PERFORMANCE PERSISTENCE THROUGH THE LENS OF CHANCE MODELS: WHEN STRONG EFFECTS OF REGRESSION TO THE MEAN LEAD TO NON-MONOTONIC PERFORMANCE ASSOCIATIONS

CHENGWEI LIU
Warwick Business School
University of Warwick
Coventry, CV4 7AL, UK

JERKER DENRELL
University of Warwick, UK

INTRODUCTION

Explaining when and why performance differences persist is an important topic in strategy and management (McGahan & Porter, 1997; Nag, Hambrick, & Chen, 2007; Rumelt, 1991; Teece, Pisano, & Shuen, 1997). Prior studies typically attribute performance differences among firms to three factors: the industry, corporate and firm effects and there is an on-going debate about how much each of these factors explains (see, Vanneste, 2017 for a meta-analysis). However, a common but often overlooked feature among these studies is that the total variances explained by these three factors together rarely exceeds 50%. Even after adjustments for sample size and variance decompsition techniques, on average there are still 45% unexplained variances for why performances differ (Vanneste, 2017). The implication is that unexplained variance, or simply chance, deserves a more central role in studying performance differences and their persistence in management.

Many management scholars have proposed that chance is an important alternative explanation for performance differences (Alchian, 1950; Aldrich, 1979; Arthur, 1989; Barney, 1986; Cohen, March, & Olsen, 1972; Denrell, 2004; Hannan & Freeman, 1989; Levinthal, 1991; Lippman & Rumelt, 1982; Nelson & Winter, 1982; Porter, 1991; Starbuck, 1994). But empirically this alternative explanation has not yet been taken seriously. We argue that management researchers could learn from other fields such as physics (Ruhla & Ruhia, 1992), genetics (Kimura, 1984) and ecology (Hubbell, 2001) where chance plays a central role in explaining observed differences (for a review, see Denrell et al., 2015). For example, management researchers can develop more rigorous theories when their empirical results are strong enough to reject the null hypotheses from "chance models" where observed performances are assumed to be the product of chance operating in structured environments (Schwab, Abrahamson, Starbuck, & Fidler, 2011; Starbuck, 1994).

Here we employ the idea of chance models to examine when and why performance differences persist in six datasets from professional sports and firm performances: National Football League (NFL, 2001-2016), National Basketball Association (NBA, 2004-2017); Major League Baseball (MLB, 1901-2016), Formula One Racing (F1, 1996-2015), return on assets of US firms (ROA, 1980-2010) and Fortune500 firms (1955-2005).

To address the "when" question, we draw on a simple chance model where performances are assumed to be a product of the more stable, dispositional factors (e.g., leaders' traits, firms' routines, corporate structure) and the more changeable, situational factors (e.g., uncontrollable circumstances, coincidence). Such a model implies "regression to the mean": more extreme performances tend to be followed by less extreme performances (Galton, 1886). For example, taller fathers tend to have shorter sons. The intuition is that the more extreme performances (e.g., heights) indicate stronger influence from dispositional factors (e.g., inherited genes) but also stronger influence from situational factors (e.g., random mutations). Since the situational factors are changeable and unlikely to persist, one should not expect the future performance to continue being so extreme. Hence, taller fathers tend to have, on average, shorter sons. This statistical account of performance (im)persistence also predicts that the effects of regressions would be stronger when the role of chance plays a more important role in performances.

Our results support our predictions: we found robust regression to the mean effects in all datasets examined and the effects are stronger when the performance is more unreliable, i.e., chance matters more in outcomes. Moreover, the regression effects can be so strong that they generate non-monotonic performance associations. For example, we show that top performing firms at year *t* (in terms of ROA, >95th percentile) tend to have a lower expected rank at year *t+1* than the "second best" at year *t* (i.e., 90th to 95th percentile). This rank reversal goes beyond conventional understanding of regression to the mean (cf. Harrison & March, 1984; Strang & Macy, 2001) and we will discuss its implication for learning and competition.

We also addressed the "why" question about performance persistence. We found strong effects of regression to the mean but they do not always produce non-monotonic performance associations. Even when they did, the locations of the non-monotonic kinks vary. We developed chance models to computationally examine how specific contextual factors, combined with chance, can produce results that are consistent with our empirical findings. These "chance explanations" provide stronger null hypotheses when testing alternative theoretical predictions than current practices (Schwab et al., 2011; Starbuck, 1994). For example, performance feedback theory suggests that top performing firms tend to decline because complacency-triggered risk aversion (Greve, 2003) or superstitious learning and inertia from prior successes (Audia, Locke, & Smith, 2000). Our chance models suggest that these interesting, causal effects may exist but exaggerated when researchers did not take into account the dull, statistical effects of regression to the mean, particularly for the "outliers" whose performances regress downward the most. The implication is that chance models are not contradictory but complementary to current practices when developing more rigorous management theories.

