Product Market Peers and Relative Performance Evaluation

Sudarshan Jayaraman, Todd Milbourn, and Hojun Seo*

November 21, 2015

Abstract

Relative Performance Evaluation (RPE) theory predicts that firms filter out common shocks (i.e., those affecting the firm and its peers) while evaluating CEO performance, and that the extent of filtering increases with the number of firms in the peer group. Despite the intuitive appeal of the theory, previous tests of RPE find weak and inconsistent evidence. We hypothesize that one reason for the mixed evidence is the inaccurate classification of peers. Rather than using static, pre-defined Standard Industry Classifications (SIC), we exploit recent advances in textual analysis and define peers based on firms' product descriptions in their 10-K filings (e.g., Hoberg and Phillips, 2015). This alternative classification not only captures common shocks to firms' product markets more effectively, but also tracks the evolving nature of these markets, as 10-Ks are updated annually. Using product market peers, we find three pieces of evidence consistent with RPE – (i) firms on average filter out common shocks to performance measures, (ii) the extent of filtering increases with the number of peers, and (iii) firms completely filter out common shocks in the presence of a large number of peers. We are able to replicate the first finding but not the other two using pre-defined industry classifications. Overall, our results suggest that a key identification strategy to testing RPE theory lies in accurately defining the peer group.

JEL codes: M40; M41; G30; J33
Keywords: Product market peers, CEO compensation, Relative Performance Evaluation

---

* Sudarshan Jayaraman: University of Rochester, e-mail: jayaraman@simon.rochester.edu. Todd Milbourn: Washington University in St. Louis, e-mail: milbourn@wustl.edu. Hojun Seo: Washington University in St. Louis, e-mail: hojun.seo@wustl.edu. We appreciate helpful comments from Dirk Black, Patricia Dechow, John Donovan, Richard Frankel, Radha Gopalan, Jerry Hoberg, Jared Jennings, Ranjani Krishnan, Partha Mohanram, Ram Natarajan, and workshop participants at 2015 Conference on Convergence of Financial and Managerial Accounting Research and Washington University in St. Louis. We thank Gerard Hoberg and Gordon Phillips for providing product market peers data. All errors remain our responsibility.

Product Market Peers and Relative Performance Evaluation

Abstract

Relative Performance Evaluation (RPE) theory predicts that firms filter out common shocks (i.e., those affecting the firm and its peers) while evaluating CEO performance, and that the extent of filtering increases with the number of firms in the peer group. Despite the intuitive appeal of the theory, previous tests of RPE find weak and inconsistent evidence. We hypothesize that one reason for the mixed evidence is the inaccurate classification of peers. Rather than using static, pre-defined Standard Industry Classifications (SIC), we exploit recent advances in textual analysis and define peers based on firms' product descriptions in their 10-K filings (e.g., Hoberg and Phillips, 2015). This alternative classification not only captures common shocks to firms' product markets more effectively, but also tracks the evolving nature of these markets, as 10-Ks are updated annually. Using product market peers, we find three pieces of evidence consistent with RPE – (i) firms on average filter out common shocks to performance measures, (ii) the extent of filtering increases with the number of peers, and (iii) firms completely filter out common shocks in the presence of a large number of peers. We are able to replicate the first finding but not the other two using pre-defined industry classifications. Overall, our results suggest that a key identification strategy to testing RPE theory lies in accurately defining the peer group.

JEL codes: M40; M41; G30; J33
Keywords: Product market peers, CEO compensation, Relative Performance Evaluation

"…despite the obvious attractive features of relative performance evaluation, it is surprisingly absent from U.S. executive compensation practices. Why shareholders allow CEOs to ride bull markets to huge increases in their wealth is an open question…we view the weak evidence of relative performance evaluation as an important puzzle for executive compensation research."

Abowd and Kaplan (1999, pg. 157)
Journal of Economic Perspectives

One of the central tenets of agency theory is that increasing the "signal-to-noise" ratio of the performance measure reduces risk without compromising the level of incentive alignment (e.g., Holmstrom, 1979). In other words, eliminating sources of variation from firm performance such as an industry wide movement in stock returns that is beyond an individual manager's influence results in a more efficient contract (i.e., one that achieves greater alignment without increasing risk). This is the idea behind Relative Performance Evaluation (RPE) where CEO compensation should not only be positively correlated with own firm performance, but also negatively correlated with industry-wide or market-wide performance to filter out common (i.e., uncontrollable) performance. RPE theory also predicts that this filtering should increase with the number of firms in the peer group, and in the limit, common performance should be completely filtered out from firm performance (Holmstrom, 1982; Gibbons and Murphy, 1990).

Despite the intuitive appeal of the theory, prior research documents weak and inconsistent evidence and refers to this lack of evidence as the "RPE puzzle" (see Abowd and Kaplan above, Antle and Smith, 1986; Barro and Barro, 1990; Gibbons and Murphy, 1990; Janakiraman, Lambert and Larcker, 1992; Prendergast, 1999; Bushman and Smith, 2001; Lambert, 2001; Frydman and Jenter, 2013). While some studies attribute the RPE puzzle to managerial rent extraction (e.g., Bertrand and Mullainathan, 2001; Garvey and Milbourn, 2006), avoidance of self-destructive behavior (e.g., Gibbons and Murphy, 1990), and encouraging collusive behavior (e.g., Aggrawal and Samwick, 1999b) others attempt to find cross-sectional variation in the use of RPE (e.g., Garvey and Milbourn, 2003; Rajgopal, Shevlin, and Zamora, 2006; Gopalan, Milbourn, and Song, 2010).

In this study, we hypothesize that the prior lack of RPE evidence is (amongst other things) on account of how the peer group is defined. In other words, we posit that pre-defined industry classifications such as Standard Industry Classification (SIC), Global Industry Classification Standard (GICS) and North American Industry Classification System (NAICS) are noisy proxies for the peer group.[1, 2] Rather than using these static definitions, we exploit recent advances in textual analysis and define peers based on firms' product descriptions in their 10-K filings (e.g., Hoberg and Phillips, 2015).[3] We contend that this product-market based definition better captures common shocks that affect the firm and its peers – which is what the theory purports to capture. An additional benefit is that these product-market based definitions reflect the changing nature of firms' business models as 10-Ks are updated annually (Hoberg and Phillips, 2015).[4] Standard industry classifications, in contrast, fail to capture these dynamic aspects of product markets.

To capture product-market peers, we use Hoberg and Phillips' (2015) Text-based Network Industry Classifications (TNIC) that are constructed based on firms' product similarities.

---

[1] These pre-defined industry classifications focus on whether firms' *production processes* are similar rather than whether firms produce *similar products*. For example, "NAICS will be erected on a production-oriented, or supply-based, conceptual framework. This means that producing units that use identical or similar production processes will be grouped together in NAICS." http://www.naics.com/info.htm

[2] We are not the first to assert that SIC classifications are an imperfect proxy for defining peers. Prior work indicates that SIC codes mis-classify firms by ignoring within-group heterogeneity (e.g., Clarke, 1989; Kahle and Walkling, 1996; Bhojraj et al., 2003; Dopuch et al., 2008; Brickley and Zimmerman, 2010; Guenther and Rosman, 1994)

[3] There are other ways to define peers than using pre-defined industry classifications. For instance, one can use peers chosen by analysts (De Franco et al., 2012), those co-searched by investors (Lee et al., 2015), those self-selected by the firm (Gong et al. 2011; Lewellen 2013). While these classifications might also reflect product market peers, the underlying criteria for selecting these peers are ex ante unobservable.

[4] Hoberg and Phillip (2015) validate that their classification better explains differences in industry characteristics such as profitability, sales growth, and market risk across industry. They show that positive (negative) industry demand shocks lead to more (less) firms entering into those industries. They also show that these classifications better reflect competitors identified by managers. Several follow-up studies find that this classification scheme provides new insights regarding a firm's product market peers. For example, Hoberg and Phillip (2010) show that M&A transactions are more likely between firms having similar product descriptions and long-term outcome such as profitability is better when the target and the acquirer have similar product descriptions ex ante, possibly due to product market synergies. Foucault and Fresard (2014) show that a firm's investment is sensitive to the stock returns of product market peers.

Hoberg and Phillips (2015) calculate product similarity scores of all possible pairs of firms in each year by parsing firms' product descriptions in annual 10-K filings. If the similarity score between a firm and its potential peer firm is above the pre-determined similarity threshold, the latter is identified as a product market peer. Thus, each firm has its own distinct set of product market peers under this scheme. In addition, the composition of the TNIC-based peer group varies over time because TNIC are based on firms' product descriptions in 10-K filings, which are updated annually. This reflects the changing nature of the firm's product markets as business strategies change. To form peers, we form quartile portfolios based on size and book-to-market (e.g., Albuquerque, 2009) within each focal firm's TNIC industry and define RPE peers as firms closest to the focal firm within the same quartile portfolio. Equal-weighted average stock returns of these peers provide a measure of peer performance.

We begin by verifying whether TNIC-based peers indeed reflect product-market factors better than those based on pre-defined industry classifications. To do so, we estimate the correlation between firm sales (and operating costs, i.e., the sum of cost of goods sold and SG&A expenses) and peer average sales (and peer average operating costs) using three alternative industry classifications – TNIC, SIC and GICS. Consistent with our expectation and prior evidence in Hoberg and Phillips (2015), we find that the correlation between firm sales and peer sales is the strongest based on TNIC as compared to SIC or GICS.[5] The differences in correlations between TNIC and SIC/GICS are not only statistically significant but also economically meaningful. For example, the correlation between firm sales and TNIC-peer average sales is 0.786 while that for SIC (GICS) is 0.243 (0.161). The evidence for operating costs (i.e., the sum of cost of goods sold

[5] Hoberg and Phillips (2015) use a slightly different methodology by examining the extent to which alternative peer group classifications (TNIC/SIC/NAICS) generate higher levels of across-industry variation in profitability, sales group and stock market betas – where greater variation indicates a more informative industry classification.

and SG&A expenses) is more pronounced – the correlation between operating costs of the firm and TNIC-based peer average operating costs is 0.876 while that based on SIC (GICS) is 0.201 (0.159). Further, consistent with TNIC classifications better capturing the evolving nature of product markets, the above results are generally stronger in more recent time periods.[6] Overall, these tests indicate that TNIC-based classifications better capture product market factors than pre-defined classifications such as SIC or GICS.

Next, we turn to the RPE evidence. We refine the set of product market peers to those with comparable size and book-to-market values as the focal firm. We do so for two reasons – (i) recent studies note that matching the focal firm to SIC-based peers using size and book-to-market helps uncover evidence in favor of RPE (e.g., Albuquerque, 2009), and (ii) using closely matched size and book-to-market peers helps filter out managers' strategic use of self-selected larger firms to pay themselves more compensation (e.g., Bizjak, Lemmon, and Naveen, 2008; Faulkender and Yang, 2010).

