Are CEOs With a Ph.D. More Innovative?*

Zhaozhao He[a]

September, 2014

ABSTRACT

Prior empirical work finds no systematic relation between CEO education and firm performance, which leads to a puzzle as to why firms hire managers with higher educational attainment. This study investigates the link between corporate innovation and managerial talent, indicated by the highest degree achievable, a Ph.D. I find that firms managed by CEOs with Ph.D. degrees (Ph.D. CEOs) produce more patents with a higher number of citations and achieve greater innovation efficiency for a given amount of innovation spending. Further evidence supports the explanation that Ph.D. CEOs have superior innovation ability, and thus are better able to transform scientific knowledge into valuable intellectual properties. Moreover, the results are robust when I control for CEO innate talent and general skills accumulated during the work experience. Finally, risk-taking preferences or compensational incentives do not appear to explain the innovative behavior of Ph.D. CEOs.

Keywords: CEO Education, Innovation, Managerial Talent, Patents, Ph.D.

JEL Classification: G30, O31, J44

*The author thanks Christopher Anderson, Donna Ginther, Babajide Wintoki, and faculty member at the University of Kansas. All remaining errors are my own.

[a]University of Kansas. Zhaozhao He can be reached at phone: (785)-864-8065, fax: (785)-864-5328, email: zzhe@ku.edu, and mailing address: 1300 Sunnyside Avenue, Lawrence, KS 66045

I. INTRODUCTION

Recent literature in corporate finance suggests that managerial traits affect corporate policies (Bertrand and Schoar, 2003; Malmendier and Nagel, 2011; Malmendier, Tate, and Yan, 2011). For instance, overconfident managers tend to be more optimistic, accept greater risk, and achieve greater innovative success (Hirshleifer, Low, and Teoh, 2012). Although CEO education is one of the few readily observables that are often used to proxy for CEO ability, a number of studies find no systematic relation between CEO education and firm performance (e.g., Gottesman, and Morey, 2006; Bhagat, Bolton, and Subramanian, 2010), which generates a quantitative puzzle as to why firms hire CEOs with a better educational background.

This study focuses on the highest educational attainment, the possession of a Ph.D. degree, to investigate the link between managerial talent and education in the context of corporate innovation. The completion of a Ph.D. program requires substantial amount of effort and knowledge acquisition, and individuals with a Ph.D. could have unusual characteristics (Chaudhuri, Ivkovic, Pollet, and Trzcinka, 2014) that may be reflected in organizational outcomes (Hambrick and Mason, 1984). More importantly, to obtain a Ph.D. degree, one needs to complete advanced coursework and apply sophisticated methodologies to execute unique research ideas. As a result, individuals holding a Ph.D. degree are likely to possess superior analytical skills and problem-solving skills. The set of skills and the amount of scientific knowledge acquired during a doctoral program might enable a manager to be better at visioning promising projects, processing new information, and transferring into breakthrough ideas (Becker, 1964, 1993).

I therefore hypothesize that CEOs with a Ph.D. degree (Ph.D. CEOs) achieve greater innovation than those without a Ph.D., potentially due to the specialized human capital (i.e., cognitive complexity, creative thinking, and receptivity to innovation) accumulated during the rigorous program that enables them to transform scientific knowledge into valuable inventions. An empirical puzzle raised by existing research is that while CEO education influences firm hiring decisions, it is not associated with greater firm performance, leading to an argument that CEO education is a poor proxy for CEO ability (Bhagat, Bolton, and Subramanian, 2010). By focusing on innovation outcomes I propose a possible solution to the CEO education puzzle: CEOs with the highest degree attainable are successful innovators.

To shed light on whether Ph.D. CEOs enhance corporate innovation, I assemble a panel of Standard & Poor's (S&P) 1,500 firms over 1992-2004 period. Among 7,403 CEOs with available biographical data in BoardEx, a proprietary database provided by Management Diagnostic Limited, 587 CEOs hold Ph.D. degrees, suggesting an ideal setting to study the innovation performance of companies managed by CEOs with a Ph.D.[1] Following the innovation literature, I measure innovation productivity using several patent-based metrics (Hall, Jaffe, and Trajtenberg, 2001). The first proxy is the number of patents filed by a firm to the United States Patent and Trademark Office (USPTO) in a given year, which are eventually granted. The second innovation measure is the total number of non-self citation counts, as citations are more precise to assess the economic importance of an innovation (Trajtenberg, 1990). The third measure, the number of citations per patent, captures the quality of each innovation product.

I find that firms with Ph.D. CEOs are significantly more likely to apply for patents than firms without Ph.D. CEOs. More importantly, CEOs with a Ph.D. degree are associated with 24% more patent grants than those without. Furthermore, patents filed by firms having Ph.D. CEOs receive 46% more subsequent citation counts, implying a greater economic importance. To glean more insight on whether CEOs with a Ph.D. are able to achieve higher quality of each patent, I examine the relation between Ph.D. CEOs and the number of citations per patent. The results show that each patent applied by firms with Ph.D. CEOs receives 6% more future citations than that by other firms. Taken together, these results support the hypothesis that Ph.D. CEOs are not only able to enhance the overall innovation productivity for their companies but also able to achieve higher impact and social value of each invention.

I recognize that endogenous matching of Ph.D. CEOs to firms might bias the observed findings. For instance, firms with more innovative projects may seek managers with a doctorate degree, which presumably signals certain managerial talent. At the same time, Ph.D. CEOs may be better able to predict

[1] Management Diagnostic Limited is a private research company specializing in collecting company social network data including a CEO's past or current employers, business relationships, organization affiliations, boards served, past universities attended, and degrees achieved.

which firms will have superior innovation performance in the future, and thus select employers accordingly. In these cases, finding a positive relation between Ph.D. CEOs and firm innovation would be a spurious outcome. I therefore apply a propensity score matching algorithm to correct for any endogenous selection on observed factors and track the innovation patterns of firms hiring Ph.D. CEOs and similar firms having non-Ph.D. CEOs over the years after the CEO appointments. I observe that firms run by CEOs with a Ph.D. achieve an increased number of patents over five years after the appointments, whereas those hiring non-Ph.D. CEOs do not. This evidence suggests that the endogeneity of CEO selection is unlikely to drive the main findings. The results are also robust when I focus on a subsample for which CEO-firm endogenous matching is less likely to be prevalent (Hirshleifer, Low, and Teoh, 2012), and when I analyze firms founded by Ph.D. CEOs, thus mitigating the potential matching mechanisms (Chaudhuri, Ivkovic, Pollet, and Trzcinka, 2014).

I then explore the potential channels through which Ph.D. CEOs enhance corporate innovation. The underlying cause of these findings could stem from a Ph.D. CEO's innate talent that is time-invariant (Falato, Li, and Milbourn, 2012), general skills accumulated during the work experience (Custodio, Ferreira, and Matos, 2013), or superior innovation ability gained through advanced knowledge acquisition. I find evidence consistent with Ph.D. CEOs' outstanding innovation ability, inconsistent with their innate talent or general skills. Moreover, I show that Ph.D. CEOs increase the effectiveness of innovation for a given amount of R&D spending, and this phenomenon is stronger for managers with high innovation ability. I argue that if boards hire managers with a Ph.D. for their innovativeness, then Ph.D. CEOs should be more sensitive to poor innovation performance. Indeed, the results show that Ph.D. CEOs are more likely to be replaced if they produce fewer patents with poorer quality than non-Ph.D. CEOs.

In the final part of this paper, I consider two alternative explanations for the main findings. First, it is possible that CEOs with a Ph.D. might be optimistic and thus more willing to undertake risky projects (Hirshleifer, Low, and Teoh, 2012). However, I do not find evidence that Ph.D. CEOs are associated with higher firm risk, measured by either stock return volatility or idiosyncratic risk. These results suggest that being a better innovator is not simply equivalent to taking more risk. The second potential explanation is that Ph.D. CEOs may have greater incentives stimulated by their compensation benefits (Lerner and Wulf, 2007; Manso, 2011; Baranchuk, Kieschnick, and Moussawi, 2014). Yet, I find that CEOs with a Ph.D. are not paid at a premium over those without. Hence, these results do not support the interpretation that the innovative behavior of Ph.D. CEOs is due to a higher level of CEO pay.

This paper mainly makes two contributions. First, this study relates to the literature on managerial traits and firm innovation. While this topic is growing in the literature, this is the first paper to investigate the effect of the highest managerial talent on corporate innovative behavior. Most existing research has focused on the role of CEO compensation in motivating innovation (Lerner and Wulf, 2007; Baranchuk, Kieschnick, and Moussawi, 2014), the effect of corporate governance (Acharya and Subramanian, 2009; Aghion, Reenen, and Zingales, 2013; He and Tian, 2013; Tian and Wang, 2014), the role of firm characteristics (Hall and Ziedonis, 2001), CEO social networks (Faleye, Kovacs, and Venkateswaran, 2013), and CEO characteristics (Hirshleifer, Low, and Teoh, 2012; Bereskin and Hsu, 2013). The only closely related paper that examines CEO education and firm innovation is Barker and Mueller (2002), who do not find the amount of education of CEOs has any material impact on firm R&D investment. The current study primarily focuses on the highest education attainable of executives in the context of corporate innovation productivity and value creation. I identify that whether CEOs hold a doctorate degree is a strong determinant of corporate innovation.

