysis of asymmetry in pay for luck. GM find significant negative asymmetry. In this section, we try to reconcile the differences between our results and theirs. We start by successfully replicating the GM finding of asymmetry for our time period. We then document that asymmetry disappears (1) if we control more effectively for firm size in their specification and (2) when we explore several plausible variations to their specification.

4.1 Replicating results from the prior literature

First, we replicate GM to the best of our ability and find the same asymmetry. In the first-stage regression, GM estimate pooled regressions of firms' annual returns on equal-weighted industry (2-digit SIC) returns, value-weighted industry returns, and year fixed effects. Based on GM's paper, our understanding is that they (1) appear to include *Execucomp* firms as the peer firms; (2) exclude firm's returns in estimating firm's industry returns; and (3) include only firms with December fiscal year end to ensure returns for every firm in the sample are measured over the same horizon. GM then multiply the predicted and residual values from the first-stage regression by the firm's lagged market capitalization to obtain the dollar values of luck and skill. In the second-stage regression, they regress ΔPay on the same set of variables that we use in our baseline, but with three differences: (1) they use dollar values of luck and skill; (2) they do not appear to winsorize luck and skill; and (3) they do not include firm size.

Row 1 of Table 6 reports the results of our attempt to replicate GM for our sample period 1992–2014.[25] As with GM, we find that the coefficient on asymmetry in pay for luck is negative *and* statistically significant. The coefficients on *Luck*, *Luck* × *cdf(Variance of luck)*, and *Luck* × *Bad luck* are 0.77, –0.74, and –0.04, respectively. In terms of economic magnitude, the results imply that, for a firm of median risk (i.e., *cdf(Variance of luck)* = 0.5), when performance due to luck increases by \$1,000, the pay increases by 40 cents (= 0.77 + 0.5 × (–0.74) = 0.40). The coefficient of –0.04 on *Luck* × *Bad luck* indicates that for the same median-risk firm, when performance due to luck decreases by \$1,000, the pay decreases by 36 cents (= 0.40 – 0.04), which is

[25] We attempt to replicate their results as carefully as possible. We apologize to GM if our understanding is not correct.

Table 6
Replicating prior literature and robustness to variations in specification

Row		Pay for good luck	Incremental pay for bad luck	N	Correlation between luck and skill
Panel A. Replicating GM					
(1)	Replicating GM: Sample period (1992–2014)	0.40*** (2.8)	**−0.04**** **(−2.1)**	17,512	-0.50
Panel B. Different time periods					
(2)	Replicating GM: (GM sample period) First stage: 1992–2001; second stage: 1992–2001	0.37 (0.8)	**−0.54**** **(−2.1)**	6,352	-0.40
(3)	Replicating GM (Post-GM sample period) First stage: 2002–2014; second stage: 2002–2014	0.12 (0.7)	**−0.004** **(−0.1)**	10,439	-0.52
(4)	Replicating GM: (GM sample period) First stage: 1992–2014; second stage: 1992–2001	0.38 (1.1)	**−0.42**** **(−2.2)**	6,313	-0.36
(5)	Replicating GM: (Post-GM sample period) First stage: 1992–2014; second stage: 2002–2014	0.41*** (2.6)	**0.01** **(0.3)**	11,199	-0.56
Panel C. Excluding largest firm-years					
(6)	Excluding top 1%	0.41** (2.6)	**−0.11*** **(−1.8)**	17,337	-0.33
(7)	Excluding top 2%	0.34** (2.2)	**−0.08** **(−1.1)**	17,162	-0.19
(8)	Excluding top 3%	0.37** (2.2)	**−0.08** **(−0.8)**	16,987	-0.16
(9)	Excluding top 5%	0.44** (2.4)	**0.10** **(0.7)**	16,636	-0.11
(10)	Excluding top 10%	0.73*** (3.5)	**0.10** **(0.5)**	15,760	-0.08
Panel D. Excluding largest firms each year					
(11)	Excluding top 1%	0.44*** (2.8)	**−0.08** **(−1.6)**	17,334	-0.34
(12)	Excluding top 2%	0.42*** (2.7)	**−0.12** **(−1.6)**	17,157	-0.22
(13)	Excluding top 3%	0.39** (2.4)	**−0.07** **(−0.7)**	16,988	-0.10
(14)	Excluding top 5%	0.42** (2.4)	**0.05** **(0.4)**	16,630	-0.10
(15)	Excluding top 10%	0.61*** (2.9)	**0.18** **(0.9)**	15,759	-0.05

Table 6
(*Continued*)