Using a firm-year panel of 22,028 observations spanning from 1996 to 2013, we investigate how changes in total CEO compensation correlates with own firm performance and peer firm average performance (i.e., a measure of common performance). First, changes in total CEO compensation are not only positively correlated with own firm stock returns (i.e., positive pay-for-performance sensitivity) but also negatively correlated with peer firm average stock returns. While the latter result is consistent with RPE, the magnitude of the own firm effect (coefficient = 0.303) is significantly larger than the absolute magnitude of the peer firm effect (coefficient = -0.115)

[6] For example, the correlation between firm COGS and TNIC-peer average COGS is 0.821 in the 1996-1999 period, 0.891 in the 2000-2004 period and 0.908 in the 2005-2008 period. In contrast, COGS correlations based on SIC are decreasing over time – 0.252 in 1996-1999; 0.188 in 2000-2004 and 0.162 in 2005-2008. These effects become weaker in the most recent post-crisis period of 2009-2013 – in particular, the TNIC-based correlations drop relative to prior periods but are still larger than those based on SIC/GICS.

indicating only partial filtering out (i.e., weak-form RPE, see Gibbons and Murphy, 1990). In other words, a 1% increase in the firm's own stock price is associated with a 30.3 basis points increase in total CEO compensation when the peer group experiences a 0% return. On the other hand, if the peer group also experiences a 1% return, CEO compensation reduces by 40% (-0.115/0.303) and he/she still enjoys a net 18.8 basis points (i.e., 0.303-0.115) increase in total annual compensation.

Second, the extent to which common performance is filtered out increases with the number of firms in the peer group (e.g., Holmstrom, 1982). The idea is that common shocks can be better estimated with less noise when the peer group consists of many firms. We find that the negative coefficient on peer stock returns becomes larger in magnitude as the number of firms in the peer group increases. This coefficient takes a value of -0.073 when the firm has a few peers, -0.123 when the firm has a moderate number of peers and -0.315 when the firm has many peers.[7] Reassuringly, the coefficient on own firm performance does not differ significantly based on the number of firms in the peer group. This coefficient takes a value of 0.264 when the firm has a few peers, 0.334 when the firm has a moderate number of peers and 0.320 when the firm has many peers, and none of them is statistically different from each other. In terms of economic significance, when the focal firm and the peer group both experience performance of a similar magnitude, the sensitivity of pay-for-own-firm-performance falls by 27.6% (-0.073/0.264) in the "few peers" group, by 36.8% (-0.123/0.334) in the "moderate peers" group, and by 98.4% (-0.315/0.320) in the "many peers" group.

The latter result represents our third finding, i.e., common performance is completely filtered out when the firm has many peers. In particular, net CEO compensation is 0.5 basis points (and insignificantly different from zero) when both the firm and the peer group experience stock

[7] "Few", "moderate" and "many" are defined based on terciles of the number of firms in the TNIC peer group.

returns of a similar magnitude – but only in firms with many peers. This evidence is referred to as "strong-form" RPE where common performance does not figure in CEO compensation and he/she is compensated solely based on idiosyncratic performance.[8] This stringent test of RPE theory has yet not been documented (to the best of our knowledge) by prior work. It also speaks to Abowd and Kaplan's (1999) opening quote about CEO rent-extraction and indicates that firms *do not* allow CEOs to walk away with millions during bull markets. However, their ability to rein in pay-for-luck hinges on the presence of many peers so as to better estimate common performance.

Our main results are robust to various sensitivity checks. First, using SIC or GICS provides evidence of weak-form RPE (e.g., Antle and Smith, 1986; Gibbons and Murphy, 1990; Aggarwal and Samwick, 1999a, 1999b; Garvey and Milbourn, 2006; Rajgopal et al., 2006; Alburquerque, 2009), but not strong-form RPE. In addition, we find that SIC/GICS based industry classifications depict evidence of weak-form RPE only because they are correlated with TNIC-based classifications. In other words, including TNIC-based peer performance in the regression drives out the explanatory power of SIC/GICS-based industry classifications.

Second, we exploit the time-varying nature of TNIC-classifications. We find that current stock performance of firms that have already exited the focal firm's product markets (i.e., past peers) as well as current stock performance of firms that will enter the focal firm's product markets in future periods (i.e., future peers) do not in general provide information about common performance while only current performance of current peers does. These results not only indicate that TNIC classifications incorporate the evolving nature of product strategies but also suggest that we are not capturing some mechanical aspect connecting the focal firm with these peers.

[8] Holmstrom (1982) specifically states that "we would expect that with many agents we would be able to achieve approximately the same solution as if there were no common uncertainty at all."

Finally, we perform cross-sectional tests to examine situations where corporate boards might not want to filter out common performance. For example, Gopalan et al. (2010) suggest that common performance should not be filtered out when a firm's exposure to common external shocks is the CEO's choice (i.e., strategic flexibility). Consistent with their theory, we find that our RPE effect is weaker in situations with greater strategic flexibility on part of the CEO.

Our study contributes to the compensation literature in three key ways. First, we hypothesize (and find confirmatory evidence) that a key identification strategy to testing RPE theory lies in accurately defining the peer group. In this sense, our study builds on the recent literature arguing that prior RPE research has failed to find the empirical evidence of RPE due to the incorrectly identified RPE peers (e.g., Albuquerque, 2009; Gong et al., 2011; Lewellen, 2013). In contrast to using static, predefined industry classifications, we exploit recent advances in textual analysis to identify product-market peers and find evidence consistent with RPE.

Second, we are able to confirm more stringent predictions of RPE theory, i.e., the extent of common performance filtering should increase with the number of peers, and that with a sufficiently large number of peers, the optimal contract should resemble one with no common uncertainty (Holmstrom, 1982). Prior attempts to find evidence supporting these predictions examine whether product market competition is positively associated with RPE but find mixed evidence (e.g., Aggarwal and Samwick, 1999b; DeFond and Park, 1999; Ali, Klasa and Yeung, 2009).[9] We provide direct evidence consistent with RPE increasing with the number of peers

[9] Aggarwal and Samwick (1999b) find a negative association between product market competition and RPE in compensation contracts. DeFond and Park (1999) find a positive association between competition and RPE in CEO turnover decisions, while Ali, Klasa, and Yeung (2009) fail to replicate DeFond and Park (1999). Ali et al. (2009) point out that the competition measure used in DeFond and Park (i.e., Sales-based HHI) is based on sales of only publicly-traded firms, resulting in a biased measure of competition. Overall, it is a still open question whether greater product market competition is positively associated with RPE in both CEO compensation and turnover decisions. Bushman and Smith (2001), for example, call for research to resolve conflicting results in Aggarwal and Samwick (1999b) and DeFond and Park (1999).

operating in the firm's product markets and with the firm optimally using RPE in the presence of a sufficiently large number of peers.

Third, our results speak to the long-standing debate about optimal contracting versus rent-extraction in explaining CEO compensation. While the presence of "pay-for-luck" is often cited as evidence in favor of CEO rent-extraction, our results suggest that this phenomenon is less prevalent in firms where market participants are privy to a relatively large number of reference points with respect to CEO compensation. A fuller exploration of the role of corporate governance in the use of product-market peers based RPE is a fruitful area for further exploration.

This paper proceeds as follows. In the next section, we discuss relevant literature and develop hypotheses. Next, we discuss empirical specifications to test RPE theory in CEO compensation contracts, and then we present the estimation results including robustness checks. Lastly, we conclude and summarize.

I. Literature Review and Hypothesis development

A. Relevant Literature

Holmstrom (1979) predicts that when the agent's efforts are unobservable and non-contractible, the second-best contracting mechanism is to provide an incentive contract where the agent's compensation is contingent on observable measures of firm performance. Consistent with this prediction, prior research shows that revisions in CEO compensation are positively associated with the firm's own stock returns (i.e., positive pay-for-performance sensitivity; Jensen and Murphy, 1990; Aggarwal and Samwick, 1999a). This incentive contract, however, imposes unnecessary risk on the risk-averse agent to the extent that firm performance is influenced by external shocks that are not under the agent's control. These uncontrollable shocks potentially decrease the utility of the agent thereby reducing contracting efficiency. One solution as Holmstrom (1982) points out is to filter out these external shocks from firm performance, thereby resulting in a greater "signal-to-noise" ratio, which in turn results in greater contracting efficiency.[10] That is, the agent should not be rewarded solely for his/her own *total* performance but rather for performance relative to that of his/her peers. This is the idea behind the Relative Performance Evaluation (RPE) theory.

Prior studies have attempted to test this RPE theory in CEO compensation contracts (e.g., Antle and Smith, 1986; Gibbons and Murphy, 1990; Jensen and Murphy, 1990; Janakiraman, Lambert, and Larcker, 1992; Aggarwal and Samwick, 1999a, among others). However, this evidence is mixed at best (Prendergast, 1999; Lambert, 2001; Frydman and Jenter, 2013), which in turn has resulted in alternative theories that seek to explain this RPE "puzzle".

For example, Bertrand and Mullainathan (2001) argue the lack of RPE is attributed to the rent-seeking behavior of managers. They argue that firms with weak corporate governance are less likely to use RPE because CEOs in these firms can affect their pay-setting process and are paid for positive external shocks but not penalized for negative external shocks (i.e., pay-for-luck). Another stream of research seeks to find cross-sectional evidence of RPE by identifying factors that alter the costs and benefits of using RPE. For instance, Gopalan et al. (2010) show that if the exposure to common external shocks is a strategic choice of the CEO (i.e., strategic flexibility), then RPE is less likely to be used in compensation contracts.

[10] Holmstrom (1982) specifically assumes *homogeneous* agents in the same team, and each agent's performance (x_i) is determined by common uncertainty parameter (η), which affects all agents in the same team, and idiosyncratic error term, which is determined by the agent-specific efforts (i.e., $x_i = \eta + e_i$). By aggregating performance of all agents in the same team, the idiosyncratic error terms are averaged out in the aggregate performance index, and thus the measure of common uncertainty parameter can be estimated.

A third stream of research seeks to test RPE by limiting the set of firms within the industry group that can be considered peers. For instance, Albuquerque (2009) argues that using the entire SIC group as peers is problematic because all firms in this group may not face common external shocks and firms' abilities to respond to common shocks is likely to vary substantially within the same industry. Albuquerque (2009) refines the set the peers within the focal firm's two-digit SIC industry group to those in the same size quartile portfolio and finds evidence consistent with RPE. In a similar vein, Dikolli, Hofmann and Pfeiffer (2011) show analytically that aggregating heterogeneous firm performance within the same industry adds significant summarization bias in the measure of common shocks, leading to the failure in detecting RPE.