Second, this paper contributes to the stand of literature on CEO education and firm performance (Gottesman, and Morey, 2006; Bhagat, Bolton, and Subramanian, 2010; Jalbert, Rao, and Jalbert, 2011). Most of the studies in this area find that CEO education level has no bearing on firm value and/or operating performance. However, in the context of institutional money management, Chaudhuri, Ivkovic, Pollet, and Trzcinka (2014) show that investment products managed by individuals with a Ph.D. achieve greater performance than matched products by non-Ph.D.s, after accounting for the management fee. The current study reveals another unexplored information embedded in the possession of a Ph.D. degree for corporate innovation. The findings suggest that CEOs with a doctorate degree are more creative and have higher innovative ability, which ultimately translates into successful innovation with greater economic significance and social value.

The reminder of the paper is organized as follows. Section II describes the data and variables. Section III presents empirical evidence on the relation between Ph.D. CEOs and innovation. Section IV addresses the endogenous matching between Ph.D. CEOs and firms. Section V explores the potential and alternative explanations for the main findings. Section VI describes additional robustness tests. Section VII concludes. Appendix provides information regarding the construction of variables.

II. DATA

A. Sample Construction

The sample examined in this paper is compiled from six different data sources. The base sample starts with all U.S. incorporated industrial firms covered in Standard and Poor's ExecuComp database from 1992 to 2004. Industrial firms are defined as companies with SIC codes outside the ranges 4900-4949 (utilities) and 6000-6999 (financials). ExecuComp provides information on CEOs, annual salary and bonus, and total compensation. Data on CEO characteristics, employment history, and educational background are obtained from BoardEx database. Patent and patent citation information comes from the National Bureau of Economic Research (NBER) patent database (Hall, Jaffe, and Trajtenberg, 2001). I then collect accounting data from Compustat North America Annual files, stock return data from Center for Research in Security Prices (CRSP) files, institutional holdings data from Thomson's CDA/Spectrum database (form 13 F), and governance data from RiskMetrics Governance database. Firms are required to have positive total assets and positive sales to be included in the sample. I further exclude firms that do not have securities assigned a CRSP security code of 10 or 11. After merging all the databases, imposing the sample screens, and deleting observations with missing values on controls, the final sample consists of 1,829 CEOs from 7,976 firm-year observations.

B. Measures of Innovation

Since the central question in this paper is whether CEOs with the highest educational level are better able to transform scientific knowledge into valuable intellectual properties, the analysis primarily focuses on firm innovation productivity. The main innovation variables are constructed using the innovation output data from the NBER patent database. Innovation activity is evaluated along three dimensions: innovation intensity, innovation importance, and the quality of each innovation.

The first measure of innovation is the number of patents filed by a firm to the United States Patent and Trademark Office (USPTO) in a given year that are eventually granted (Hirshleifer, Low, and Teoh, 2012; He and Tian, 2013), which also gauges the firm's overall innovation intensity. The second measure of innovation is the total number of non-self citation counts that are summed across all patents applied by the firm during the year. Citation counts are more precise to assess the economic importance of an innovation than simple patent counts (Trajtenberg, 1990). One potential concern is the time truncation bias of the citations because patents granted during the later years of the sample period have less time to accumulate citations. As such, citation count of each patent is multiplied by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005) to adjust this bias. So the total number of weighted citation counts is the sum of the weighted, non-self citations across all patents filed by a firm during the year. The third measure, the number of citations per patent, captures the quality of each innovation product.

Following the innovation literature, I code patent and citation counts to zero for firm-years without patent-related information in the NBER database. The sample period ends at 2004, two years before the last year in the database, because there is a 2-year lag between patent application and patent grant (Hall, Jaffe, and Trajetenberg, 2001). To further mitigate the potential time truncation issues, all the regressions incorporate year fixed effects (Hirshleifer, Low, and Teoh, 2012).

C. Ph.D. CEO

To identify CEOs with a Ph.D. degree, I use the director education file in BoardEx, which contains information on CEOs' graduating schools, degree types (i.e., undergraduate, Master, or Ph.D.), and graduation years. For each CEO, I also collect available information on the employment history, including dates of employment and the role title. In almost all of the cases, Ph.D. degrees are obtained during the years predate the current employment years. To be conservative, I drop the observations that graduation years of doctorate are after the current employment years.[2] Out of the 1,829 distinct CEOs in the sample, 141 CEOs hold Ph.D. degrees; these CEOs are thus classified as Ph.D. CEOs, and all others are non-Ph.D. CEOs. The fraction of Ph.D. CEOs in the sample is similar to those reported in Cohen, Frazzini, and Malloy (2008). I also gather the information on the major of each degree and classify all the fields of study into five categories: Engineering and Science, Business, Medicine, Law and others.

D. Control Variables

Following prior work in the innovation literature, I control for a number of firm and CEO characteristics that are likely to affect firms' innovative behavior. Regarding CEO traits, older CEOs tend to be more conservative, and younger CEOs tend to be more aggressive in corporate growth strategies (Faleye, Kovacs, and Venkateswaran, 2013; Serfling, 2014). In a similar vein, longer-tenured CEOs might have little incentives to innovate because the payoff from an innovation might not be realized before their retirement (Barker and Mueller, 2002). Therefore, *CEO age* and *tenure* are included in the regression analysis to account for these possibilities. Further, I control for *CEO delta* and *CEO vega* of the option holdings to take into account the relation between CEO incentives and corporate innovation (Coles, Daniel, and Naveen, 2006; Lerner and Wulf, 2006; Francis, Hasan, and Sharma, 2011; Manso, 2011). *CEO delta* is defined as the dollar increase in the CEO's wealth for a 1% increase in stock price, a measure of CEO's incentives to increase stock price and thus maximize shareholder wealth. *CEO vega* is the dollar increase in the CEO's option holdings for a 1% increase in stock return volatility, a proxy for the risk-taking incentives. *CEO delta* and *CEO vega* are computed using the one-year approximation

[2] There are only two CEOs in this situation. The first one is Philip M. Condit, who received a Ph.D. in engineering from the Tokyo University of Science in 1997 and was the CEO of the Boeing Company from 1996 to 2003. The second one is Richard C. Adkerson, who is the president, CEO and Vice Chairman of the Freeport-McMoRan Copper & Gold Inc. since 2008, received an Honorary Doctor of Science degree from Mississippi State University in 2010.

method of Core and Guay (2002). I also incorporate *CEO stock ownership* in the analysis because CEOs with greater at-risk wealth in firms tend to be more long-term oriented, and therefore more willing to develop innovative projects (Barker and Mueller, 2002). Lastly, I include a dummy variable for CEOs with an MBA degree.

With respect to firm-level controls, firm *Size* is calculated as the natural logarithm of book value of assets expressed in 2006 dollars in millions, since large firms are more likely to innovate due to the benefits of economic scale and scope (Hall and Ziedonis, 2001). *Age* is the number of years the firm is listed with a non-missing stock price on Compustat; recent studies show that young firms are more innovation intensive than mature firms (Brown, Fazzari, and Petersen, 2009). To capture the nonlinear effects of firm age on innovation productivity, I include a squared term of the natural logarithm of firm age. Innovative firms are usually associated with higher growth opportunities, lower leverage (Hall, 2005, 2009), higher cash holdings (Lyandres and Palazzo, 2014; He and Wintoki, 2014), and lower asset tangibility than less innovative firms. I thus include *Mktbk*, which is calculated as total assets minus book value of common equity plus the market value of common equity, divided by book value of assets, *Annualized stock return*, computed as the firm's daily stock return averaged over the fiscal year, *Leverage*, defined as the ratio of long-term debt plus debt in current liabilities to total assets, *Cash-to-assets*, the ratio of cash and marketable securities over book assets, and *Tangibility*, calculated as the net property, plant, and equipment over book assets. Moreover, past operating performance might influence a firm's innovation ability. To account for this, I control for sales growth and return on assets. Bushee (1998), and Aghion, Van Reenen, and Zingales (2009) show that institutional ownership affects managers' innovative incentives. So I incorporate the fraction of shares held by institutions over fiscal year, which is constructed as the arithmetic mean of the four quarterly institutional holdings. Detailed variable definitions are reported in the Appendix. To mitigate the effect of outliers on the results, all variables are winsorized at the 1^{st} and 99^{th} percentiles of their empirical distributions. Dollar values are CPI adjusted into 2006 dollars.

E. Descriptive Statistics

Panel A of Table I presents the summary statistics for all the variables used in the analysis for the full sample. The sample exhibits characteristics comparable to those in prior studies in the innovation literature (Hirshleifer, Low, and Teoh, 2012; Faleye, Kovacs, and Venkateswaran, 2013). First, the medians of innovation variables are all zero, suggesting that the distributions of patents or citation counts are right skewed. I take a number of approaches to mitigate the impact of the right skewness of these measures. First, I use the natural logarithm of patents, citations, and citations per patent as the dependent variables in the analysis. Second, I winsorize all of these variables at the 99th percentile to alleviate the impact of outliers on the findings. Third, I employ a Tobit model to estimate regressions so that firm-years with zero values for the dependent variables are left censored at zero. As shown, an average firm in the sample has 14 granted patents every year, and these patents receive a number of 76 non-self citations, or 194 non-self citations adjusted for the truncation bias. Moreover, each patent has been cited by 2.3, 5.8 times using raw citations or weighted citations, respectively.