Row		Pay for good luck	Incremental pay for bad luck	N	Correlation between luck and skill
Panel E. Minor perturbations to GM specification					
(16)	In the second stage, use winsorized (rather than unwinsorized) luck and skill	0.10 (1.3)	**0.00** **(0.0)**	17,512	-0.33
(17)	In the second stage, luck and skill as rate of return (instead of dollar values)	1,019.11*** (3.2)	**770.91** **(0.8)**	17,512	-0.02
(18)	In the first stage, estimate regressions by executive (instead of pooled regression)	0.50*** (4.0)	**0.02** **(0.5)**	16,502	-0.21
(19)	In the second stage, cluster standard errors by executive (rather than by firm)	0.40*** (2.9)	**−0.04** **(−1.4)**	17,512	-0.50
(20)	In the first stage, use equal-weighted industry returns (instead of both equal and value)	0.35** (2.1)	**−0.04**** **(−2.1)**	17,512	-0.55
(21)	In the first stage, use value-weighted industry returns (instead of both equal and value)	0.34** (2.3)	**−0.01** **(−0.5)**	17,512	-0.45
(22)	In 2^{nd} stage, exclude the four skill terms	0.23 (1.6)	**0.01** **(0.4)**	17,512	NA
(23)	In the second stage, use Δlog(Pay) instead of ΔPay as dependent variable [$\times 10^{-4}$]	0.64*** (3.3)	**−0.04** **(1.3)**	17,512	-0.50
(24)	In the second stage, use log(Pay) instead of ΔPay as dependent variable [$\times 10^{-4}$]	1.32*** (7.5)	**−0.04*** **(−1.7)**	17,512	-0.50

The table reports results of replicating prior literature (Garvey and Milbourn 2006; GM) on asymmetry in pay for luck to the best of our ability (Row 1) and robustness to GM baseline specification (other rows). In the first stage, we estimate pooled regression of firm's annual returns on equal-weighted industry (2-digit SIC) returns, value-weighted industry returns, and year fixed effects. We (1) include only *Execucomp* firms as peer group; (2) do not include firm's returns in estimating firm's industry returns; and (3) include only firms with December fiscal year end. The predicted value is luck and residual value is skill, both expressed as rate of return. We then multiply the predicted and residual values by the firm's lagged market capitalization to obtain the dollar values of luck and skill. In the second-stage regression, we regress ΔPay on the same set of variables that we use in our baseline, but with three differences: (1) use dollar values of luck and skill; (2) do not winsorize luck and skill; and (3) do not include firm size. Row 1 of panel A reports the results. In panel B, we estimate the regression from Row 1 for different periods. *Luck* and *Skill* are estimated using data for 1992–2001 and 2002–2014, respectively, for Rows 2 and 3, and are estimated using the entire sample for Rows 4 and 5. In panel C, each row reports results from dropping the largest X% of firm-years in terms of market capitalization from the sample. For example, Row 6 reports the results where the top 1% (by market cap) of firm-year observations are dropped. In panel D, each row shows the results from dropping the largest X% of firms each year in terms of market capitalization from the sample. Panel E shows the results for plausible variations to the baseline GM model in Row 1. *Row 16*: use winsorized (instead of unwinsorized) luck and skill terms; values are winsorized at the 1st and 99th percentile levels. *Row 17*: use luck and skill as rates of returns (instead of multiplying these by lagged market capitalization to obtain dollar values). *Row 18*: estimate the first-stage regression by executive (instead of pooled). *Row 19*: cluster standard errors by executive (rather than by firm). *Row 20*: use only equal-weighted industry returns (instead of both equal and value-weighted industry returns) and year fixed effects in the first stage. *Row 21*: use only value-weighted industry returns (instead of both equal and value-weighted industry returns) and year fixed effects in the first stage. *Row 22*: exclude the four terms containing skill. *Row 23*: $\Delta log(Pay)$ is the dependent variable (instead of ΔPay) in the second-stage regression. *Row 24*: *log(Pay)* is the dependent variable (instead of ΔPay) in the second-stage regression. t-statistics in panel C are based on standard errors adjusted for heteroscedasticity and firm-level clustering. NA indicates not applicable. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

a decline of 10% in sensitivity. GM find the corresponding numbers to be 79 cents and 60 cents (a decline of 25% in the sensitivity).