While the above studies use firms in the same industry as RPE peers, another stream uses peer firms that are self-disclosed by the firm (e.g., Murphy, 1999; Bannister and Newman, 2003; Carter et al., 2009; Gong et al., 2011; Lewellen, 2013).[11] Similar to the above, these studies also argue that using all firms in the same industry might lead to noisy measure of common external shocks, resulting in the failure to detect RPE in the data.

B. Product Market Peers

In this study, we hypothesize that the prior lack of RPE evidence is (amongst other things) on account of the use of pre-defined industry classifications such as SIC, GICS and NAICS to define the peer group. We reason that these pre-defined industry classifications group firms based on production functions rather than the similarity in products or outputs (e.g., Bhojraj et al., 2003,

[11] For example, Gong et al. (2011) use compensation disclosures mandated by SEC after 2006 and examine the RPE theory. Interestingly, Gong et al. (2011) do not find evidence of RPE following the method used in Albuquerque (2009), but find evidence of weak-form evidence of RPE using self-selected RPE peers by the firm. Similar to Gong et al. (2011), Lewellen (2013) collects a firm's significant competitors disclosed in the firm's 10-K filings, and finds evidence consistent with RPE.

Guenther and Rosman, 1994).[12] RPE theory assumes homogenous agents in the same team that share the same common uncertainty parameter (Holmstrom, 1982). Arguably, the empirical counterpart of common uncertainty parameter is likely to be common demand and supply shocks that affect all firms producing similar products, rather than, for example, having similar production functions. The distinction is important because having similar production functions does not necessarily imply those firms producing similar products (e.g., Bernard and Skinner, 1996; Brickley and Zimmerman, 2010).

In addition, pre-defined industry classifications rarely change over time and consequently do not capture the evolving nature of the firm's product markets as the firm's product offerings change (Hoberg and Phillips, 2015). A firm enters or exits its peers' product market space if the latter starts or stops producing similar products (not whether or not it uses similar production processes). Although the firm's product market peers also change accordingly in this case, traditional industry classifications do not reflect this as these classifications do not evolve rapidly. Accordingly, pre-defined industry classifications fail to capture this dynamic nature of evolving product markets. If so, RPE tests relying on these pre-defined industry classifications might fail to detect consistent evidence. Gibbons and Murphy (1990, pg. 49) allude to this possibility as "…our inability to detect an industry effect after controlling for market movements may reflect the inappropriateness of industry definitions based on SIC codes for purposes of relative performance evaluation".

[12] http://www.naics.com/info.htm. The Census Department states "NAICS was developed to classify units according to their production function. NAICS results in industries that group units undertaking similar activities using similar resources but does not necessarily group all similar products or outputs." Prior studies also note how these standard classifications ignore within-industry heterogeneity in production processes, resulting in misclassification (e.g., Clarke, 1989; Kahle and Walkling, 1996; Bhojraj et al., 2003; Dopuch et al., 2008; Hoberg and Phillips, 2015).

Hence, we argue that the empirical analysis should identify peer firms producing similar products who face similar demand and supply shocks as RPE peers. To this end, we use Text-based Network Industry Classifications (TNIC) recently developed by Hoberg and Phillips (2015) to identify RPE peers. Hoberg and Phillips (2015) identify peer firms based on the pairwise product similarity scores among firms by parsing firms' product descriptions in annual 10-K filings (Item 1 or 1A). They argue that firms producing similar products are more likely to be peer firms competing in the same product markets.

Hoberg and Phillips (2015) specifically convert each firm's product description in 10-K filings into a word vector and calculate product cosine similarity scores for every pair of firms (i.e., the distance between two word vectors for every pair of firms). For example, a firm *i*'s product similarity score with a firm *j* is calculated as the dot product of the word vector of the firm *i*, which consists of vocabularies describing the firm *i*'s products, and that of the firm *j*. This cosine product similarity score between firm *i* and firm *j* is bounded in [0,1] and increases with the number of same words that both firm *i* and firm *j* use, implying that firm pairs with high cosine similarity scores are likely to operate in the similar product markets. Firm *j* is classified as firm *i*'s product market peer if product similarity score between firm *i* and firm *j* is above a pre-specified minimum similarity threshold.[13] This classification yields a group of product market peers for every firm, which allows peer group composition to vary year-to-year and firm-by-firm. Hoberg and Phillips (2015) argue that this procedure can capture the notion that the most appropriate peer firms are firms producing similar products. In addition, Hoberg and Phillips (2015) also argue that TNIC captures the changing nature of product markets over time because all firms' update their product

[13] Hoberg and Phillips (2015) state that "Although one can use any minimum similarity threshold to construct a classification, we focus on thresholds generating industries with the same fraction of membership pairs as SIC-3 industries, allowing us to compare our industries to SIC-3 in an unbiased fashion."

descriptions annually and the updates are required to be correct and timely by SEC. Hence, we test the RPE theory using TNIC-based peers. This leads to our first RPE prediction:

H1: ***Firms base CEO compensation not only on own firm performance but also filter out performance of their product market peers.***

C. Implication of the Number of Product Market Peers in RPE

While prior studies also find evidence in favor of the RPE hypothesis above, we go one step further and devise more stringent tests based on RPE theory. In particular, we hypothesize that the extent of RPE (i.e., filtering out of peer performance) increases with the number of firms in the peer group. In addition, we predict that common performance is completely filtered out in the presence of a large number of peers (i.e., strong-form RPE).

Holmstrom (1982) predicts that if the number of agents is sufficiently large enough to infer the precise value of common uncertainty parameter, the principal can completely filter out common uncertainty in evaluating the agent's performance. On the other hand, if the number of agents in a team is small, then idiosyncratic performance of agents is not sufficiently eliminated in the aggregation process, resulting in the principal only partially filtering out common uncertainty in evaluating the agent's performance (Gibbons and Murphy, 1990). These hypotheses are stated as follows:

H2: ***The extent of filtering of common performance in CEO compensation increases with the number of product market peers.***

H3: ***Firms completely filter out common performance in CEO compensation in the presence of a large number of product market peers.***

II. Research Design

A. Empirical Specification

We use the empirical specification proposed by Holmstrom and Milgrom (1987) and widely used in prior RPE studies (e.g., Gibbons and Murphy, 1990).

$$\begin{aligned}\Delta\, ln\,(Total\ comp) = \alpha + \beta_1 Firm\ Ret + \beta_2 Peer\ Ret + \beta_3 \Delta Size + \beta_4 \Delta BM \\ + \beta_5 SalesGrowth + \beta_6 Vol + \beta_7 Tenure + \beta_8 Age + \beta_9 Own \\ + \sum Ind - Year + \varepsilon\end{aligned} \tag{1}$$

The dependent variable is the annual change of (the log) of total CEO compensation, measured as the sum of salary, bonus, grant-date fair value of stock and option grants, long term incentive payouts, other annual compensation, and all other annual compensation (i.e., variable TDC1 in ExecuComp). *Firm Ret* captures firm *i*'s own stock price performance and defined as the annual buy-and-hold stock returns including dividends. *Peer Ret* captures average stock performance of firm *i*'s product market peers and is measured as equal-weighted average of annual stock returns of product market peers excluding firm *i*. To define product market peers, we choose one quarter of TNIC peers based on the closeness of size (i.e., market value of equity) and book-to-market within each focal firm's TNIC group and year (e.g., Albuquerque, 2009; Lys and Sabino, 1992).[14] Firms in the same quartile portfolio as firm *i* (excluding firm *i*) are defined as firm *i*'s product market peers in period *t*.

[14] Specifically, we take the following steps to choose the closest RPE peers in terms of size and BM. First, we merge the latest market value of equity and BM of TNIC peers as of the beginning of the focal firm's fiscal period. We drop peer firm observations with missing values of market value of equity, BM, and annual stock returns. We also drop the 1st and 99th annual stock returns to mitigate the influence of extreme observations. We rank firms (including the focal firm) using size and BM and divide those ranks by the number of peers in each TNIC group and year (i.e., ranks range from 0 to 1). Similar to Hoberg and Phillips (2015), we compute the pair-wise distance score using size and book-to-market rank scores (i.e., square root of [(size rank of the focal firm – size rank of the peer firm)^2 + (BM rank of the focal firm – BM rank of the peer firm)^2]). This approach assumes equal weights on size and BM in selecting RPE peers. Finally, we choose one quarter of TNIC peers that are closest to firm *i*'s in terms of size and BM. Lys and Sabino (1992) show that researchers can maximize the power of their tests by placing 27% of the sample on each of the extreme portfolios. We require a minimum of two peers for each focal firm in each year.

Following prior studies, we include changes in total assets (Δ*Size*) and changes in the book-to-market ratio (Δ*BM*) to control size and growth options (Smith and Watts, 1992). We also include idiosyncratic volatility (*Vol*) following Aggarwal and Samwick (1999a), (the log of) CEO tenure (*Tenure*), (the log of) CEO age (*Age*), and CEO stock ownership (*Own*) to control for the effects of firm and CEO characteristics on firms' compensation policies. We include sales growth because sales is an important explicit performance metric used in compensation contracts along with stock returns (e.g., Huang, Marquardt, and Zhang 2015). We also include industry-year fixed effects to control unobservable time-varying, industry-specific factors (Gormley and Matsa, 2014).[15]

The coefficient on *Firm Ret* in eq. (1) is expected to be positive (i.e., pay-for-performance sensitivity), while that on *Peer Ret* captures the RPE effect and is expected to be negative (i.e., lower compensation for greater common performance). In addition, optimal contracting theory predicts that the sum of the coefficients on *Firm Ret* and *Peer Ret* is statistically zero if common performance is completely filtered out while compensating the CEO, and thus he/she is evaluated solely on idiosyncratic performance (Holmstrom, 1982).

B. Data and Descriptive Statistics

We retrieve market values of equity and stock returns data from CRSP, financial statement data from Compustat, and CEO compensation data from ExecuComp. We adjust delisting returns following Beaver, McNichols, and Price (2007). Following Garvey and Milbourn (2006), we use a sample of ExecuComp firms with at least two consecutive years of data for each CEO during our sample period and require non-negative CEO tenure. We also delete observations with missing financial and compensation data. The above data requirements yield a sample of 22,028 firm-year

[15] For industry fixed effects, we use fixed industry classifications that are also constructed based on the product similarity scores (Hoberg and Phillips, 2015) to be consistent with our use of TNIC.

observations. The sample period ranges from 1996 to 2013 because TNIC are only available for this sample period (Hoberg and Phillips, 2015).