[Insert Table I about here]

Turning into CEO characteristics, the average CEO is 48 years old when first became CEO. Most of the individuals have held one CEO position to date, and have fewer than two degrees. 26.9% of the executives in the sample also hold an MBA degree. Moreover, the average CEO is 55 years old and has been in office for 7.8 years. 13.3% of these individuals are founders or co-founders of their companies, 40% serve as board chair, and 98.5% are male. The means (medians) of total compensation package, salary and bonus, are $5.4 ($2.6) million and $1.5 ($1.1) million, respectively. And the means (medians) of CEO delta and vega are $1.0 ($0.3) million and $0.1 ($0.05) million, respectively. Lastly, the average (median) CEO owns 3.1% (0.4%) of the firm's shares.

Of greater interest are the descriptive statistics for subsamples classified into Ph.D. CEOs and non-Ph.D. CEOs, as shown in Panel B of Table I. Consistent with the hypothesis, firms managed by CEOs with Ph.D. degrees have significantly greater innovative productivity than other firms, as the Wilcoxon rank sum tests indicate that both means and medians of all the innovation measures between the two groups are significantly different at a 1% level. Specifically, the average firm with a Ph.D. CEO has 8 more patents, 25 more citations, and 116 more adjusted citations than an average firm without. In terms of innovation quality, each granted patent filed by firms with Ph.D. CEOs receives significantly greater average number of raw citations or weighted citations.

It is also worthwhile to compare the characteristics of managers with and without doctorate degrees. To begin with, Ph.D. CEOs take less time to first become CEOs than non-Ph.D. CEOs (46.8 years vs. 47.9 years). The median Ph.D. CEO is about two years younger when promoted to the top job than the median non-Ph.D. CEO. Intuitively, Ph.D. CEOs have significantly higher number of degrees than their counterparts (2.8 vs. 1.6). It is interesting to see that executives with the highest academic education level are significantly less likely to hold an MBA degree than all others. Ph.D. CEOs are on average younger, have longer tenure, and are more likely to be founders. As to compensation and salary, an average Ph.D. CEO earns $2.4 million more in total compensation, but slightly less in salary and bonus, than a chief executive without a Ph.D. Finally, Ph.D. CEOs have higher mean (median) delta, higher mean (median) vega, and higher average (median) stock ownership than non-Ph.D. CEOs.

Looking at the firm-level controls, we see that Ph.D. CEOs tend to work in substantially larger firms with fewer tangible assets, higher market-to-book ratios, higher sales growth, but lower return on assets, than CEOs without a Ph.D. In terms of capital structure, firms with Ph.D. CEOs use less debt but more cash as financing. There is little difference in the institutional holdings between the two groups of subsamples. Lastly, CEOs with doctorate degrees are more likely to work in younger firms with higher stock return volatility, higher idiosyncratic risk, and higher stock return, than their counterparts.

Table II gives the descriptive statistics for the distribution of Ph.D. CEOs in the sample. Panel A displays firm-years with and without Ph.D. CEOs across the Fama French 12 Industry classification. As seen, Ph.D. CEOs are more prevalent in business equipment (37.17%) and health sectors (23.16%), suggesting that Ph.D. CEOs are more likely to work in innovative industries. Panel B shows the presence of Ph.D. CEOs relative to non-Ph.D. CEOs by R&D intensity measured by R&D-to-assets ratio. I find a monotonic increase in the presence of Ph.D. CEOs with the R&D intensity. Specifically, the fraction of

Ph.D. CEO observations is only 4.5% of firm-years with zero R&D; this fraction increases to 16.11% among firms in the top quartile of R&D intensity. Panel C presents the distribution of Ph.D. degrees awarded by the top U.S. institutions.[3] The six elite universities that graduate the most Ph.D. CEOs – Stanford University, Harvard University, Columbia University, New York University, University of California Berkeley and Princeton – jointly account for 20% of total Ph.D. degrees in the entire sample.

[Insert Table II about here]

III. PH.D. CEOS AND INNOVATION

To investigate whether CEOs with Ph.D. degrees are better innovators, I estimate the following regression taking the form:

$$Innovation_{i,t+1} = \propto \; + \beta_0 * Ph.D.\ CEO_{i,t} \; + \; \beta_1 * X_{i,t} + \; \beta_2 * Y_{i,t} + \epsilon_{i,t} \; + \gamma_j \; + \; d_t \qquad (1)$$

where *i* indexes firm and *t* indexes year. The dependent variables are the innovation measures discussed in section II.B: *log (1+patents), log (1+citations),* and *log (1+citations per patent).* In most of the cases, each regression is estimated with a Tobit model. Since these measures are forwarded by one year, I exclude firms that have new CEOs in the next year. The key variable of interest is *Ph.D. CEO*, a dummy variable equal to one if the CEO holds doctorate degrees as identified in the BoardEx, and zero otherwise. *X* is a vector of firm-level controls including *Size, log (age),* $[\log(age)]^2$, *Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, and Institutional ownership. Y* is a vector of variables that control for CEO characteristics, containing *log (1+CEO tenure), log (1+ CEO age), CEO MBA degree, log (1+ CEO delta), log (1+ CEO vega),* and *CEO stock ownership.* Equation (1) includes industry (γ_j) and year fixed effects (d_t), where industries are defined at the two-digit SIC code level. Standard errors are clustered by firm (Petersen, 2009).

[3] Notice that the total number of doctorate degrees is greater than the total number of Ph.D. CEOs because some CEOs have two to three doctorate degrees.

A. Patenting Activity

To begin with, I first analyze whether firms with Ph.D. CEOs are more likely to apply for patents than those with non-Ph.D. CEOs by estimating a logit model with the dependent variable equal to one if a firm applies for patents during the year, and zero otherwise. The control variables are those specified in Equation (1). The results, as reported in column (1) of Table III, show that Ph.D. CEOs are associated with a significantly higher likelihood of applying for patents than non-Ph.D. CEOs. Thus, as the starting point of the patenting activity, the propensity of patent application is much greater for firms run by CEOs with Ph.D. degrees than other firms.

[Insert Table III about here]

Next, I examine the main hypothesis that Ph.D. CEOs are better able to transform scientific knowledge into innovative products, and thus their firms display greater innovation intensity proxied by the patent counts. The dependent variable is the natural logarithm of one plus the number of patents filed by a firm in the next year. Equation (1) is estimated with a Tobit model. The results are reported in column (2) of Table III.

As predicted, the result shows that Ph.D. CEOs are associated with significantly higher patent counts. The estimated coefficient of *Ph.D. CEO dummy* is 0.529 and significant at a 1% level. To better understand the economic magnitude, I estimate the marginal effect of each variable.[4] Since the estimated marginal effect of *Ph.D. CEO dummy* is 0.411 conditional on the uncensored observations, this implies that among the firms with non-zero patenting activity, companies with a Ph.D. CEO apply for and are eventually granted 24% ($e^{0.529*0.411} - 1$) more patents compared with those without a Ph.D. CEO, when other covariates are evaluated at their sample means. The unconditional marginal effect of Ph.D. CEO is 0.172, which translates into a 9% ($e^{0.529*0.172} - 1$) higher patent count for firms executed by a Ph.D. CEO relative to all other firms in the entire sample, confirming the notion that Ph.D. CEOs are able to achieve higher patent grants than non-Ph.D. CEOs.

[4] The regression coefficients of variables in a Tobit model cannot be interpreted as the marginal effects of the variables.

Among other variables, we see that a firm's patenting activity is significantly and positively associated with firm size, market-to-book ratio, cash holdings, institutional ownership, and CEO vega. The number of patents is negatively correlated with firm leverage. These results are consistent with those documented in prior studies (Francis, Hasan, and Sharma, 2011; Hirshleifer, Low, and Teoh, 2012; Aghion, Reenen, and Zingales, 2013; Faleye, Kovacs, and Venkateswaran, 2013). Lastly, there is no evidence on whether CEOs hold an MBA degree relates to firm innovation.

B. Patent Citations

The results above suggest that companies run by executives with doctorate degrees exhibit greater innovation intensity compared to those without Ph.D. degrees. As the simple number of patents is an imperfect measure of innovation success, I then turn to the number of citations that a firm has received on its patents to examine whether innovations produced by firms with Ph.D. CEOs also generate greater economic importance. The dependent variable is the one-year forward natural logarithm of one plus the number of non-self citations, adjusted or unadjusted for the truncation bias. The results estimated using Equation (1) are reported in columns (3) – (4) of Table III.

Using raw citation counts, column (3) show that Ph.D. CEOs are able to bring 46% ($e^{0.817*0.47} - 1$) more subsequent citations to the patents applied for, conditional on being cited, and 12% ($e^{0.817*0.13} - 1$) more citations compared with all non-Ph.D. CEOs in the sample.[5] After taking into account the time truncation bias, the coefficient on *Ph.D. CEO dummy* increase in magnitude with a larger t-statistic, as reported in column (4); the effect also become stronger. Precisely, Ph.D. CEOs increase the weighted citation counts by 97% (16%), conditional (unconditional) on being cited.[6] Taken together, these results have provided supporting evidence that firms with Ph.D. CEOs produce a higher number of patent grants, which also create a greater overall economic impact, than those with non-Ph.D. CEOs.

[5] The conditional and unconditional marginal effects of *Ph.D. CEO dummy* are 0.47 and 0.13, respectively.
[6] The conditional and unconditional marginal effects of *Ph.D. CEO dummy* are 0.64 and 0.14, respectively.

C. Citations per Patent

While the number of patents and citations measure the overall innovation productivity of a firm, the number of citations per patent evaluates the quality of each innovation output. To glean more insight on whether managers with doctorate degrees are able to achieve higher significance of each patent, I next estimate Equation (1) with the dependent variable using the natural logarithmic transformation of one plus citations per patent in the following year, for weighted and un-weighted citation counts. The results are displayed in columns (5) – (6) of Table III.