4.2 Effect of the sample period

While Row 1 reports the results based on our sample period of 1992–2014, GM's sample period is from 1992 to 2001. In Rows 2 and 3, we redo our estimation from Row 1 separately for the periods 1992–2001 and 2002–2014. Specifically, to estimate luck and skill (first stage), we use observations only from 1992 to 2001 for Row 2, and from 2002 to 2014 for Row 3. When we use GM's original sample period, we find that there continues to be an asymmetry in pay for luck (Row 2). When we reestimate these results for the post-2001 period, however, we find that the asymmetry is economically and statistically close to zero (Row 3). In Rows 4 and 5, we repeat the estimate of Rows 2 and 3, but now we estimate luck and skill using the entire sample (1992–2014). As before, we find asymmetry only in the first period.

In addition, we estimate asymmetry in luck over nonoverlapping 5-year periods (1993–1997, 1998–2002, 2003–2007, and 2008–2012). We start with 1993 because we wish to include 5 years of data in each subsample and we do not have change in pay for 1992. We find asymmetry in pay for luck only in the first two 5-year periods, which coincide approximately with GM's sample period. Online Appendix Table A12 offers the results.[26]

4.3 Effect of excluding large firms

It is puzzling as to why, if there is truly negative asymmetry in pay for luck in the data, it shows up in less than 3% of the specifications we have estimated so far. Therefore, we take a closer look at the GM methodology.

Luck, by definition, is random; hence, theoretically luck should have zero correlation with skill. Empirically, while luck and skill measured as rates of return are uncorrelated by definition (because they are the predicted and residual values from a pooled regression), we find that the dollar measures of luck and skill have a correlation of –50%. Thus, the seemingly innocuous step of multiplying luck and skill, expressed as rates of return, by lagged market capitalization to obtain the corresponding dollar measures, appears to be a problem. Conceptually, there is nothing wrong with the use of dollar values of luck and skill. For example, Aggarwal and Samwick (1999) use the dollar returns obtained by multiplying stock return by lagged market capitalization to explain variation in ΔPay. We also think that it is unfair to impose the theoretically high hurdle of zero correlation on empirical proxies for luck and skill. Nevertheless, the correlation is high enough to warrant further exploration.[27]

[26] In addition to nonoverlapping 5-year windows, we estimate the regressions over rolling 5-year windows corresponding to the post-GM period of 2002–2014. We do not find (results not tabulated) asymmetry in 6 of the 9 rolling 5-year windows.

[27] Our baseline estimates separate regressions for each executive rather than use the entire panel. By definition, luck and skill are uncorrelated at the executive level, as theory would suggest. Thus, we do not encounter the same problem.

The correlation is introduced when we multiply by lagged firm size (market capitalization), so we sort firms into quintiles based on lagged size and examine the correlation between *GM luck* and *GM skill* in each quintile.[28] We find that the negative correlation between the GM measures of luck and skill increases with size. Specifically, we find that the correlation increases (although not strictly monotonically) from 1.3% for firms in the smallest quintile of size to –50% for firms in the largest quintile. Even within the largest quintile, the correlation becomes increasing negative as firm size gets larger (the correlation is –0.06 up to the 98^{th} percentile of size, but is –0.53 for firms in the top-2 percentile). This raises the question of whether we would find any asymmetry when we exclude the largest firms.

Panel C of Table 6 reports the results when we exclude the largest firm-year observations by size. Rows 6–10 report the results from estimating the GM regression, but after excluding the largest 1%, 2%, 3%, 5%, and 10%, respectively, of firm-year observations. We find that the coefficient of *Luck* × *Bad luck* becomes progressively less negative and turns positive while the absolute value of t-statistic gradually decreases in magnitude. There is no asymmetry even when we exclude just 2% of observations in terms of market capitalization (coeff. = –0.08; t-statistic = –1.1). Thus, fewer than 2% of observations in terms of size appear to cause the asymmetry.

Excluding the top X% (the top 1%, 2%, etc.) of firm-years in the sample as we do above makes sense because we are estimating pooled OLS regressions and if we wish to ensure that outliers do not affect the results, we should consider the top X% of the entire sample. But the drawback to this is that we will likely be eliminating more observations in recent years as the largest X% of the sample by market capitalization are likely to be recent firm-year observations. To account for this, we now estimate the same regression, but exclude the top X% of firms by market capitalization in each year. Rows 11–15 of panel D show the results. The results are similar to those in panel C, but we now find that the asymmetry in pay for luck disappears even when we exclude just the top 1% of the sample each year. Thus, there is no asymmetry in 99% of the sample.

Overall, our findings here call into question the generalizability of the asymmetry result reported in prior literature; even when we use the exact same methodology used in the literature, we find the result is not robust.