Our intended contribution is threefold. First, we provide a new way of examining performance persistence by borrowing ideas from natural sciences. Chance models emphasize the impacts of unsystematic variances that management researchers usually consider nuisance to be eliminated in their explanations. Our results echo a widely shared belief in natural sciences that chance in structured environment can produce systematic patterns (Bak, 1997; Newman, Barabasi, & Watts, 2006; Sober, 1993). Our findings also illustrate how chance models can in fact strengthen the quality of theoretical development in management. Brushing off chance explanations can lead to overestimation of systematic factors. Chance models deserve more attention from management researchers and we illustrate two of the many possible ways of developing and utilizing them.

Second, we show that the effects of regression to the mean are prevalent in various contexts but their effects are idiosyncratic. The effects can be so strong in certain performance ranges that they generate non-monotonic performance associations and rank reversals. In particular, top performing firms seem to be increasingly sensitive to regression to the mean - we show a decreasing persistence among top performing firms in the past thirty years based on ROA and Fortunte500 firms data. This suggests that the main mechanism that produces exceptional performances is increasingly likely to be factors beyond managers' control. Some firms could get lucky in one year but most of these "outliers" soon fail to beat the effects of regression to the mean (cf. Peters & Waterman, 1982). Managers who ignore these findings and instead believe that firms can be "built to last" or move from "good to great" by design are to commit illusion of control. We developed several chance models to replicate these non-monotonic performance associations we found and we hope future researchers could build on our ideas and develop more thoughtful chance models to further explore when and how regression to the mean occurs.

Third, the non-monotonic performance associations and the implied rank reversals have important implications for searching for strategic opportunities. People tend to believe that higher performances indicate higher skill, romanticize these exceptional performers and expect their performances to persist, ignoring the effects of regression to the mean (Dawes, 1979; Kahneman, 2011; Kahneman & Tversky, 1973; Meindl, Ehrlich, & Dukerich, 1985). Our findings are against laypersons' intuition and may imply collective false expectations and misevaluations, particularly around the ranges where performance non-monotonicity occurs. This may be bad news because systematic over- and under-estimations are likely. But this can mean good news to "strategic arbitragers" who are ready to exploit others' predictable mistakes (Liu, 2017; Zuckerman, 2012).

In particular, we show how the non-monotonic patterns we found can be translated into profits using data from sport betting and discuss when they imply strategic opportunities beyond sport. People tend to believe that higher performances indicate higher skill and higher future performance. The non-monotonic performance associations identified suggest a counterintuitive result that people may not comprehend or resist to accommodate. This implies a possible arbitrage opportunity: strategic arbitragers can buy low (e.g., acquiring extremely poor performing actors or resources because they are likely under-estimated by others) and sell high (e.g., releasing exceptionally performing actors or resources because they are likely over-estimated by others). But there is often a "limit to arbitrage" (Benner & Zenger, 2016; Litov, Moreton, & Zenger, 2012; Shleifer & Vishny, 1997) – one cannot profit from having unique insights if one's strategy or performance depend on stakeholders who are likely to mistake luck for skill when evaluating the outliers. In other words, social constraint can make the inefficiencies in evaluations to persist, e.g., lucky but incompetent executives continue to receive huge bonus even when their performances become systematically worse than their currently lower performing counterpart, i.e., the second best.

The limits to arbitrage due to interactive social dynamics may be weaker in some contexts. For example, the non-monotonic associations we identified in NFL and MLB may entail profitable arbitrage opportunities because bookmakers and sport gamblers may not yet incorporate fully the regression effects, and particularly the non-monotonic performance associations. This implies that the odds for the exceptionally performing teams may be overly favorable whereas the odds for the extremely poor performing teams may be overly unfavorable.

More importantly, limits to arbitrage are much weaker in this case, as long as one has sufficient funding to develop an investment portfolio that utilizes the non-monotonic performance association identified.

To examine this prediction, we downloaded the historical odd data from www.oddsportal.com. We then identified the highest and lowest performing teams in NFL, where performance non-monotonicity is most significant. We then simulated various investment portfolios based on bets on the salient winners' future season games to lose and on the salient losers' future season games to win. The idea is to exploit bookmakers' underestimation of salient losers and the overestimation of salient winners. The simulation results show that the investment portfolios have a positive return when the bet focuses only on the first game in the next season. This is not surprising in hindsight – bookmakers adjusted their odds based on the observed performances during a season so the odds become increasingly reliable and reflect the actual quality of the team (against its component) over time. However, information about team quality is more ambiguous before the first game of a season and hence it is influenced more by last season's outcomes, making the arbitrage opportunities most likely to exist. Interestingly, this arbitrage strategy works better for betting on salient losers' first season game to win. The reason is not because these underdogs are more likely to win, but because the odd is more favourable (e.g., 10/1 if this last season's loser wins this first game) than a similar strategy that focuses on the salient winners (e.g., 3/1 if this last season's winner loses this first game). Stated differently, strategic arbitrages work more effectively for the underdogs than the stars because the degree of underestimation is greater than the degree of overestimation. One possible explanation is about the asymmetrical attention people pay to the winners versus the losers (Pontikes & Barnett, 2017): winners attract more attention and scrutiny so their odds can be overestimated but not as much as the degree of the underestimation against the losers. Overall this exercise shows that our theory is not merely an abstract possibility but can be translated into a behavioral insight and an alternative source of superior profit (Fang & Liu, Forthcoming).

REFERENCES AVAILABLE FROM THE AUTHOR(S)