As shown, the coefficients of *Ph.D. CEO dummy* are significant at a 1% level in both columns. In terms of economic significance, each patent applied by firms with Ph.D. CEOs receives 6% (4%) more future citations than that by other firms, conditional (unconditional) on being cited. The effect is also stronger when using the weighted citation counts; the corresponding effect is 16% (7%). Consistent with previous findings, these results imply that Ph.D. CEOs are not only able to enhance the overall innovation productivity for their companies but also able to achieve higher impact and social value of each invention.

IV. ENDOGENOUS MATCHING BETWEEN PH.D. CEOS AND FIRMS

The evidence so far is consistent with Ph.D. CEOs enhancing both quantity and quality of firm innovation. Nevertheless, endogeneity in the assignment of CEOs to firms often plague empirical tests. Indeed, Table II shows that CEOs with doctorate degrees are concentrated in innovative industries and hired by firms that innovate more. This endogenous matching of CEOs to firms can be explained in two ways. First, firms with promising opportunities for innovative projects might seek managers with a doctorate degree, which presumably signals certain managerial talent.[7] Second, Ph.D. CEOs might be better able to predict which firms will have superior innovation performance in the future, and thus select

[7] I also construct a subsample of one year before and one year after CEO turnover cases to examine the effect of a former Ph.D. CEO on the firm's future hiring decisions. The results are reported in the Internet Appendix Table A.I. I find that firms do proactively select new CEOs with Ph.D. degrees if former CEOs have Ph.D.s, in line with the findings documented by Bhagat, Bolton, and Subramanian (2010).

employers accordingly. Under these circumstances, the positive relation between the presence of a Ph.D. CEO and firm innovation would be spurious.

A. Propensity Score Matching Algorithm

To tackle this selection issue, I conduct a propensity score matching analysis and compare the innovation patterns of firms that are managed by Ph.D. CEOs with similar firms that are not. Specifically, I estimate a Logit model of the *Ph.D. CEO dummy,* which equals one for firms with Ph.D. CEOs and zero otherwise, to generate the likelihood (i.e., propensity score) that a firm appoints a Ph.D. CEO using observable firm and CEO characteristics including *log (1+patents), Size, Tangibility, Mktbk, ROA, Cash-to-assets, Annualized stock return, log (1+CEO tenure), log (1+ CEO age)* and *CEO founder*, as well as industry and year dummies.[8]

Panel A of Table IV presents the parameter estimates of the Logit model. The results show that six out of ten of the covariates significantly predict the propensity of having a Ph.D. CEO run the company. Moreover, the p-value from the χ^2 test is less than 0.0001, confirming the model fitness. Using the predicted probabilities from the model, I then perform a propensity score match with replacement and with a radius of 1%.[9] Specifically, each firm with a Ph.D. CEO is matched to firms hiring non-Ph.D. CEOs at the time of appointment with the propensity scores within ±1% of each other. This procedure generates 38 treated firms (firms appointing Ph.D. CEOs) and 133 matched firms. I then track the innovation patterns of the treatment (156 firm-years) and control groups (525 firm-years) during the post appointment period.

[Insert Table IV about here]

To get a sense of how well the matching algorithm works, Panel B of Table IV shows balance tests of all the control variables at the time of hiring for the two groups. Overall, the matching process has

[8] To account for the possibility that Ph.D. CEOs are more likely to be entrepreneurs, who are relatively more optimistic and thus innovate more (Puri and Robinson, 2009), I include *CEO founder*, a dummy equal to one if the CEO is identified as "founder" or "co-founder" of the company.
[9] The results hold when I use different cutoffs (e.g., using 2.5% or 5% as the radius for propensity score matching).

removed most of the observed differences in the two subsets of firms when making the hiring decisions, except that firms appointing Ph.D. CEOs are slightly larger and hold less cash than the counterparts. Importantly, there are no significant differences in the innovation productivity measured by *log (1+patents)* between the treated and control firms in the years of hiring, which should facilitate meaningful comparison of the innovation outcomes during the CEO employment.

Figure I tracks the innovation patterns, measured by the number of patents, of treated and matched firms since the years of CEO appointments. In stark contrast, firms that hire Ph.D. CEOs increase the number of patent applications from 6 to 12 on average over the five-year period, whereas comparable firms exhibit a slightly downward trend of patenting activities.

[Insert Figure I about here]

Panels C and D of Table IV present the univariate test results and multivariate regression analysis. In each panel, the outcome variables are the natural logarithm of one plus the number of patents, the natural logarithm of one plus the number of weighted citations, and the natural logarithm of one plus the number of weighted citations per patent, respectively. Panel C displays the average differences in the innovation measures between treatments and controls. Mirroring the above results, firms run by Ph.D. CEOs have 2.1 ($e^{0.77}$) times more patents, 2.6 ($e^{0.96}$) times more weighted citations, and 1.5 ($e^{0.38}$) times more weighted citations per patent than comparable firms by non-Ph.D. CEOs. All the differences are significant at a 1% level.

Panel D shows the OLS regression estimates using the matched samples. To save space, only the coefficient estimate and t-statistic associated with the *Ph.D. CEO dummy* is reported for each regression, while controlling for all the matching variable as well as industry and year fixed effects. In line with the findings in the univariate tests, the positive and significant coefficient on *Ph.D. CEO* for all innovation proxies confirms that Ph.D. CEOs enhance corporate innovation. Overall, these results provide convincing evidence that the endogeneity of CEO selection is unlikely to drive the main findings.

B. Subsamples Addressing the CEO-firm Endogenous Matching

To provide additional evidence, I focus on two subsamples. The first subsample contains firm-years for which CEO-firm endogenous matching is less likely to be prevalent. To do so, I restrict the sample to firm-years in which CEOs have been employed at their firms for at least three years. The idea is that since CEO education is a persistent trait, but firm characteristics are time-varying, the matching effects between certain CEO traits and a firm's opportunity sets tend to be strongest when the CEO is newly appointed (Hirshleifer, Low, and Teoh, 2012). Another rationale is that this restriction helps to identify the true management style by allowing managers to "imprint their mark" on a given company (Bertrand and Schoar, 2003). The results of estimating Equation (1) are presented in Panel A of Table V. We continue to see a significant and positive relation between the presence of Ph.D. CEOs and the innovative behavior of their firms.

[Insert Table V about here]

As an alternative approach, I turn to the analysis of firms founded by Ph.D. CEOs, thus eliminating the potential matching mechanisms elaborated earlier (Chaudhuri, Ivkovic, Pollet, and Trzcinka, 2014). As Panel B of Table V shows, I again find that companies run by Ph.D. CEOs exhibit significantly greater innovation activities than other firms. For all innovation measures, the coefficient on *Ph.D. CEO dummy* is statistically significant at a 1% level, despite the small sample size of 1,103 observations. Taken together, these results constitute strong evidence that the differences in innovation performance do not seem to stem from the endogenous matching between Ph.D. CEOs and innovative firms.

V. POTENTIAL EXPLANATIONS

Having established the fact that hiring a Ph.D. CEO leads to an increase in a firm's innovation productivity, this section aims to explore the potential mechanisms through which Ph.D. CEOs enhance corporate innovation. The underlying cause of these findings could stem from a Ph.D. CEO's innate talent, general skills accumulated during the work experience, or superior innovation ability gained through advanced knowledge acquisition. The following analysis attempts to pin down which factor is truly at play.

A. Do Ph.D. CEOs Have Higher Innate Talent, General Skills, or Innovation Ability?

It is conceivable that talented managers are more innovative. To proxy for CEO talent, I create a *High latent ability* dummy, which equals one if the age at which the individual took the first CEO position is in the bottom decile among all the CEOs in the sample, and zero otherwise. This is based on the idea that CEOs who have secured the first executive-level positions earlier in their careers exhibit greater talent (Falato, Li, and Milbourn, 2011; Faleye, Kovacs, and Venkateswaran, 2013). If the possession of Ph.D.s is simply a reflection of unobserved, time-invariant talent, then the effect of Ph.D. degrees should be subsumed when controlling for such managerial talent. Columns (1), (4), and (7) of Panel A in Table VI show results of estimating Equation (1) with the proxy of innate ability included. We see that this is not the case. For all innovation measures, the coefficient of *Ph.D. CEO dummy* remains quantitatively and qualitatively similar to the corresponding one in the baseline analysis, while the *High latent ability* dummy is not statistically significant across all specifications.

[Insert Table VI about here]

I next turn into the possibility that the greater innovation performance of Ph.D. CEOs is attributable to their general managerial skills gathered over their lifetime work experience. Indeed, Custodio, Ferreira, and Matos (2014) show that generalist CEOs innovate more than specialist CEOs because general skills are transferable across firms should innovative projects fail. Accordingly, the *General ability index* is the first factor of the principle components analysis of five variables capturing a CEO's generic skills: 1) the number of different positions in the past, 2) the number of firms where the CEO worked, 3) the number of industries where the CEO worked, 4) a dummy variable indicating whether the CEO held a CEO position before, and 5) a dummy indicating whether the CEO worked for a conglomerate firm (Custodio, Ferreira, and Matos, 2013, 2014). Columns (2), (5), and (8) of Panel A in Table VI exhibit the results of estimating Equation (1) after accounting for general managerial skills. The coefficient of *General ability* has a positive sign but is statistically insignificant across all specifications. Rather, the coefficient of *Ph.D. CEO dummy* is comparable to the corresponding one in Table III, and is still associated with statistical significance at the 1% level, indicating that the positive impact of CEOs' doctorate degrees on firms' innovative outputs found earlier is not driven by CEO general skills.