[28] We do not wish to imply that multiplying by firm size solely contributes to the high correlation between luck and skill. The choice of first-stage decomposition methodology also affects the correlation. For example, we find that the average absolute correlation between the monthly measures of luck and skill in our baseline specification is zero when we use luck and skill as rates of return (which is what we would expect). When we use dollar values of luck and skill, the average of the absolute correlation (across executives) is 8.8%, showing the effect of multiplication by firm size on the correlation. When we use a different decomposition methodology, however, the results change. For example, when we use the baseline but estimate pooled regressions instead (like in Row 1 of Table 3), the correlation between luck and skill measured as rates of return is 0%, but the correlation between luck and skill measured as dollar values is -32%. This clearly indicates that both the decomposition methodology to estimate luck and skill as rate of return and the multiplication by firm size to obtain dollar values affect the correlation.

4.4 Effect of minor specification changes

In this section, we examine whether the GM result is robust to other, relatively minor, changes in specification.[29] As with the robustness to our baseline, we take the baseline GM specification and change one parameter at a time. Panel E of Table 6 reports the results.

First, we winsorize luck and skill. Row 16 reports the results. We find that when we use winsorized values, the asymmetry is no longer significant. Typically, larger dollar values of luck and skill are likely to be associated with larger firms. In fact, in our data, we find a significant overlap between the observations that are in the top 1% of GM luck and those that are in top 1% in terms of size. The overlap is over 50%. When we consider the top 2% by size, the overlap is even higher (=80%). The disappearance of asymmetry when we use winsorized values of luck and skill is consistent with our earlier evidence that asymmetry seems to be caused by large firms.

Second, we repeat the GM estimation, but use luck and skill estimated as rates of return. That is, we follow the exact GM methodology, but we use the luck and skill values obtained as rates of return (from the first stage) instead of multiplying them by firm size to get the dollar equivalents. Row 17 reports the results. We now find no asymmetry.

Third, instead of estimating pooled regressions in the first stage, we estimate executive by executive regressions. Row 18 reports the results. We find no asymmetry.

Fourth, we cluster the standard errors by executive rather than by firm. Row 19 reports the results. Once again, we find that the asymmetry term becomes statistically insignificant (although, by definition, economically it remains the same, like in the base-case GM specification).

Fifth, we follow the GM methodology with one change: GM use equal-weighted industry returns, value-weighted industry returns, and year fixed effects in the first-stage regression. We use only equal-weighted industry returns and year fixed effects or only value-weighted industry returns and year fixed effects in the first-stage. All other calculations are the same as those employed by GM. We do this because we find the use of both value-weighted and equal-weighted together as two separate independent variables a little puzzling; these two variables have a correlation of 82% and capture the same underlying economic variable. Further, prior work on RPE uses either equal- or value-weighted returns as the industry benchmark, not both. Rows 20 and 21 show the results. We find that, when we include only equal-weighted industry returns and year fixed effects, the GM result of negative asymmetry is evident in the data. When we use only value-weighted industry returns and year fixed effects, however, the asymmetry in pay for luck disappears.

Next, we drop the skill terms to determine whether there is asymmetry in pay for luck. In theory, because luck and skill are uncorrelated, we do not need the skill terms. We find no asymmetry when we exclude the skill terms (Row 22).

[29] We are grateful to a referee for suggesting this.

Finally, we use the exact GM methodology with one difference: in the second stage, we use Δ*log(Pay)* and *log(Pay)* respectively as the dependent variables instead of Δ*Pay*. Rows 23 and 24 present the results. We find no asymmetry when we use Δ*log(Pay)* but continue to find asymmetry when we use *log(Pay)*.

Overall, we find that the prior literature's finding of asymmetry, while intuitively appealing, is empirically not a robust result.

5. Conclusions

Current literature finds that CEOs' annual pay increases in response to performance attributable to good luck, but does not decrease as much in response to performance attributable to bad luck. Some studies find this asymmetry in pay for luck to be stronger in firms with poor governance and argue that their results are consistent with managerial rent extraction.

Our interest in revisiting this result of negative asymmetry stems from our observation that researchers have tremendous degrees of freedom in choosing the appropriate specification to test for asymmetry. In fact, over 3 million specifications are possible. We seek to examine whether the finding of asymmetry holds across numerous plausible specifications.