Table 1 presents descriptive statistics. The mean (median) of total compensation is $5.23 million ($2.96 million), which shows significant right skewness as in prior compensation studies (e.g., Albuquerque, 2009). Hence, we take the natural logarithm of total compensation to reduce skewness. Table 2 presents correlations among main variables. We note that changes in total compensation are positively correlated with most performance measures. Changes in total compensation are also positively correlated with sales growth (0.15) and negatively correlated with book-to-market ratio (-0.08), suggesting that large and growth firms incur greater compensation costs to hire talented managers (Smith and Watts, 1992).

III. Empirical Results

A. Validity Check

We begin by validating our main assumption that TNIC better captures a firm's demand and supply shocks in product markets relative to SIC and GICS. Using average sales as a measure of demand shocks, we run a regression of firm *i*'s sales on peer firm average sales based on TNIC, SIC, and GICS (excluding firm *i*). This test allows us to examine the correlation between firm *i*'s sales and peer sales holding the effect of other industry classifications fixed. With regard to supply shocks, we estimate correlations between firm *i's* operating costs (i.e., the sum of cost of goods sold and SG&A expenses) and average operating costs based on TNIC, SIC, and GICS.

Table 3 tabulates these results. In panel A, we examine sales correlations. In column (1) of Panel A, we find that the coefficient on *Average Sales* (*TNIC*) is 0.786 while the coefficients on *Average Sales* (*SIC*) and *Average Sales* (*GICS*) are 0.243 and 0.161, respectively. This result is consistent with our expectation and Hoberg and Phillip (2015) and suggests that TNIC better captures firms' demand shocks than SIC or GICS. Correlations based on operating cost are even stronger. In column (1) of Panel B, the coefficient on *Average Costs* (*TNIC*) is 0.876 while the coefficient on *Average Costs* (*SIC*) and *Average Costs* (*GICS*) is 0.201 and 0.159, respectively. This evidence suggests that TNIC better captures firms' supply shocks than SIC or GICS.

We further examine whether the above results are stronger in more recent periods. If TNIC better captures the evolving nature of product markets and is updated annually, we would expect the TNIC-based correlations to be stronger in more recent periods. To this end, we partition our full sample into four subsamples based on 5-year periods, and estimate sales and operating costs correlations for each period (columns (2) through (5) in each panel). In general, we find that sales and operating costs correlations based on TNIC are increasing in more recent periods. In particular, in Panel B, we find that the correlation between firm *i*'s operating costs and TNIC-peer average operating costs is 0.821 in the 1996-1999 period, 0.891 in the 2000-2004 period and 0.908 in the 2005-2008 period. In contrast, COGS correlations based on SIC are decreasing over time – 0.252 in 1996-1999; 0.188 in 2000-2004 and 0.162 in 2005-2008. These effects become weaker in the most recent post-crisis period of 2009-2013 – in particular, the TNIC-based correlations drop relative to prior periods but are still larger than those based on SIC/GICS.

Overall, findings in Table 3 suggest that TNIC better captures a firm's supply and demand shocks as evidenced by stronger correlations between firms' sales (operating costs) and TNIC-peer average sales (average operating costs). In the next section, we present results of our main RPE hypothesis.

B. Main Results

Table 4 presents the results of eq. (1). We start with the full-sample results in column (1). Consistent with positive pay-for-performance sensitivity, the coefficient on *Firm Ret* is positive (coeff. = 0.303) and significant at the 1% level. Furthermore, consistent with our first RPE prediction, the coefficient on *Peer Ret* is negative (-0.115) and also significant at the 1% level. The evidence supports the weak-form version of RPE (i.e., partial filtering of common performance, Gibbons and Murphy 1990), given that the absolute value of the coefficient on *Peer Ret* is significantly less than that on *Firm Ret* (F-stat = 44.38). Stated in economic terms, CEO compensation increases by 30.3 basis points when the firm experiences a 1% increase in its own stock price, and its peers experience a 0% stock return during the fiscal year. However, if peer stock returns also increased by 1%, the CEO still experiences an 18.8 basis points (0.303-0.115) increase in annual total compensation.

To test our second prediction that the extent of common performance filtering increases with the number of peers (Holmstrom, 1982), we divide the sample into three subsamples based on the number of product market peers (Hoberg and Phillips, 2015) and estimate eq. (1) within each subsample. Firm-year observations that belong to the first, second, and third tercile of the number of product market peers are classified as the Few, Moderate, and Many group, respectively. We expect the coefficient on *Peer Ret* to become increasingly more negative as we move from the Few group to the Moderate group to the Many group. This is exactly what we find – the coefficient on *Peer Ret* is -0.073 in the Few peer group, -0.123 in the Moderate peer group and -0.315 in the Many peer group. The coefficient on *Peer Ret* is significantly different between the Moderate and the Many groups at the 10% level, and also between the Few and the Many groups at the 5% level. Reassuringly, the coefficient on *Firm Ret* does not vary significantly across these subsamples (0.264, 0.334 and 0.320). This is comforting as there is no a priori reason for the extent of pay-for-performance sensitivity to differ based on the size of the peer group – only the extent of RPE is predicted to. In terms of economic significance, when both the firm and its peers experience the same magnitude of common stock return, the positive pay-for-performance sensitivity decreases by 27.6% (-0.073/0.264) in the Few peer group, by 36.8% (-0.123/0.334) in the Moderate peer group, and by 98.4% (-0.315/0.320) in the Many peer group.

The latter result is consistent with our third prediction of complete filtering (i.e., strong-form RPE of Holmstrom, 1982). In particular, the sum of the coefficients on *Firm Ret* and *Peer Ret* is indistinguishable from zero (F-stat = 0.000 p-value = 0.956), suggesting that common performance is completely filtered out while evaluating the CEO. Overall, these results suggest that firms use RPE in rewarding their CEOs, and the extent of RPE usage depends on the presence of a large enough number of product market peers.

IV. Additional Tests

A. Alternative Classifications

In this section, we replicate our results using pre-defined industry classifications. In Panel A of Table 5, we use three-digit SIC codes and calculate *Peer Ret* (*SIC*) based on the same method used in the construction of our main peer return variable using TNIC.[16] In column (1) of Table 5, we find the same result documented in column (1) of Table 4 (i.e., weak-form evidence of RPE). This result is consistent with prior RPE research (e.g., Gibbons and Murphy, 1990; Albuquerque, 2009). Next, we partition the sample into three subsamples based on the number of firms in the same SIC industry and estimate equation (1) within each subsample. While there is some evidence

[16] TNIC is comparable with three-digit SIC because the pre-specified minimum product similarity threshold use in constructing TNIC is set to generate industries with the same fraction of industry pairs as three-digit SIC industries (Hoberg and Phillips, 2015). Results using NAICS are similar and are not tabulated.

that the extent of filtering increases with the number of peers, it is far from monotonic. In other words, there is no difference in the extent of common performance filtering between the Moderate peer and Many peer groups. Even this weak evidence, as we document below, manifests because of the correlation between SIC and TNIC-based classifications. Once the latter is included in the regression, it drives out the former. Moreover, we are also unable to find evidence consistent with complete filtering (i.e., our third prediction) in the Many peers subsample. In particular, the coefficient on *Firm Ret* is 0.285 while that on *Peer Ret* is -0.158, indicating that the CEO continues to enjoy a 12.7 basis points increase in annual compensation even when both the firm and the peer group experience a 1% stock return during the year (F-statistic = 4.14 p-value 0.042).

In Panel B, we use GICS codes to define peers. GICS codes are the most recent and improved industry classification method developed by MSCI Inc. and S&P (e.g., Bhojraj et al., 2003). Here again, while we find evidence consistent with weak-form RPE in the full-sample, we are unable to find evidence consistent with our other two predictions. In particular, the coefficient on *Peer Ret* does not show monotonicity as we move from few peers to moderate peers to many peers – rather it is most negative and significant in the moderate peers subsample. Further, there is no evidence of complete filtering of common performance in the many peers subsample – CEOs continue to enjoy a 14.2 basis points increase in annual compensation when the firm and the peer group both enjoy a 1% annual stock return.

If SIC and GICS are poor proxies for the firm's peer group, why do we observe evidence consistent with partial filtering (i.e., weak-form evidence of RPE) using these classifications? And also some weak evidence of monotonicity using SIC-based peers? We conjecture that this is due to these proxies being correlated with TNIC-based classifications. To examine this possibility, we include both *Peer Ret* based on TNIC and *Peer Ret* based on these alternative pre-defined industry classifications simultaneously in the same regression. Table 6 presents these results. Consistent with our conjecture, we find that in a multivariate specification, only the coefficient on *Peer Ret* (*TNIC*) remains negative and significant, while those on *Peer Ret* (*SIC*) and *Peer Ret* (*GICS*) become insignificant. These results suggest that proxies for peer performance based on pre-defined industry classifications appear to provide evidence consistent with RPE only because they are correlated with product-market based TNIC industry classifications.

B. Dynamic Peer Groupings

As noted earlier, one of the key advantages of using TNIC to identify RPE peers is that TNIC captures the evolving nature of product markets. Therefore, we can examine whether current stock returns of past, current, and future product market peers contain information about common performance. For example, consider *past* peer firm *j* that was the product market peer of firm *i* in period *t-1* but not in period *t* (i.e., firm *j* exited firm *i*'s product space in period *t-1*). In this case, firm *j*'s current stock return in period *t* is less likely to contain information regarding common demand and supply shocks that firm *i* faces in period *t*. Similarly, if firm *k* is not a product market peer of firm *i* in period *t* but is expected to be a peer in period *t+1* (i.e., future peer), then the stock returns of firm *k* in period *t* are also less likely to contain relevant information about common shocks that firm *i* is experiencing in period *t*. In reality, entering new product markets takes time, and hence firm *k* is most likely taking some activities to enter the new product market in the current period *t* (e.g., investments), resulting in firm *k*'s stock returns in period *t* presumably containing information regarding common external shocks. Foucault and Fresard (2014) adopt this approach and show that past (future) peers' stock price is not (weakly) associated with the focal firm's investment while present peers' stock price is informative to the focal firm's investment.

Similar to Foucault and Fresard (2014), we classify firm-year observations into four sets of peer firms: (1) past peers, (2) new peers, (3) current peers, and (4) future peers. We define *Past Peers* as firms that were firm *i*'s product market peers in period *t-1* but are not in the same TNIC group in period *t*. *New Peers* are firm *i*'s product market peers in period *t* but were not in the same TNIC group in period *t-1*. *Current Peers* are firm *i*'s product market peers in period *t-1* as well as in period *t.* Lastly, we define *Future Peers* as firms that will be firm *i*'s product market peers in period *t+1* but are not in the same TNIC group in period *t*. We then calculate equal-weighted stock returns of each set of peers using stock returns in period *t* and replace *Peer Ret* in equation (1) with each of these stock returns.