I now examine the most credible explanation that Ph.D. CEOs possess superior innovation ability that is developed over the period of knowledge acquisition. Motivated by Faleye, Kovacs, and Venkateswaran (2013), I define a CEO with *High innovation ability* if the number of patents filed by the firm is greater than the average number of patents filed by firms in the same industry during the year. I then carry out a Tobit regression of Equation (1) with the inclusion of *High innovation ability* dummy. The results are presented in columns (3), (6), and (9) of Panel A in Table VI. As expected, CEOs with high innovation ability are associated with substantially enhanced innovation outcome (as found by Faleye, Kovacs, and Venkateswaran, 2013). Strikingly, we see that the coefficient of *Ph.D. CEO dummy* becomes marginally significant for all innovation measures once the *High innovation ability* is incorporated. These results are consistent with the notion that Ph.D. CEOs have better innovation ability to transform advanced knowledge into valuable intellectual properties.

Further, I expect the effect of Ph.D. on innovation activities to be more pronounced among CEOs with high innovation ability than those with low ability to innovate. I thus split the sample into high versus low innovation ability of CEOs. The results are displayed in Panel B of Table VI. As would be predicted, the coefficient on *Ph.D. CEO* is bigger and more significant for managers with high innovation ability across all specifications than the corresponding coefficient for managers with low innovation ability. These results confirm the interpretation that it is the innovation ability acquired by Ph.D. CEOs that causes substantial opportunities for innovation.

A.I. Evidence from Innovation Efficiency

To provide further evidence supporting the innovation ability of CEOs with Ph.D.s, I now investigate whether Ph.D. CEOs increases the effectiveness of innovation for a given amount of R&D

spending, and whether this effect is still stronger for managers with high innovation ability. Table VII reports the results of estimating Equation (1) with the addition of natural logarithm of one plus R&D-to-assets ratio for the full sample (Panel A), and for subsamples split by CEO innovation ability (Panel B).

[Insert Table VII about here]

Panel A shows that using the full sample, the coefficients on *Ph.D. CEO dummy* are comparable to those reported in Table III, in terms of both magnitude and statistical significance. These results suggest that CEOs with doctorate degrees are better able to utilize R&D investment to produce more successful inventions, which ultimately contributes to the increased overall innovation productivity, than CEOs without a Ph.D.

In the subsample analysis in Panel B, we see that the coefficients on *Ph.D. CEO dummy* for managers with high innovation ability are all positive and significant (although smaller than the corresponding ones in Panel A), whereas none of the coefficients on *Ph.D. CEO dummy* are significant for executives with low ability to innovate. This evidence, again, supports the argument that Ph.D. CEOs, especially those possess superior innovation ability, achieve greater innovation outcome for a given level of innovation spending, than non-Ph.D. CEOs.

A.II. Evidence from CEO Turnover

The findings thus far point out a potential solution to the CEO education puzzle (Bhagat, Bolton, and Subramanian, 2010), that CEOs with the most prestigious educational level are able to translate scientific knowledge into valuable innovations, which could explain why boards hire managers with a Ph.D. If this explanation is valid, then Ph.D. CEOs might be more sensitive to poor innovation performance. To examine this possibility, I estimate a logit model of CEO turnover on Ph.D. CEO taking the following form:

$$CEO\ Turnover_{i,t} = \propto\ + \beta_0 * Ph.D.\ CEO_{i,t}\ +\ \beta_1 * Ph.D.\ CEO_{i,t} * Z_{i,t} +\ \beta_2 * X_{i,t} +\ \beta_3 * Y_{i,t} + \epsilon_{i,t}\ +\ \gamma_j\ +\ d_t \qquad (2)$$

where *CEO turnover* is a binary variable that equals one for CEO turnover years. *Ph.D. CEO* is a dummy equal to one if the CEO holds doctorate degrees, and zero otherwise. Z is a vector of innovation performance including *log (1+patents), log (1+weighted citations),* and *log (1+weighted citations per patent).* The explanatory variables of interest are the interaction terms of $Ph.D.\ CEO$ and Z, which capture the differential CEO turnover-innovation performance sensitivity for CEOs with and without Ph.D.s. X and Y represent the whole set of firm-level and manager-level controls as specified in Equation (1). The model also includes industry (γ_j) and year (d_t) fixed effects. Table VIII exhibits the findings.

[Insert Table VIII about here]

Column (1) shows that, in general, CEOs with a Ph.D. are significantly less likely to be fired than CEOs without a Ph.D. This result contrasts sharply to the positive relation between CEO general ability and turnover probability (also reproduced in the Internet Appendix) documented by Custodio, Ferreira, and Matos (2014), providing another piece of evidence that the possession of a doctorate degree reflects a different dimension of ability that is not generic. Columns (2) – (4) include the interaction terms of *Ph.D. CEO dummy* and innovation performance. The significant and negative coefficient on the interaction term for all innovation proxies implies that Ph.D. CEOs are more likely to be replaced if they produce fewer patents with poor quality, than non-Ph.D. CEOs. Moreover, the F-statistic testing the joint significance of the interaction term and the *Ph.D. CEO dummy* is significant at a 1% level for all three innovation measures, confirming the negative CEO turnover-innovation performance sensitivity for Ph.D. CEOs.

In addition, I also examine the CEO turnover-operating performance sensitivity by estimating a similar model specified in Equation (2). I measure operating performance using ROA and industry-demeaned ROA. The results are reported in the Internet Appendix Table A.II. I do not find any significant difference in the sensitivity of CEO turnover to firm operating performance for Ph.D. CEOs and non-Ph.D. CEOs.

B. Alternative explanations

I consider two alternative explanations for the main findings. First, instead of having superior innovation ability, CEOs with a Ph.D. may be optimistic and thus more willing to undertake risky projects. Second, Ph.D. CEOs are more innovative because they have greater incentives stimulated by their compensation benefits.

B.I. Do Ph.D. CEOs take more risk?

To answer the question as to whether Ph.D. CEOs have a higher preference of risk-taking, I estimate an OLS regression of firm risk on *Ph.D. CEO dummy* and control variables based on Hirshleifer, Low, and Teoh (2012), and Serfling (2014). Following these studies, I measure firm risk using stock return volatility and idiosyncratic risk volatility. *Stock return volatility* is defined as the standard deviation of a firm's daily stock returns averaged over the fiscal year. *Idiosyncratic risk* is computed as the standard deviation of the error terms estimated from regressing daily stock returns on Fama and French three factors (Fama and French, 1993), and then averaged over the fiscal year. All regressions incorporate industry and year fixed effects. Table IX reports the results of this exercise.

[Insert Table IX about here]

Interestingly, there is no evidence that Ph.D. CEOs are associated with higher firm risk, measured by either stock return volatility (columns 1 and 2), or idiosyncratic risk (columns 3 and 4), with or without controlling for managerial incentives. In all the tests, the coefficient on *Ph.D. CEO dummy* is economically small and statistically insignificant. Therefore, it is also unlikely that the effect of Ph.D. CEOs on corporate innovation derives from CEO overconfidence (Hirshleifer, Low, and Teoh, 2012). Put differently, these results imply that being a better innovator is not simply equivalent to taking more risk.

B.II. Are Ph.D. CEOs paid more?

Another possible explanation for the innovative behavior of Ph.D. CEOs is that they receive greater compensational incentives. I thus investigate whether Ph.D. CEOs earn a wage premium over non-Ph.D. CEOs. To do so, I carry out an OLS regression of total compensation on *Ph.D. CEO dummy*, a similar set of controls according to Engelberg, Gao, and Parsons (2012), as well as industry and year fixed effects. The dependent variable is the natural logarithmic transformation of one plus total compensation converted into 2006 dollars. The results are shown in Table X.

[Insert Table X about here]

In column (1), the coefficient on *Ph.D. CEO* is positive but is not statistically significant, indicating CEOs with Ph.D. degrees are not paid at a premium over those without. Further, in columns (2) – (4), there is no consistent evidence that the relation between Ph.D. and CEO pay varies with the quality of corporate governance, as the interaction terms of governance proxies (G-index, institutional ownership, and CEO chairman) and *Ph.D. CEO dummy* are not always significant. Hence, these results do not support the interpretation that the innovative behavior of Ph.D. CEOs is due to a higher level of CEO pay.

As an extension, I examine whether Ph.D. CEOs are able to enhance firm value and operating performance. The dependent variable of the OLS regression is either Tobin's Q or industry-adjusted ROA. I estimate the regression model on all firms, and on subsamples split by CEO innovation ability. The results, as reported in the Internet Appendix Table A.III, show no evidence that Ph.D. CEOs are significantly associated with higher firm value or operating performance, consistent with the findings in Gottesman and Morey (2006), and Bhagat, Bolton, and Subramanian (2010). This evidence suggests that higher innovative output achieved by Ph.D. CEOs does not necessarily translate into greater firm performance. An alternative interpretation is that if there is an optimal matching between Ph.D. CEOs and innovative firms, then there will be no cross-sectional differences in firm performance conditional on CEO type (Custodio, Ferreira, and Matos, 2014).