Among others, we consider (1) alternative methodologies to decompose performance into luck and skill (executive-specific regressions, pooled regressions), (2) alternative luck factors (industry returns, market returns, Carhart four factors), (3) alternative pay measures (pay change, level of pay, or rate of change in pay), (4) alternative performance measures (stock returns, accounting returns, combination of both), (5) alternative performance evaluation periods (current year's performance, 12-month performance preceding the largest equity grant, 3-year performance), and (6) alternative subsamples (based on governance, compensation, time periods, and industry groups). Overall, fewer than 2% of the 205 specifications we explore show significant negative asymmetry (at the 0.10 level). This result is in contrast to prior literature that finds strong evidence of negative asymmetry.

We do find negative asymmetry as documented in prior literature by replicating the exact specification used in prior literature. This result, however, does not hold out of sample, is not particularly robust to changes in specification, and seems to be caused by extremely large (top 1%–2%) firms.

Overall, our primary contribution in this paper is to document that there is no asymmetry in pay for luck. Testing for asymmetry in weakly governed firm is intuitively appealing. Researchers need to be careful, however, to use the right specification and to test for robustness given the wide latitude of choices available.

References

Aggarwal, R. K., and A. A. Samwick. 1999. The other side of the trade-off: The impact of risk on executive compensation. *Journal of Political Economy* 107:65–105.

Bertrand, M., and S. Mullainathan. 2001. Are CEOs rewarded for luck? The ones without principals are. *Quarterly Journal of Economics* 116:901–32.

Bettis, C., J. Bizjak, J. Coles, and S. Kalpathy. 2010. Stock and option grants with performance-based vesting provisions. *Review of Financial Studies* 23:3849–88.

Bizjak, J. M., M. L. Lemmon, and L. Naveen. 2008. Does the use of peer groups contribute to higher pay and less efficient compensation? *Journal of Financial Economics* 90:152–68.

Carhart, M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52:57–82.

Coles, J. L., N. D. Daniel, and L. Naveen. 2006. Managerial incentives and risk-taking. *Journal of Financial Economics* 79:431–68.

———. 2014. Co-opted boards. *Review of Financial Studies* 27:1751–96.

Core, J. E., W. R. Guay, and R. E. Verrecchia. 2003. Price versus non-price performance measures in optimal CEO compensation contracts. *Accounting Review* 78:957–81.

Coughlan, A., and R. Schmidt. 1985. Executive compensation, management turnover, and firm performance: An empirical investigation. *Journal of Accounting and Economics* 7:43–66.

Denis, D., Denis, J., and A. Sarin. 1997. Ownership structure and top executive turnover. *Journal of Financial Economics* 45:193–221.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3–56.

Frydman, C., and D. Jenter. 2010. CEO compensation. *Annual Review of Financial Economics* 2:75–102.

Gabaix, X., and A. Landier. 2008. Why has CEO pay increased so much? *Quarterly Journal of Economics* 123:49–100.

Garvey, G., and T. Milbourn. 2006. Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad. *Journal of Financial Economics* 82:197–225.

Gibbons, R., and K. J. Murphy. 1990. Relative performance evaluation for chief executive officers. *Industrial and Labor Relations Review* 43:30S–51S.

Gompers P., J. Ishii, and A. Metrick. 2003. Corporate governance and equity prices. *Quarterly Journal of Economics* 118:107–56.

Gopalan R., T. Milbourn, and F. Song. 2010. Strategic flexibility and the optimality of pay for sector performance. *Review of Financial Studies* 23:2060–98.

Hall, B. J., and J. B. Liebman. 1998. Are CEOs really paid like bureaucrats? *Quarterly Journal of Economics* 113:653–92.

Hoberg, G., and G. Phillips. 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies* 23:3773–811.

———. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124:1423–65.

McConnell, J., and H. Servaes. 1990. Additional evidence on equity ownership and corporate value. *Journal of Financial Economics* 27:595–612.

Morck, R., A. Shleifer, and R. Vishny. 1988. Management ownership and market valuation: An empirical analysis. *Journal of Financial Economics* 20:293–315.

Murphy, K. J. 1985. Corporate performance and managerial remuneration: An empirical investigation. *Journal of Accounting and Economics* 7:11–42.

———. 1999. Executive compensation. In *Handbook of Labor Economics*, eds. O. C. Ashenfelter and D. Card, 2485–563. Amsterdam, the Netherlands: Elsevier.

———. 2013. Executive compensation: Where we are, and how we got there. In *Handbook of the economics of finance*, eds. G. Constantinides, M. Harris, and R. Stulz, 211–356. Amsterdam, the Netherlands: Elsevier.

Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–80.

Weisbach, M. 1988. Outside directors and CEO turnover. *Journal of Financial Economics* 20:421–60.