Table 7 reports the results. Consistent with our expectations, in column (1) of Panel A, the current period returns of past peers (*Peer Ret_Past*) is not associated with current changes in total CEO compensation, suggesting that current stock returns of past peers do not contain information regarding common shocks. In columns (2) and (3), we find that the coefficients on *Peer Ret_New* and *Peer Ret_Current* are significantly negative at 1% level (i.e., both of them are product market peers in period *t*). In column (4), we find that *Peer Ret_Future* is not statistically significant, suggesting that current stock returns of firms that are expected to enter firm *i*'s product markets in the next period also do not contain information concerning common shocks. Overall, this time-series evidence not only corroborates our RPE hypothesis, but also mitigates concerns that we are capturing some mechanical feature linking our focal firm to these product-market peers.

C. Cross-sectional Test: Strategic Flexibility and RPE

Finally, we investigate cross-sectional variation in the use of RPE. Gopalan et al. (2010) propose a model showing that the use of RPE decreases if firms want to provide strategic flexibility to their CEOs. Gopalan et al. argue that "the board of directors is not primarily concerned with how hard the CEO is actually working, but whether she has the vision to choose the right strategy for deploying the firm's assets. In doing so, the CEO's concern is with the firm's strategic direction in lieu of its surrounding market environment." Put differently, if the effect of common external shocks on firm performance is not random but under the CEO's control, then the effect of common external shocks should not be excluded in evaluating the CEO's efforts.

Following Gopalan et al., we use two alternative proxies to identify firms that offer greater strategic flexibility to the CEO. First, we use the firm's market-to-book ratio. Firms with high market-to-book ratios are more likely to have greater growth options and thus are more likely to provide their CEOs with greater strategic flexibility to allow more discretion in exercising those options. We classify firm-years with market-to-book ratios above the median as offering greater strategic flexibility to the CEO. Second, we use the peer-adjusted stock returns during the previous year as a measure of CEO talent. Gopalan et al. predicts that RPE is reduced for more talented CEOs due to the decreasing disutility of effort for more talented CEOs. We classify firm-years with positive peer-adjusted stock returns during period *t-1* as having more talented CEOs because firms managed by more talented CEOs are more likely to exhibit better peer-adjusted stock performance.[17]

Gopalan et al. further argue that if less RPE allows CEOs to have greater strategic flexibility, we expect to observe some evidence that CEOs with less RPE exploit the strategic flexibility to a greater extent at the firm level. Hence, following Gopalan et al., we identify firms

[17] Gopalan et al. also examine whether multi-segment firms (based on the SIC industry) are less likely to use RPE. We do not examine this variable because segment information in Compustat is only based on SIC industries, which we do not rely on in our study. Gopalan et al. also use R&D expenditures to test the theory. In untabulated tests, we find that the extent of RPE in firms with high R&D expenditures is not significantly different from RPE in firms with low R&D expenditures. This result could be attributed to R&D expenditures being a noisier measure of the firm's growth options because a significant portion of firms in Compustat universe does not report R&D expenditures separately.

with positive asset growth in subsequent period as exploiting their strategic flexibility to a greater extent and examine whether firms with positive (negative) asset growth in period *t+1* are less (more) likely to use RPE in period *t*.

Table 8 presents the results. In column (1) and (2), we divide the full sample into two subsamples based on median market-to-book. We find that the coefficient on *Peer Ret* for high market-to-book firms in column (2) is statistically greater at the 5% level than that for low market-to-book firms in column (1), suggesting that firms providing their CEOs with greater strategic flexibility use less RPE. In column (3) and (4), we use the peer-adjusted stock returns as a measure of the CEO talent, and find similar results. The coefficient on *Peer Ret* is statistically greater at 10% level in column (4) (i.e., greater CEO talent) relative to the coefficient in column (3) (i.e., lower CEO talent). In column (5) and (6), we use the asset growth rate in period *t+1* to investigate whether CEOs with less RPE exploit their strategic flexibility in subsequent period to a greater extent at the firm level. Consistent with our expectations, we find that firms with asset growth in period *t+1* filter out common shocks to a lesser extent in period *t* (i.e., less RPE) as evidenced by the significantly greater coefficient on *Peer Ret* in column (6) (i.e., asset growth in period *t+1*) relative that in column (7) (i.e., asset decline in period *t+1*). In sum, the results in Table 8 are in general consistent with Gopalan et al. and suggest that the use of RPE is attenuated by the board's desire to promote strategic flexibility on part of the CEO.

V. Conclusion

This study re-examines the RPE puzzle using product-market peers identified by textual analysis of firms' product descriptions in 10-K filings (Hoberg and Phillip, 2015). In contrast to the mixed evidence of RPE documented in prior studies, we find three pieces of evidence consistent with RPE – (i) firms on average filter out common shocks to performance measures, (ii) the extent of filtering increases with the number of peers, and (iii) firms completely filter out common shocks in the presence of a large number of peers. We are able to replicate the first finding but not the other two using pre-defined industry classifications. Overall, our results suggest that a key identification strategy to testing RPE theory lies in accurately defining the peer group.

APPENDIX: Variable Definitions

Variable	Definition
Age	*Age* is defined as CEO age variable in ExecuComp as of the beginning of period *t*. When this variable is used in the regression, we take the natural logarithm of this variable.
Δ*BM*	Δ*BM* is measured as the change in firm *i*'s Book-to-Market in year *t-1*. Book-to-Market is measured as book value of equity divided by market value of equity. Book value of equity is measured by shareholders' equity plus deferred tax and investment credit minus preferred stock. Market value of equity is obtained from CRSP and is calculated by the number of common shares outstanding multiplied by share price.
Firm Ret	*Firm Ret* is measured as the natural logarithm of one plus firm *i*'s annual buy-and-hold stock return in period *t*.
# of Peers	*# of Peers* is measured as the number of product market peers for firm *i* in period *t*.
Own	*Own* is calculated as the number of shares owned by CEO excluding option divided by the number of shares outstanding for firm *i* as of the beginning of period *t*.
Peer Ret	*Peer Ret* is measured as the natural logarithm of one plus equal-weighted annual returns of firm *i*'s product market peers in period *t*. To define product market peers, we choose one quarter of TNIC peers based on the closeness of market value of equity and book-to-market as of the beginning of the fiscal period within the same TNIC group as firm *i* in period *t*. We require firm *i* to have a minimum of two peer firms in each period.
SalesGrowth	*SalesGrowth* is measured as the natural logarithm of sales in period *t* divided by the natural logarithm of sales in period *t-1*.
Δ*Size*	Δ*Size* is measured as the change in the natural logarithm of firm *i*'s total assets in period *t-1*.

Tenure	*Tenure* is defined as the difference between the *BECAMECEO* variable in ExecuComp and the date of fiscal year-end for firm *i* as of the beginning of period *t* divided by 365. When this variable is used in the regression, we take the natural logarithm of this variable.
Total Comp	*Total Comp* is *TDC1* in ExecuComp, which is measured by the sum of salary, bonus, long-term incentive payouts, fair value of stock and option grants, and all other compensation for firm *i* in period *t*.
Δ*Total Comp*	Δ*Total Comp* is measured as changes in total CEO compensation for firm *i* in period *t*.
Δ*ln*(*Total Comp*)	Δ*ln*(*Total Comp*) is measured as the change in the natural logarithm of one plus total CEO compensation for firm *i* in period *t*.
Vol	*Vol* measures idiosyncratic return volatility and defined as the standard deviations of residuals from the regression of firm *i*'s monthly returns on monthly equal-weighted average returns of product market peers (described above) using preceding past 24 months (a minimum of 8 observations is required).

REFERENCES

Abowd, J. M., and Kaplan, D. S., 1999. Executive Compensation: Six Questions That Need Answering. *Journal of Economic Perspectives*, *13*(4), 145-168.

Aggarwal, R., and Samwick, A., 1999a. The Other Side of the Trade-Off: The Impact of Risk on Executive Compensation. *Journal of Political Economy* 107, no. 1: 65–105.

Aggarwal, R., and Samwick, A., 1999b. Executive Compensation, Strategic Competition, and Relative Performance Evaluation: Theory and Evidence. *The Journal of Finance* 54, no. 6.

Albuquerque, A., 2009. Peer Firms in Relative Performance Evaluation. *Journal of Accounting and Economics* 48, no. 1: 69–89.

Ali, A., Klasa, S., and Yeung, E., 2009. The limitations of industry concentration measures constructed with Compustat data: Implications for finance research. *Review of Financial Studies*, *22*(10), 3839-3871.

Antle, R., and Smith, A., 1986. An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives. *Journal of Accounting Research* 24, no. 1: 1–39.

Bannister, J., and Newman, H., 2003. Analysis of Corporate Disclosures on Relative Performance Evaluation. *Accounting Horizons* 17, no. 3: 235–246.

Barro, J. R., and Barro, R. J., 1990. Pay, Performance, and Turnover of Bank CEOs. *Journal of Labor Economics*, 448-481.

Beaver, W., McNichols, M., and Price, R., 2007. Delisting returns and their effect on accounting-based market anomalies. *Journal of Accounting and Economics*, *43*(2), 341-368.

Bernard, V., and Skinner, D., 1996. What Motivates Managers' Choice of Discretionary Accruals? *Journal of Accounting and Economics* 22, no. 1: 313–325.

Bertrand, M., and Mullainathan, S., 2001. Are CEOs Rewarded for Luck? The Ones Without Principals Are. *The Quarterly Journal of Economics* 116, no. 3: 901–932.

Bhojraj, S., Lee, C., and Oler, D., 2003. What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research* 41, no. 5: 745–774.

Bizjak, J. M., Lemmon, M. L., and Naveen, L. 2008. Does the use of peer groups contribute to higher pay and less efficient compensation? *Journal of Financial Economics*, *90*(2), 152-168.

Brickley, J., and Zimmerman, J., 2010. Corporate Governance Myths: Comments on Armstrong, Guay, and Weber. *Journal of Accounting and Economics* 50, no. 2: 235–245.

Bushman, R., and Smith, A., 2001. Financial Accounting Information and Corporate Governance. *Journal of Accounting and Economics* 32, no. 1: 237–333.

Carter, M., Ittner, D., and Zechman, S., 2009. Explicit Relative Performance Evaluation in Performance-vested Equity Grants. *Review of Accounting Studies* 14, no. 2–3: 269–306.

Clarke, R., 1989. SICs as Delineators of Economic Markets. *Journal of Business*: 17–31.

DeFond, M., and Park, C., 1999. The Effect of Competition on CEO Turnover. *Journal of Accounting and Economics* 27, no. 1: 35–56.