VI. ROBUSTNESS CHECKS

Finally, I perform a number of robustness tests to validate the main findings. First, I exclude high-tech industries where innovation is deemed important to mitigate the concern that the results might be driven by CEOs in innovation intensive industries. Second, I examine the period leading up to the technology boom to alleviate the concern that the results are contaminated by the exceptional period for innovation intensive firms. The results are reported in Panels A and B of Table XI.

[Insert Table XI about here]

The results continue to show a positive relation between Ph.D. CEOs and innovation productivity after excluding the high-tech industries (Panel A), or restricting to the period prior to the incidence of technology bubble in 1998 (Panel B). All the coefficient estimates of *Ph.D. CEO dummy* remain significant at the conventional statistical levels, suggesting that the observed results are not driven by high-tech industries or the exceptional period of technology boom and bust in the 1990s.

Another concern is that the ability to innovate might only hold true for certain types of Ph.D.s, for example, CEOs with a degree in technical field (Baranchuk, Kieschnick, and Moussawi, 2014). I therefore create a binary variable equal to one if the CEO holds a doctorate degree in Engineering or Science field, and another binary variable indicating whether the CEO has a Ph.D. in Medicine. Panel C of Table XI shows that CEOs with a Ph.D. in Engineering or Science field are relatively more innovative than an average Ph.D. CEO, but those with a Ph.D. in Medicine are not. More importantly, the effects of Ph.D. CEOs on firms' innovative behavior are not diminished, supporting the generalization of the inference on CEOs with doctorate degrees and corporate innovation. Nevertheless, I control for CEO's technical background by incorporating dummy variables indicating whether the CEO has ever obtained a degree in Engineering and Science, or Medicine at any academic levels. Panel D presents the results. For each innovation proxy, the coefficient on *Ph.D. CEO dummy* is positive and statistically significant at a 10% level, implying that Ph.D. CEOs specializing in non-technical field also exhibit superior innovation performance than those non-Ph.D. CEOs without technical background.

I also adopt an instrumental variables approach to address the issue of unobserved heterogeneity. The validity of this method relies on the requirement that the selected instruments are correlated with CEOs holding doctorate degrees but uncorrelated with their firms' innovative behavior. I employ thee instruments. The first instrument is the number of degrees the CEO has obtained. While it is conceivable that Ph.D.s have received more degrees than non-Ph.D.s, there is little reason to believe that the number of degrees held by a CEO has a direct impact on the firm's innovation. The second instrument is a dummy indicating whether the CEO earned an MBA degree, as MBA programs have little to do with innovative skills, and getting an MBA is correlated with the decision to pursue a Ph.D.[10] The third one is the number of executive positions the CEO has held to date, as it's not clear a priori that a CEO has CEO experience before will innovate more or less than an individual first secures the executive-level positon. Table XII shows the results. Column (1) shows that all of the instruments are significantly associated with the *Ph.D. CEO dummy*. The high F-statistic of 36.49 and the Anderson canonical correlation statistic of 772.31 (p-value < 0.0001) indicate that the problem of weak instruments is not a concern. Columns (2) to (4) present the second-stage results. As seen, the effect of the presence of a Ph.D. CEO on corporate innovation remains positive and significant at a 1% level across all innovation measures. The over-identification test fails to reject the null that the instruments are valid at the 5% level.

[Insert Table XII about here]

Other robustness checks are presented in the Internet Appendix. I briefly describe them here. Instead of a Tobit model, I employ two alternative models – a negative binomial and an OLS regression.[11] In Table A.IV, the results estimated from the negative binomial show that the coefficient on *Ph.D. CEO dummy* is positive and significant at a 5% level or less for all innovation measures, consistent with the main findings.[12] The OLS regression estimates remain qualitatively and quantitatively similar.

Since the innovation process could take longer than one year, I rerun all the regressions using two-year forward dependent variables while excluding firm-years with a new CEO in the following two years. The results show comparable coefficient estimates of *Ph.D. CEO dummy* as in the baseline regressions.

[10] This is also applied in Faleye, Kovacs, and Venkateswaran (2013), who use a CEO's MBA degree as an instrument for the CEO's connections to examine corporate innovation.
[11] Recommended by Hall, Jaffe, and Trajtenberg (2001), the negative binomial model relaxes the assumption that the mean of the parameter equals the variance.
[12] Specifically, the estimated marginal effects based on standard errors using Delta method suggest that firms managed by Ph.D. CEOs receive 5.8 more patent grants, 80.5 more weighted citations, and 1.5 more weighted citations per patent.

To address the potential omitted variable bias that the possession of Ph.D. degrees might simply proxy for other factors that potentially affect firm innovative strategies, I incorporate additional CEO characteristics as controls. I first consider CEO bargaining power, since powerful CEOs might be more willing to take on innovative opportunities (Faleye, 2009). I measure CEO duality using an indicator equal to one if the CEO is labeled “chairman” in the BoardEx. Second, I include a dummy equal to one if the CEO is the founder of the company. I notice that only founder CEOs have a positive effect on the patent count, with no qualitative or quantitative change in the main result. In the unreported analysis, I also control for gender of the CEO because prior studies find that men tend to take more risks than females (Byrnes, Miller, and Schafer, 1999; Barber and Odean, 2001). The results are virtually unchanged after including this variable.

VII. CONCLUSION

This paper investigates the link between managerial talent and both ability and education in the context of corporate innovation. Focusing on the possession of a Ph.D. degree, this study finds that firms managed by Ph.D. CEOs achieve greater innovation performance than those by non-Ph.D. CEOs. Specifically, over the period of 1992-2004, CEOs with doctorate degrees are associated with higher number of patent applications and patent citations, better quality of each invention, and greater innovation efficiency for given amount of R&D investment, than non-Ph.D. CEOs. The results are robust to a battery of tests, including a propensity score matching method, subsample analyses, different estimation models, and an instrumental variables approach.

Further examination reveals that the underlying cause of the relation between Ph.D. CEOs and corporate innovation stems from their superior innovation ability. This explanation is also supported by the fact that the enhanced innovation performance is more pronounced for Ph.D. CEOs with greater innovation ability. These results provide convincing evidence that Ph.D. CEOs are better able to transform scientific knowledge into valuable intellectual properties. The findings in this paper therefore provide a clear justification for corporate boards hiring CEOs with Ph.D. degrees (Gottesman and Morey, 2006; Bhagat, Bolton, and Subramanian, 2010).

Finally, alternative interpretations such as innate talent, general skills, excessive risk-taking, or compensational incentives do not seem to explain Ph.D. CEOs' innovative behavior, suggesting that the possession of a Ph.D. captures a different facet of managerial ability, that is, the innovation ability.

REFERENCE

Acharya, V. V., & Subramanian, K. V. (2009). Bankruptcy codes and innovation. *Review of Financial Studies*, hhp019.

Aghion, P., Van Reenen, J., & Zingales, L. (2013). Innovation and Institutional Ownership. *The American Economic Review*, *103*(1), 277-304.

Baranchuk, N., Kieschnick, R., & Moussawi, R. (2014). Motivating innovation in newly public firms. *Journal of Financial Economics*, *111*(3), 578-588.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116, 261-292.

Barker III, V. L., & Mueller, G. C. (2002). CEO characteristics and firm R&D spending. *Management Science*, *48*(6), 782-801.

Becker, G. S. (1964), *Human capital* (Columbia University Press, New York, NY).

Becker, G. S. (1993), *Human capital: A theoretical and empirical analysis, with special reference to education, 3rd ed*. (University of Chicago Press, Chicago, IL).

Bereskin, F. L., & Hsu, P. H. (2013). Bringing in Changes: The Effect of New CEOs on Innovation. *Available at SSRN 1684329*.

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

Bhagat, S., Bolton, B., & Subramanian, A. (2010). CEO education, CEO turnover, and firm performance. *University of Colorado-Boulder working paper*.

Brown, J. R., Fazzari, S. M., & Petersen, B. C. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *The Journal of Finance*, *64*(1), 151-185.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological bulletin*, *125*(3), 367.

Bushee, B. J. (1998). The influence of institutional investors on myopic R&D investment behavior. *Accounting review*, 305-333.

Chaudhuri, Ranadeb and Ivkovich, Zoran and Pollet, Joshua Matthew and Trzcinka, Charles, (2013) What a Difference a Ph.D. Makes: More than Three Little Letters. Available at SSRN: http://ssrn.com/abstract=2344938

Cohen, L., Frazzini, A., & Malloy, C. (2008). The Small World of Investing: Board Connections and Mutual Fund Returns. *Journal of Political Economy, 116*(5), 951-979.

Coles, J. L., Daniel, N. D., & Naveen, L. (2006). Managerial incentives and risk-taking. *Journal of Financial Economics*, *79*(2), 431-468.

Core, J., & Guay, W. (2002). Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting Research*, *40*(3), 613-630.

Custódio, C., Ferreira, M. A., & Matos, P. (2013). Generalists versus specialists: Lifetime work experience and chief executive officer pay. *Journal of Financial Economics*, *108*(2), 471-492.

Custodio, C., Ferreira, M. A., & Matos, P. P. (2014). Do General Managerial Skills Spur Innovation?. *ECGI-Finance Working Paper*, (376).

Engelberg, J., Gao, P., & Parsons, C. A. (2012). The Price of a CEO's Rolodex. *Review of Financial Studies*, hhs114.

Falato, A., Li, D., & Milbourn, T. (2012). Which skills matter in the market for CEOs? Evidence from pay for CEO credentials. *Unpublished working paper. Federal Reserve Board and Washington University in St. Louis*.

Faleye, O. (2009). Classified boards, stability, and strategic risk taking. *Financial Analysts Journal*, 54-65.

Faleye, O., Kovacs, T., & Venkateswaran, A. (2013). Do better-connected CEOs innovate more?. *Journal of Financial and Quantitative Analysis, forthcoming*.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3-56.

Francis, B., Hasan, I., & Sharma, Z. (2011). Incentives and innovation: Evidence from CEO compensation contracts. *Bank of Finland Research Discussion Paper*, (17).

Gottesman, A. A., & Morey, M. R. (2006). Does a Better Education Make For Better Managers? An Empirical Examination of CEO Educational Quality and Firm Performance. Available at SSRN: http://ssrn.com/abstract=564443

Hall, B. H. (2005). The financing of innovation. *The Handbook of Technology and Innovation Management*, 409-430.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools* (No. w8498). National Bureau of Economic Research.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of economics*, 16-38.

Hall, B. H., & Lerner, J. (2009). *The financing of R&D and innovation* (No. w15325). National Bureau of Economic Research.

Hall, B. H., & Ziedonis, R. H. (2001). The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995. *RAND Journal of Economics*, 101-128.

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

He, J. J., & Tian, X. (2013). The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics*, *109*(3), 856-878.

He, Z., & Wintoki, B. M., (2014), The Cost of Innovation: R&D and High Cash Holdings in U.S. Firms. Available at SSRN: http://ssrn.com/abstract=2322029

Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are overconfident CEOs better innovators?. *The Journal of Finance*, *67*(4), 1457-1498.

Jalbert, T., Rao, R., & Jalbert, M. (2011). Does school matter? An empirical analysis of CEO education, compensation, and firm performance. *International Business & Economics Research Journal (IBER)*, *1*(1).

Kimberly, J. R., & Evanisko, M. J. (1981). Organizational innovation: The influence of individual, organizational, and contextual factors on hospital adoption of technological and administrative innovations. *Academy of management journal*, *24*(4), 689-713.

Lerner, J., & Wulf, J. (2007). Innovation and incentives: Evidence from corporate R&D. *The Review of Economics and Statistics*, *89*(4), 634-644.

Lyandres, E., & Palazzo, B. (2014). Strategic cash holdings and R&D competition: Theory and evidence. *Available at SSRN 2017222*.

Malmendier, U., & Nagel, S. (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?. *Quarterly Journal of Economics*, *126*(1), 373-416.

Malmendier, U., Tate, G., & Yan, J. (2011). Overconfidence and early-life experiences: the effect of managerial traits on corporate financial policies. *The Journal of finance*, *66*(5), 1687-1733.

Manso, G. (2011). Motivating innovation. *The Journal of Finance*, *66*(5), 1823-1860.

Serfling, M. A. (2014). CEO age and the riskiness of corporate policies. *Journal of Corporate Finance*, *25*, 251-273.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of financial studies*, *22*(1), 435-480.

Puri, M., & Robinson, D. T. (2009). The economic psychology of entrepreneurship and family business. *Available at SSRN 1422140*.

Tian, X., & Wang, T. Y. (2014). Tolerance for failure and corporate innovation. *Review of Financial Studies*, *27*(1), 211-255.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics*, 172-187.

Table I. Summary Statistics

This table reports descriptive statistics for variables of interest for the full sample in Panel A and for subsamples by firms having a CEO with Ph.D. degrees (Ph.D. CEOs) in Panel B. The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with available data on educational background of the CEOs in BoardEx, accounting data in Compustat, patent data from the NBER patent database, and stock return data from CRSP. All continuous variables are winsorized at their 1st and 99th percentiles. Dollar values are CPI-adjusted and expressed in 2006 dollars. In panel B, firms are divided based on whether the CEO on board holds Ph.D. degrees. t-tests (Wilcoxon-Mann-Whitney tests) are conducted to test for differences between the means (medians) for firms with Ph.D. CEOs and firms with non-Ph.D. CEOs. Variable definitions are reported in Appendix A. *, **, and*** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. Full Sample

Variable:	Mean	S.D.	25th Percentile	Median	75th Percentile
Dependent Variables					
Patent application dummy	0.433	0.495	0.000	0.000	1.000
No. of patents	14.414	49.384	0.000	0.000	4.000
Citation (raw) count	75.824	305.834	0.000	0.000	8.000
Citation (weighted) count	194.332	738.433	0.000	0.000	30.700
Citations per patent	2.280	6.800	0.000	0.000	1.330
Citations (weighted) per patent	5.470	13.300	0.000	0.000	6.120
CEO Characteristics:					
Age to CEO (years)	47.800	8.950	42.000	48.000	54.000
No. of CEO positions to date	1.290	0.661	1.000	1.000	1.000
No. of degrees	1.670	1.070	1.000	2.000	2.000
CEO MBA degree	0.269	0.443	0.000	0.000	1.000
CEO age	54.700	7.860	49.000	55.000	60.000
CEO tenure	7.830	7.890	2.000	5.000	11.000
CEO founder	0.133	0.340	0.000	0.000	0.000
CEO chairman	0.399	0.490	0.000	0.000	1.000
CEO male	0.985	0.123	1.000	1.000	1.000
Total compensation($000)	5,425.00	16,204.00	1,277.00	2,578.00	5,451.00
Salary and bonus ($000)	1,497.00	1,638.00	668.00	1,105.00	1,802.00
CEO delta ($000)	1,047.29	2,453.81	117.45	296.12	821.43
CEO vega ($000)	140.96	243.27	18.58	54.49	145.86
CEO stock ownership	0.031	0.066	0.001	0.004	0.021
Firm Characteristics:					
Assets ($mil)	5,728.00	23,428.00	496.000	1,229.00	3,716.00
Tangibility	0.298	0.210	0.137	0.246	0.416
Mktbk	2.250	1.620	1.290	1.720	2.550
Sales growth	0.160	0.398	0.016	0.099	0.216
ROA	0.046	0.117	0.022	0.056	0.094
Leverage	0.208	0.166	0.051	0.198	0.325
Cash-to-assets	0.143	0.175	0.019	0.065	0.206
R&D-to-assets	0.034	0.059	0.000	0.004	0.047
Institutional ownership	0.511	0.305	0.331	0.591	0.747
log(age)	2.890	0.825	2.300	3.000	3.640
Stock return volatility	0.031	0.016	0.020	0.028	0.038
Idiosyncratic risk	0.029	0.016	0.018	0.026	0.036
Annualized stock return	0.252	0.694	-0.139	0.210	0.559
G-index	9.200	2.710	7.000	9.000	11.000

Panel B. Subsample by Firms with Ph.D. CEOs

	Ph.D. CEO Obs=616					Non-Ph.D. CEO Obs=7,360		
Variable	Mean		Median		S.D.	Mean	Median	S.D.
Dependent Variables								
Patent application dummy	0.688	***	1.000	***	0.464	0.412	0.000	0.492
No. of patents	21.600	***	3.000	***	52.000	13.800	0.000	49.100
Citation (raw) count	98.900	***	3.000	***	283.000	74.000	0.000	308.000
Citation (weighted) count	302.000	***	15.000	***	819.000	186.000	0.000	731.000
Citations per patent	3.590	***	0.444	***	8.550	2.180	0.000	6.630
Citations (weighted) per patent	9.170	***	3.000	***	16.800	5.180	0.000	13.000
CEO Characteristics:								
Age to CEO (years)	46.800	***	46.000	***	10.000	47.900	48.000	8.860
No. of CEO positions to date	1.230	***	1.000		0.527	1.290	1.000	0.671
No. of degrees	2.780	***	3.000	***	1.040	1.580	2.000	1.020
CEO MBA degree	0.122	***	0.000	***	0.327	0.281	0.000	0.449
CEO age	55.300	**	56.000	**	7.480	54.700	55.000	7.890
CEO tenure	10.000	***	8.000	***	8.040	7.650	5.000	7.850
CEO founder	0.351	***	0.000	***	0.478	0.116	0.000	0.320
CEO chairman	0.428		0.000		0.495	0.397	0.000	0.489
CEO male	0.982		1.000		0.132	0.985	1.000	0.122
Total compensation($000)	7,667.00	***	2,983.00	***	21,255.00	5,244.00	2,556.00	15,713.00
Salary and bonus ($000)	1,492.00	***	955.00	***	1,899.00	1,497.00	1,117.00	1,615.00
CEO delta ($000)	1,299.29	***	474.13	***	2,552.67	1,027.37	282.72	2,444.87
CEO vega ($000)	169.43	***	81.38	***	253.84	138.73	52.82	242.29
CEO stock ownership	0.038	***	0.006	***	0.070	0.031	0.004	0.066
Firm Characteristics:								
Assets ($mil)	11,878.00	***	1,008.00	***	50,210.00	5,233.00	1,244.00	19,670.00
Tangibility	0.234	***	0.179	***	0.188	0.303	0.253	0.211
Mktbk	2.820	***	2.040	***	2.030	2.200	1.700	1.570
Sales growth	0.222	***	0.139	***	0.547	0.155	0.097	0.383
ROA	0.020	***	0.051	**	0.165	0.048	0.057	0.112
Leverage	0.178	***	0.153	***	0.170	0.210	0.202	0.165
Cash-to-assets	0.252	***	0.185	***	0.238	0.134	0.060	0.166
R&D-to-assets	0.073	***	0.052	***	0.081	0.031	0.001	0.056
Institutional ownership	0.502		0.561	*	0.308	0.512	0.593	0.305
log(age)	2.570	***	2.480	***	0.848	2.910	3.040	0.818
Stock return volatility	0.037	***	0.035	***	0.018	0.031	0.027	0.016
Idiosyncratic risk	0.035	***	0.033	***	0.017	0.029	0.025	0.015
Annualized stock return	0.329	**	0.274		0.847	0.246	0.207	0.680
G-index	8.460	***	8.000	***	2.620	9.260	9.000	2.700

Table II. Descriptive Statistics for the Prevalence of Ph.D. CEOs

This table presents descriptive statistics for the prevalence of Ph.D. CEO in the sample. Panel A shows the distribution of firm-years by Ph.D. CEO across Fama French 12 industry groups. Firms are divided based on whether the CEO on board holds Ph.D. degrees. Panel B presents the sample distribution by the presence of Ph.D. CEO and R&D intensity. R&D intensity is the R&D-to-assets ratio. For observations with non-zero R&D-to-assets ratios, firm-years are then sorted by the R&D intensity into quartiles. Panel C reports the top U.S. universities ranked by the number of doctoral degrees held by the CEOs in the sample. Ph.D. CEO is a dummy equal to one for firms having CEOs with doctorate degrees, and zero otherwise. The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with available data on educational background of the CEOs in BoardEx, accounting data in Compustat, patent data from the NBER patent database, and stock return data from CRSP.