De Franco, G., Hope, O. K., and Larocque, S., 2012. Analysts' choice of peer companies. *Review of Accounting Studies*, 1-28.

Dikolli, S., Hofmann, C., and Pfeiffer, T., 2011. Relative Performance Evaluation and Peer-performance Summarization Errors. *Review of Accounting Studies*: 1–32.

Dopuch, N., Mashruwala, R., Seethamraju, C., and Zach, T., 2012. The Impact of a Heterogeneous Accrual-generating Process on Empirical Accrual Models. *Journal of Accounting, Auditing & Finance* 27, no. 3: 386–411.

Faulkender, M., and Yang, J., 2010. Inside the black box: The role and composition of compensation peer groups. *Journal of Financial Economics*, *96*(2), 257-270.

Foucault, T., and Fresard, L., 2014. Learning from peers' stock prices and corporate investment. *Journal of Financial Economics*, 111(3), 554-577.

Frydman, C., and Jenter, D., 2010. CEO Compensation. *Annual Review of Financial Economics*, *2*(1), 75-102.

Garvey, G., and Milbourn, T., 2003. Incentive Compensation When Executives Can Hedge the Market: Evidence of Relative Performance Evaluation in the Cross Section. *The Journal of Finance* 58, no. 4: 1557–1582.

Garvey, G., and Milbourn, T., 2006, Asymmetric Benchmarking in Compensation: Executives Are Rewarded for Good Luck but Not Penalized for Bad. *Journal of Financial Economics* 82, no. 1: 197–225.

Gibbons, R., and Murphy, K., 1990. Relative Performance Evaluation for Chief Executive Officers. *Industrial and Labor Relations Review*: 30S–51S.

Gormley, T. A., and Matsa, D. A., 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies*, 27(2), 617-661.

Gong, G., Li, L. Y., and Shin, J. Y., 2011. Relative Performance Evaluation and Related Peer Groups in Executive Compensation Contracts. *The Accounting Review* 86, no. 3 (2011): 1007–1043.

Gopalan, R., Milbourn, T., and Song. F., 2010. Strategic Flexibility and the Optimality of Pay for Sector Performance. *Review of Financial Studies* 23, no. 5: 2060–2098.

Guenther, D., and Rosman, A., 1994. Differences Between COMPUSTAT and CRSP SIC Codes and Related Effects on Research. *Journal of Accounting and Economics* 18, no. 1: 115–128.

Hoberg, G., and Phillips, G., 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies*, *23*(10), 3773-3811.

Hoberg, G., and Phillips, G., 2015. Text-Based Network Industries and Endogenous Product Differentiation, *Journal of Political Economy*, forthcoming

Holmstrom, B., 1979. Moral Hazard and Observability. *The Bell Journal of Economics*: 74–91.

Holmstrom, B., 1982. Moral Hazard in Teams. *The Bell Journal of Economics*: 324–340.

Holmstrom, B., and Milgrom, P., 1987. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica: Journal of the Econometric Society*: 303–328.

Huang, R., Marquardt, C. A., and Zhang, B., 2015. Using sales revenue as a performance measure, Working paper.

Janakiraman, S., Lambert, R., and Larcker, D., 1992. An Empirical Investigation of the Relative Performance Evaluation Hypothesis. *Journal of Accounting Research* 30, no. 1: 53–69.

Jensen, M., and Murphy, K., 1990. Performance Pay and Top-Management Incentives. *The Journal of Political Economy* 98, no. 2: 225–264.

Kahle, K. and Walkling, R., 1996. The Impact of Industry Classifications on Financial Research. *Journal of Financial and Quantitative Analysis* 31, no. 3.

Lambert, R., 2001. Contracting Theory and Accounting. *Journal of Accounting and Economics* 32, no. 1: 3–87.

Lee, C. M., Ma, P., and Wang, C. C. (2015). Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics*, *116*(2), 410-431.

Lewellen, S., 2013, Executive compensation and peer effects, *working paper*, Yale University

Lys, T., and Sabino, J. S., 1992. Research design issues in grouping-based tests. *Journal of Financial Economics*, *32*(3), 355-387.

Milbourn, T., 2003. CEO Reputation and Stock-based Compensation. *Journal of Financial Economics* 68, no. 2: 233–262.

Murphy, K., 1999. Executive Compensation. *Handbook of Labor Economics* 3: 2485–2563.

Prendergast, C., 1999. The Provision of Incentives in Firms. *Journal of Economic Literature* 37, no. 1: 7–63.

Rajgopal, S., Shevlin, T., and Zamora, V., 2006. CEOs' Outside Employment Opportunities and the Lack of Relative Performance Evaluation in Compensation Contracts. *The Journal of Finance* 61, no. 4: 1813–1844.

Smith, C. W., and Watts, R. L., 1992. The investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Financial Economics*, *32*(3), 263-292.

Table I
Descriptive Statistics

This table reports descriptive statistics for all sample firms with available information. The sample period ranges from 1996 to 2013. *Total Comp* is *TDC1* in ExecuComp, which is measured by the sum of salary, bonus, long-term incentive payouts, fair value of stock and option grants, and all other compensation for firm *i* in period *t*. Δ*Total Comp* is measured as changes in total CEO compensation for firm *i* in period *t*. Δ*ln*(*Total Comp*) is measured as the change in the natural logarithm of one plus total CEO compensation for firm *i* in period *t*. *Firm Ret* is measured as the natural logarithm of one plus firm *i*'s annual buy-and-hold stock return in period *t*. *Peer Ret* is measured as the natural logarithm of one plus equal-weighted annual returns of firm *i*'s product market peers in period *t*. To define product market peers, we choose one quarter of TNIC peers based on the closeness of market value of equity and book-to-market as of the beginning of the fiscal period within the same TNIC group as firm *i* in period *t*. *Peer Ret* (*SIC*) and *Peer Ret* (*GICS*) are measured using SIC and GICS, respectively. *# of Peers* is measured as the number of product market peers for firm *i* in period *t*. *# of Peers* (*SIC*) and *# of Peers* (*GICS*) are the number of SIC and GICS peers, respectively. Δ*Size* is measured as the change in the natural logarithm of firm *i*'s total assets in period *t-1*. Δ*BM* is measured as the change in firm *i*'s Book-to-Market in period *t-1*. *SalesGrowth* is measured as the natural logarithm of sales in period *t* divided by the natural logarithm of sales in period *t-1*. *Tenure* is defined as the difference between the *BECAMECEO* variable in ExecuComp and the date of fiscal year-end for firm *i* as of the beginning of period *t* divided by 365. *Age* is defined as CEO age variable in ExecuComp as of the beginning of period *t*. *Own* is calculated as the number of shares owned by CEO excluding option divided by the number of shares outstanding for firm *i* as of the beginning of period *t*. All variables are also defined in the Appendix and all continuous variables are winsorized at the 1st and 99th percentiles.

	N	Mean	Std	Q1	Median	Q3
Total Comp	22,028	5,233.76	9,619.31	1,406.94	2,960.54	6,094.10
Δ*Total Comp*	22,028	179.38	10,489.85	-527.16	129.57	1,016.74
Δ*ln*(*Total Comp*)	22,028	0.05	0.65	-0.20	0.06	0.35
Firm Ret	22,028	0.04	0.48	-0.16	0.10	0.30
Peer Ret	22,028	0.07	0.31	-0.08	0.11	0.26
Peer Ret (*SIC*)	20,072	0.06	0.34	-0.11	0.10	0.26
Peer Ret (*GICS*)	20,326	0.06	0.32	-0.10	0.10	0.26
# of Peers	22,028	19.65	28.02	3.00	9.00	23.00
# of Peers (*SIC*)	20,072	20.00	27.74	3.00	6.00	26.00
# of Peers (*GICS*)	20,326	16.65	23.20	5.00	10.00	19.00
Δ*Size*	22,028	0.10	0.22	0.00	0.07	0.17
Δ*BM*	22,028	0.01	0.31	-0.09	0.00	0.10
SalesGrowth	22,028	0.08	0.20	-0.01	0.07	0.16
Vol	22,028	0.09	0.05	0.06	0.08	0.12
Tenure	22,028	7.87	7.50	2.67	5.51	10.51
Age	22,028	55.12	7.21	50.00	55.00	60.00
Own	22,028	0.02	0.06	0.00	0.00	0.01

Table II
Correlation

This table presents Pearson (Above) / Spearman (Below) correlations. Correlations that are significant at 1% level are bolded. The sample period ranges from 1996 to 2013. *Total Comp* is *TDC1* in ExecuComp, which is measured by the sum of salary, bonus, long-term incentive payouts, fair value of stock and option grants, and all other compensation for firm *i* in period *t*. Δ*Total Comp* is measured as changes in total CEO compensation for firm *i* in period *t*. Δ*ln*(*Total Comp*) is measured as the change in the natural logarithm of one plus total CEO compensation for firm *i* in period *t*. *Firm Ret* is measured as the natural logarithm of one plus firm *i*'s annual buy-and-hold stock return in period *t*. *Peer Ret* is measured as the natural logarithm of one plus equal-weighted annual returns of firm *i*'s product market peers in period *t*. To define product market peers, we choose one quarter of TNIC peers based on the closeness of market value of equity and book-to-market as of the beginning of the fiscal period within the same TNIC group as firm *i* in period *t*. *Peer Ret* (*SIC*) and *Peer Ret* (*GICS*) are measured using SIC and GICS, respectively. *# of Peers* is measured as the number of product market peers for firm *i* in period *t*. *# of Peers* (*SIC*) and *# of Peers* (*GICS*) are the number of SIC and GICS peers, respectively. Δ*Size* is measured as the change in the natural logarithm of firm *i*'s total assets in period *t-1*. Δ*BM* is measured as the change in firm *i*'s Book-to-Market in period *t-1*. *SalesGrowth* is measured as the natural logarithm of sales in period *t* divided by the natural logarithm of sales in period *t-1*. *Tenure* is defined as the natural log of the difference between the *BECAMECEO* variable in ExecuComp and the date of fiscal year-end for firm *i* as of the beginning of period *t* divided by 365. *Age* is defined as the natural log of CEO age variable in ExecuComp as of the beginning of period *t*. *Own* is calculated as the number of shares owned by CEO excluding option divided by the number of shares outstanding for firm *i* as of the beginning of period *t*. All variables are also defined in Appendix and all continuous variables are winsorized at the 1st and 99th percentiles.

	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	Δ*ln*(*Total Comp*)	-	**0.19**	**0.07**	**0.07**	**0.07**	0.00	**-0.08**	**0.15**	-0.03	0.03	0.00	0.00
(2)	*Firm Ret*	**0.21**	-	**0.56**	**0.56**	**0.57**	**-0.10**	**0.03**	**0.16**	**-0.05**	0.01	0.03	0.00
(3)	*Peer Ret*	**0.07**	**0.56**	-	**0.74**	**0.78**	**-0.10**	**0.04**	**0.04**	**-0.05**	0.01	**0.05**	-0.01
(4)	*Peer Ret* (*SIC*)	**0.08**	**0.56**	**0.74**	-	**0.79**	**-0.10**	**0.03**	**0.04**	**-0.03**	0.01	**0.04**	-0.01
(5)	*Peer Ret* (*GICS*)	**0.08**	**0.58**	**0.77**	**0.78**	-	**-0.10**	**0.04**	**0.03**	-0.02	0.01	**0.03**	-0.01
(6)	Δ*Size*	0.00	**-0.07**	**-0.09**	**-0.09**	**-0.09**	-	**0.08**	**0.35**	0.00	**0.08**	**-0.06**	**0.04**
(7)	Δ*BM*	**-0.12**	**0.03**	**0.03**	**0.04**	**0.04**	**0.06**	-	**-0.18**	**-0.02**	**0.02**	0.00	0.01
(8)	*SalesGrowth*	**0.18**	**0.17**	**0.03**	**0.03**	**0.03**	**0.39**	**-0.21**	-	-0.02	**0.06**	**-0.05**	**0.04**
(9)	*Vol*	-0.01	**-0.03**	**-0.03**	**-0.02**	-0.01	-0.03	**0.04**	0.01	-	-0.01	**-0.16**	**0.11**
(10)	*Tenure*	0.00	0.01	0.01	0.01	0.01	**0.12**	**0.03**	**0.08**	0.01	-	**0.36**	**0.35**
(11)	*Age*	-0.01	0.02	**0.03**	**0.03**	**0.02**	**-0.03**	0.00	**-0.05**	**-0.16**	**0.33**	-	**0.13**
(12)	*Own*	-0.01	0.01	0.00	0.00	0.00	**0.08**	0.02	**0.07**	**0.23**	**0.48**	**0.13**	-

Table III
Revenue / Cost Correlations

Panel A presents estimation results from the regression of firm *i*'s sales in period *t* on average sales using various peer group definitions in period *t* (excluding firm *i*). Panel B presents estimation results from the regression of firm *i*'s operating costs (the sum of cost of goods sold and SG&A expenses) in period *t* on average operating costs using various peer group definitions in period *t* (excluding firm *i*). In each panel, column (1) reports estimation results using full sample, and column (2) through (5) present estimation results conditional on time periods denoted in each column. All variables are winsorized at 1% and 99% levels and standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

Panel A Sales Correlations

	Dependent Variable: *Sales*				
Independent Variables	(1)	(2)	(3)	(4)	(5)
	Full Sample	1996 - 1999	2000 - 2004	2005 - 2008	2009 - 2013
Average Sales (TNIC)	**0.786*****	**0.800*****	**0.813*****	**0.812*****	**0.705*****
	(37.631)	**(29.786)**	**(32.818)**	**(25.417)**	**(19.792)**
Average Sales (SIC)	0.243***	0.227***	0.222***	0.228***	0.314***
	(10.694)	(8.280)	(7.978)	(6.977)	(8.630)
Average Sales (GICS)	0.161***	0.150***	0.151***	0.163***	0.198***
	(7.674)	(5.337)	(5.789)	(4.817)	(5.217)
Constant	-0.170***	-0.178***	-0.158***	-0.175***	-0.176***
	(-14.142)	(-11.242)	(-10.311)	(-10.675)	(-10.244)
# of observations	67,556	18,115	21,205	13,716	14,520
Adjusted R-squared	0.543	0.517	0.540	0.569	0.552

Panel B Cost Correlations

	Dependent Variable: *Costs*				
Independent Variables	(1)	(2)	(3)	(4)	(5)
	Full Sample	1996 - 1999	2000 - 2004	2005 - 2008	2009 - 2013
Average Costs (TNIC)	**0.876*****	**0.821*****	**0.891*****	**0.908*****	**0.888*****
	(44.826)	**(29.419)**	**(36.759)**	**(29.577)**	**(27.039)**
Average Costs (SIC)	0.201***	0.252***	0.188***	0.162***	0.204***
	(10.068)	(8.921)	(7.284)	(5.362)	(6.815)
Average Costs (GICS)	0.159***	0.135***	0.149***	0.198***	0.169***
	(8.178)	(4.814)	(6.358)	(6.179)	(5.232)
Constant	-0.276***	-0.242***	-0.271***	-0.308***	-0.297***
	(-17.294)	(-12.317)	(-13.243)	(-13.373)	(-12.935)
# of observations	67,556	18,115	21,205	13,716	14,520
Adjusted R-squared	0.447	0.426	0.440	0.462	0.459

Table IV
Tests of Relative Performance Evaluation Hypothesis

This table presents results obtained from the regression of the change in the natural logarithm of total CEO compensation on own firm stock returns (*Firm Ret*), peer firm average returns based on TNIC (*Peer Ret*), control variables, and industry-year fixed effects over the sample period between 1996 and 2013. In column (2), (3), and (4) the full sample is divided into three subsamples based on the number of firms in the same TNIC. Results testing strong-form evidence of RPE and the coefficient differences are summarized toward the bottom of the table. All continuous variables are winsorized at 1% and 99% levels and are defined in the Appendix. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

	Dependent Variable: Δln(*Total Comp*)			
Independent Variables	(1)	(2)	(3)	(4)
		# of TNIC Peers		
	Full Sample	Few	Moderate	Many
Firm Ret (β_1)	0.303***	0.264***	0.334***	0.320***
	(19.505)	(9.167)	(10.511)	(11.068)
***Peer Ret* ($\boldsymbol{\beta_2}$)**	**-0.115*** **	**-0.073** **	**-0.123** **	**-0.315*** **
	(-4.156)	**(-2.019)**	**(-2.216)**	**(-3.472)**
Δ*Size*	-0.045	-0.113**	-0.018	0.005
	(-1.444)	(-2.015)	(-0.324)	(0.087)
Δ*BM*	-0.099***	-0.082**	-0.098**	-0.125***
	(-4.885)	(-2.409)	(-2.515)	(-2.991)
SalesGrowth	0.278***	0.403***	0.222***	0.225***
	(8.030)	(5.994)	(3.561)	(3.795)
Vol	-0.055	-0.126	0.033	-0.068
	(-0.500)	(-0.571)	(0.136)	(-0.313)
Tenure	0.023***	0.013	0.029**	0.021**
	(4.286)	(1.090)	(2.561)	(2.161)
Age	-0.076***	0.024	-0.162**	-0.063
	(-2.727)	(0.419)	(-2.379)	(-1.199)
Own	-0.003	0.037	-0.033	0.357**
	(-0.041)	(0.280)	(-0.242)	(2.520)
Constant	0.299***	-0.077	0.619**	0.252
	(2.701)	(-0.337)	(2.292)	(1.227)
Strong RPE F-Stat	**44.380**	**21.560**	**13.710**	**0.000**
p-value ($\boldsymbol{\beta_1 + \beta_2 = 0}$)	**0.000**	**0.000**	**0.000**	**0.956**
Industry-Year FE	Yes	Yes	Yes	Yes
# of observations	22,028	7,376	7,280	7,372
Adjusted R-squared	0.090	0.137	0.122	0.097
Coefficient Difference	$\Delta\beta_1$	p-value	$\Delta\beta_2$	p-value
Few versus Moderate	0.070	(0.106)	-0.050	(0.451)
Moderate versus Many	-0.015	(0.736)	-0.191*	(0.072)
Few versus Many	0.056	(0.172)	-0.242**	(0.017)

Table V
RPE tests using alternative industry classifications

This table presents results obtained from the regression of the change in the natural logarithm of total CEO compensation on own firm stock returns (*Firm Ret*), peer firm average returns based on SIC (*Peer Ret* (*SIC*)) in Panel A, or peer firm average returns based on GICS (*Peer Ret* (*GICS*)) in Panel B, control variables, and industry-year fixed effects over the sample period between 1996 and 2013. In column (2), (3), and (4) the full sample is divided into three subsamples based on the number of firms in the same industry used in each panel. Results testing strong-form evidence of RPE and the coefficient differences are summarized toward the bottom of each panel. All continuous variables are winsorized at 1% and 99% levels and are defined in the Appendix. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

Panel A SIC

	Dependent Variable: $\Delta ln(Total\ Comp)$			
Independent Variables	(1)	(2)	(3)	(4)
		# of SIC Peers		
	Full Sample	Few	Moderate	Many
Firm Ret	0.294***	0.316***	0.310***	0.285***
	(17.602)	(9.606)	(10.523)	(11.284)
***Peer Ret* (*SIC*)**	**-0.096*** **	**-0.004**	**-0.101** **	**-0.158** **
	(-2.907)	**(-0.085)**	**(-1.985)**	**(-2.385)**
Δ*Size*	-0.046	-0.181***	-0.036	0.004
	(-1.385)	(-2.735)	(-0.646)	(0.081)
Δ*BM*	-0.106***	-0.062	-0.076**	-0.171***
	(-4.866)	(-1.616)	(-2.356)	(-3.969)
SalesGrowth	0.263***	0.371***	0.220***	0.249***
	(6.924)	(4.368)	(3.683)	(4.628)
Vol	0.027	0.345	-0.031	-0.084
	(0.240)	(1.514)	(-0.161)	(-0.483)
Tenure	0.023***	0.016	0.020**	0.032***
	(4.052)	(1.298)	(2.133)	(3.389)
Age	-0.080***	-0.114*	-0.121**	-0.038
	(-2.769)	(-1.748)	(-2.239)	(-0.858)
Own	-0.038	-0.011	-0.072	-0.012
	(-0.500)	(-0.095)	(-0.480)	(-0.107)
Constant	0.304***	0.423*	0.486**	0.121
	(2.661)	(1.649)	(2.288)	(0.694)
Strong RPE F-Stat	**32.310**	**31.520**	**15.230**	**4.140**
p-value ($\beta_1 + \beta_2 = 0$)	**0.000**	**0.000**	**0.000**	**0.042**
Industry-Year FE	Yes	Yes	Yes	Yes
# of observations	20,072	6,466	6,866	6,740
Adjusted R-squared	0.072	0.100	0.102	0.074
Coefficient Difference	$\Delta\beta_1$	p-value	$\Delta\beta_2$	p-value
Few versus Moderate	-0.006	(0.891)	-0.097	(0.156)
Moderate versus Many	-0.024	(0.522)	-0.057	(0.497)
Few versus Many	-0.030	(0.451)	-0.154*	(0.064)

Panel B GICS

Independent Variables	Dependent Variable: Δ*ln*(*Total Comp*) (1)	(2)	(3)	(4)
		# of GICS Peers		
	Full Sample	Few	Moderate	Many
Firm Ret	0.297***	0.300***	0.301***	0.301***
	(18.661)	(10.754)	(10.334)	(11.779)
***Peer Ret* (GICS)**	**-0.113*** **	**-0.037**	**-0.177*** **	**-0.159** **
	(-3.497)	**(-0.907)**	**(-2.684)**	**(-2.299)**
Δ*Size*	-0.066**	-0.121**	-0.136**	0.004
	(-2.129)	(-2.097)	(-2.455)	(0.081)
Δ*BM*	-0.105***	-0.029	-0.163***	-0.136***
	(-5.284)	(-0.906)	(-4.951)	(-3.491)
SalesGrowth	0.290***	0.364***	0.334***	0.213***
	(7.902)	(5.386)	(5.316)	(3.873)
Vol	0.134	0.201	0.127	0.168
	(1.287)	(1.078)	(0.672)	(0.922)
Tenure	0.019***	0.026***	0.012	0.015
	(3.477)	(2.839)	(1.169)	(1.627)
Age	-0.067**	-0.067	-0.092*	-0.028
	(-2.475)	(-1.338)	(-1.728)	(-0.590)
Own	-0.058	-0.274**	0.126	-0.007
	(-0.768)	(-2.304)	(0.783)	(-0.050)
Constant	0.250**	0.231	0.366*	0.098
	(2.350)	(1.180)	(1.724)	(0.532)
Strong RPE F-Stat	**30.790**	**36.130**	**3.330**	**4.650**
p-value ($\beta_1 + \beta_2 = 0$)	**0.000**	**0.000**	**0.069**	**0.031**
Industry-Year FE	Yes	Yes	Yes	Yes
# of observations	20,326	7,018	6,549	6,759
Adjusted R-squared	0.081	0.100	0.089	0.088
Coefficient Difference	$\Delta\beta_1$	p-value	$\Delta\beta_2$	p-value
Few versus Moderate	0.001	(0.977)	-0.140*	(0.075)
Moderate versus Many	-0.000	(0.993)	0.017	(0.857)
Few versus Many	0.001	(0.982)	-0.123	(0.134)

Table VI
Comparison with alternative industry classifications

This table presents results obtained from the regression of the change in the natural logarithm of total CEO compensation on own firm stock returns (*Firm Ret*), peer firm average returns based on TNIC (*Peer Ret*), control variables, and industry-year fixed effects over the sample period between 1996 and 2013. In column (1), we additionally include the *Peer Ret* (*SIC*) variable in the regression. In column (2), we include additionally the *Peer Ret* (*GICS*) variable instead of the *Peer Ret* (*SIC*) variable. In column (3), we simultaneously include the *Peer Ret*, *Peer Ret* (*SIC*), and *Peer Ret* (*GICS*) variables in the same regression. In this table, we include year fixed effects but do not include industry fixed effects. All continuous variables are winsorized at 1% and 99% levels and are defined in the Appendix. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

	Dependent Variable: $\Delta ln(Total\ Comp)$		
Independent Variables	(1)	(2)	(3)
Firm Ret	0.301***	0.302***	0.304***
	(20.168)	(20.046)	(19.964)
Peer Ret	**-0.073*****	**-0.069*****	**-0.062****
	(-2.858)	**(-2.645)**	**(-2.299)**
Peer Ret (*SIC*)	-0.036	-	-0.024
	(-1.549)	-	(-0.960)
Peer Ret (*GICS*)	-	-0.042	-0.030
	-	(-1.626)	(-1.061)
Δ*Size*	-0.054*	-0.054*	-0.054*
	(-1.927)	(-1.927)	(-1.930)
Δ*BM*	-0.125***	-0.124***	-0.124***
	(-6.886)	(-6.861)	(-6.863)
SalesGrowth	0.327***	0.327***	0.326***
	(10.687)	(10.672)	(10.656)
Vol	-0.266***	-0.265***	-0.264***
	(-3.043)	(-3.034)	(-3.024)
Tenure	0.017***	0.017***	0.017***
	(3.575)	(3.558)	(3.564)
Age	-0.055**	-0.056**	-0.056**
	(-2.276)	(-2.314)	(-2.294)
Own	-0.039	-0.037	-0.038
	(-0.614)	(-0.582)	(-0.596)
Constant	0.330***	0.334***	0.333***
	(3.350)	(3.389)	(3.375)
Year FE	Yes	Yes	Yes
# of observations	21,015	21,015	21,015
Adjusted R-squared	0.065	0.065	0.065

Table VII
Dynamic peer groups

This table presents the results obtained from the regression of the change in the natural logarithm of total CEO compensation on own firm stock returns (*Firm Ret*), peer firm average returns in period *t* based on past, new, current, or future peers, control variables, and industry-year fixed effect over the sample period between 1996 and 2013. We define *Past Peers* in column (1) as firms that were firm *i*'s product market peers in period *t-1* but are not in the same TNIC group in period *t*. *New Peers* in column (2) are firm *i*'s product market peers in period *t* but were not in the same TNIC group in period *t-1*. *Current Peers* in column (3) are firm *i*'s product market peers in period *t-1* as well as in period *t*. Lastly, we define *Future Peers* in column (4) as firms that will be firm *i*'s product market peers in period *t+1* but are not in the same TNIC group in period *t*. We calculate equal-weighted average stock returns of each set of peers in period t and replace these the *Peer Ret* variable in equation (1) with each of these calculated peer firm average returns. All continuous variables are winsorized at 1% and 99% levels and are defined in the Appendix. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

	Dependent Variable: Δ*ln*(*Total Comp*)			
Independent Variables	(1)	(2)	(3)	(4)
Firm Ret	0.305***	0.297***	0.303***	0.312***
	(11.819)	(12.223)	(13.957)	(10.859)
Peer Ret_Past	**-0.053**	-	-	-
	(-1.169)	-	-	-
Peer Ret_New	-	**-0.124*****	-	-
	-	**(-2.740)**	-	-
Peer Ret_Current	-	-	**-0.163*****	-
	-	-	**(-3.455)**	-
Peer Ret_Future	-	-	-	**-0.068**
	-	-	-	**(-1.537)**
Δ*Size*	0.027	0.010	-0.007	-0.038
	(0.513)	(0.202)	(-0.162)	(-0.672)
Δ*BM*	-0.133***	-0.158***	-0.094***	-0.137***
	(-3.482)	(-4.607)	(-3.135)	(-3.478)
SalesGrowth	0.187***	0.252***	0.281***	0.264***
	(3.403)	(4.617)	(6.008)	(4.470)
Vol	-0.111	-0.097	-0.014	-0.136
	(-0.484)	(-0.454)	(-0.084)	(-0.569)
Tenure	0.028**	0.027***	0.017**	0.035***
	(2.516)	(2.661)	(2.219)	(3.042)
Age	-0.178***	-0.105*	-0.041	-0.054
	(-3.005)	(-1.931)	(-1.017)	(-0.861)
Own	0.059	0.031	0.092	-0.076
	(0.304)	(0.192)	(0.898)	(-0.530)
Constant	0.680***	0.400*	0.154	0.200
	(2.916)	(1.869)	(0.980)	(0.816)
Industry-Year FE	Yes	Yes	Yes	Yes
# of observations	8,534	9,192	12,098	8,272
Adjusted R-squared	0.072	0.094	0.092	0.077

Table VIII
Cross-sectional variation: strategic flexibility and RPE

This table presents the results obtained from the regression of the change in the natural logarithm of total CEO compensation on own firm stock returns (*Firm Return*), peer firm average returns based on TNIC (*Peer Ret*), control variables, and industry-year fixed effect over the sample period between 1996 and 2013. In column (1) and (2), firm-year observations are divided into two groups based on the median value of market-to-book ratio. In column (3) and (4), observations are divided into two groups based peer-adjusted stock return in period *t-1*. In column (5) and (6), observations are divided into two groups based on the firm's asset growth rate in period *t+1*. All continuous variables are winsorized at 1% and 99% levels and are defined in the Appendix. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

	Dependent Variable: $\Delta ln(Total\ Comp)$					
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Market-to-Book		*Peer-Adjusted Return*		*Asset Growth*	
	Low	High	Negative	Positive	Negative	Positive
Firm Ret	0.296***	0.306***	0.271***	0.332***	0.340***	0.256***
	(13.423)	(11.911)	(11.494)	(13.741)	(15.620)	(10.163)
Peer Ret	**-0.179***	**-0.036**	**-0.177***	**-0.070***	**-0.196***	**-0.056**
	(-4.252)	**(-0.859)**	**(-3.976)**	**(-1.796)**	**(-4.634)**	**(-1.409)**
Δ*Size*	-0.087*	-0.002	-0.108**	-0.046	-0.094**	-0.024
	(-1.881)	(-0.044)	(-2.075)	(-1.077)	(-2.088)	(-0.542)
Δ*BM*	-0.081***	-0.173***	-0.062**	-0.128***	-0.084***	-0.098***
	(-3.546)	(-2.879)	(-2.015)	(-3.814)	(-2.994)	(-3.210)
SalesGrowth	0.228***	0.363***	0.266***	0.275***	0.238***	0.324***
	(4.936)	(6.197)	(5.030)	(5.981)	(5.105)	(5.845)
Vol	0.075	-0.396*	0.078	-0.207	-0.167	0.158
	(0.442)	(-1.920)	(0.376)	(-1.017)	(-0.953)	(0.809)
Tenure	0.017**	0.028***	0.033***	0.013	0.020**	0.026***
	(2.199)	(3.015)	(3.084)	(1.345)	(2.198)	(2.607)
Age	-0.075*	-0.069	-0.039	-0.086*	-0.007	-0.124**
	(-1.677)	(-1.403)	(-0.668)	(-1.699)	(-0.148)	(-2.262)
Own	0.038	0.011	-0.171	0.136	-0.091	0.079
	(0.366)	(0.095)	(-1.104)	(1.148)	(-0.859)	(0.639)
Constant	0.309*	0.265	0.118	0.370*	0.048	0.463**
	(1.755)	(1.350)	(0.516)	(1.852)	(0.247)	(2.143)
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
# of observations	11,014	11,014	10,093	11,935	11,901	10,127
Adjusted R-squared	0.112	0.084	0.095	0.097	0.095	0.110
Coefficient difference	ΔCoeff.	p-value	ΔCoeff.	p-value	ΔCoeff.	p-value
Firm Ret	0.010	(0.766)	0.061*	(0.070)	-0.084**	(0.012)
Peer Ret	0.143**	(0.016)	0.107*	(0.060)	0.141**	(0.014)