Panel A. Distribution of Firm-years by Ph.D. CEO across Fama French 12 Industry Groups

Ph.D. CEO Obs=616				Non-Ph.D. CEO Obs=7,360			
Industry	Pct.	Industry	Pct.	Industry	Pct.	Industry	Pct.
Business Equipment	37.17	Chemicals	4.99	Business Equipment	20.39	Chemicals	4.80
Health	23.16	Consumer Durables	1.39	Manufacturing	17.79	Energy	4.29
Other	8.04	Telecommunication	1.39	Shops	16.09	Consumer Durables	3.62
Manufacturing	7.77	Consumer Nondurables	1.25	Other	13.20	Telecommunication	2.19
Energy	7.49	Utilities	n/a	Consumer Nondurables	9.27	Utilities	n/a
Shops	7.35	Money	n/a	Health	8.37	Money	n/a

Panel B. Distribution of Firm-years by Ph.D. CEO and R&D Intensity

	R&D=0	R&D>0			
		Q1	Q2	Q3	Q4
Ph.D. CEO Obs (%)	4.50	7.64	7.66	14.99	16.11
Non-Ph.D. CEO Obs (%)	95.50	92.36	92.34	85.01	83.89

Panel C. Distribution of Ph.D. CEOs by Universities

University	No. of Ph.D. Degrees	Pct.	University	No. of Ph.D. Degrees	Pct.
Stanford University	10	6.49	University of Michigan	2	1.30
Harvard University	6	3.90	University of Chicago	2	1.30
Columbia University	5	3.25	University of Texas Austin	2	1.30
New York University	4	2.60	University of Pennsylvania	1	0.65
UC Berkeley	3	1.95	Yale	1	0.65
Princeton University	3	1.95	Other	115	74.68

Table III. Ph.D. CEOs and Innovation: Baseline Results

This table presents the regression results on the relation between Ph.D. CEOs and innovation, proxied by patenting activities, patent citations and citations per patent. The dependent variables are *Patent Application Dummy* in the logit model (column 1), the natural logarithm of one plus total number of patents (column 2), the natural logarithm of one plus total number of citations (column 3), the natural logarithm of one plus total number of weighted citations (column 4), the natural logarithm of one plus non-self citations per patent (column 5), and the natural logarithm of one plus weighted, non-self citations per patent (column 6). *Patent Application Dummy* equals one if a firm applies for patents in that year, and zero otherwise. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Citation (raw) count* is the total number of non-self citations summed across all patents filed in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Citations per patent* is the total number of non-self citations received on the firm's patents, scaled by the number of patents filed. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed. The key variable of interest is *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	Patent Application	$\log(1+\text{patents})_{t+1}$	$\log(1+\text{citations})_{t+1}$	$\log(1+\text{weighted citations})_{t+1}$	$\log(1+\text{citations per patent})_{t+1}$	$\log(1+\text{weighted citations per patent})_{t+1}$
Model	Logit	Tobit	Tobit	Tobit	Tobit	Tobit
	(1)	(2)	(3)	(4)	(5)	(6)
Ph.D. CEO	0.665***	0.529***	0.817***	1.080***	0.322***	0.520***
	(2.851)	(2.703)	(2.771)	(2.873)	(2.899)	(2.972)
Size	0.548***	0.734***	0.994***	1.227***	0.280***	0.443***
	(7.707)	(9.172)	(8.386)	(8.305)	(6.547)	(6.751)
log(age)	-0.861**	-0.615	-0.505	-0.932	-0.186	-0.410
	(-2.019)	(-1.414)	(-0.760)	(-1.102)	(-0.708)	(-1.015)
$[\log(age)]^2$	0.179**	0.133	0.106	0.192	0.035	0.080
	(2.214)	(1.571)	(0.819)	(1.159)	(0.684)	(1.025)
Tangibility	0.083	0.836	1.143	1.227	0.076	0.133
	(0.173)	(1.533)	(1.322)	(1.141)	(0.228)	(0.264)
Mktbk	0.131***	0.115***	0.151***	0.204***	0.048**	0.087***
	(3.205)	(3.151)	(2.766)	(2.989)	(2.197)	(2.625)
Sales growth	0.081	0.057	0.164	0.236	0.133**	0.193*
	(0.762)	(0.638)	(1.110)	(1.226)	(2.039)	(1.935)
ROA	-1.146***	-0.467	-0.579	-0.894	-0.203	-0.459
	(-3.115)	(-1.259)	(-1.039)	(-1.225)	(-0.927)	(-1.305)
Leverage	-0.694*	-0.801**	-1.195**	-1.481*	-0.349	-0.542
	(-1.928)	(-2.142)	(-2.027)	(-1.953)	(-1.512)	(-1.488)
Cash-to-assets	1.564***	1.890***	2.688***	3.498***	0.946***	1.556***
	(4.236)	(4.899)	(4.437)	(4.570)	(4.030)	(4.334)
Annualized stock return	-0.017	0.036	0.074	0.043	0.025	0.014
	(-0.355)	(0.844)	(1.003)	(0.448)	(0.819)	(0.306)
Institutional ownership	0.343*	0.346	0.731**	0.893**	0.342**	0.494**
	(1.746)	(1.606)	(2.101)	(2.025)	(2.512)	(2.354)
log(1+CEO tenure)	0.063	0.093	0.109	0.138	0.041	0.062
	(1.011)	(1.453)	(1.125)	(1.117)	(1.094)	(1.060)
log(1+CEO age)	-0.164	-0.253	-0.299	-0.338	-0.041	-0.038
	(-0.407)	(-0.571)	(-0.440)	(-0.398)	(-0.163)	(-0.097)
CEO MBA degree	0.088	0.170	0.172	0.244	0.034	0.080
	(0.727)	(1.333)	(0.845)	(0.953)	(0.446)	(0.676)
log(1+CEO delta)	-0.074	-0.058	-0.099	-0.112	-0.036	-0.053
	(-1.076)	(-0.815)	(-0.933)	(-0.824)	(-0.847)	(-0.822)
log(1+CEO vega)	0.122**	0.150***	0.224***	0.286***	0.078***	0.126***
	(2.550)	(2.583)	(2.675)	(2.629)	(2.577)	(2.698)
CEO stock ownership	-0.742	0.271	0.771	0.506	-0.086	-0.244
	(-0.541)	(0.203)	(0.382)	(0.191)	(-0.100)	(-0.184)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	7,976	7,976	7,976	7,976	7,976	7,976
Pseudo R^2	0.3664	0.2377	0.25	0.2223	0.3001	0.2476

Table IV. Propensity Score Matching

This table presents the results on the relation between Ph.D. CEOs and firm innovation using a propensity score matched sample. Treated firms are defined as firms with a Ph.D. CEO. Each treated firm is matched to firms with a non-Ph.D. CEO using propensity score matching (with replacement and a radius of 1%). Firms are matched in the first year of CEO appointment by *log(1+ patents), Firm Size, Tangibility, Mktbk, ROA, Cash-to-assets, Annualized stock return, log(1+ CEO age)* and *CEO founder*. Panel A reports the parameter estimates from the Logit model used in estimating the propensity scores for the treatment and control groups. The dependent variable is *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. Panel B presents the results of balance tests of matched treatment and control groups in the first year of CEO appointment. Panel C shows the differences in mean innovation measures between treated and matched firms. Innovation measures including the number of patents, the number of weighted citations and the number of weighted citations per patent in the next year, all transferred into natural logarithm. Panel D presents the OLS regression results using the matched samples. The dependent variables are the three innovation proxies outlined above. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted). The OLS regressions control for all the matching variables, year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. Pre-matching Propensity Score Estimation

Dependent Variable	Ph.D. CEO
Model	Logit
	(1)
log(1+patents)	0.070**
	(2.318)
Size	0.142***
	(3.814)
Tangibility	-0.492
	(-1.552)
Mktbk	0.035
	(1.446)
ROA	-0.696***
	(-2.659)
Cash-to-assets	1.510***
	(5.908)
Annualized stock return	0.048
	(0.889)
log(1+CEO tenure)	0.075
	(1.469)
log(1+CEO age)	1.072***
	(4.079)
CEO founder	1.051***
	(9.748)
Industry FE	Yes
Year FE	Yes
N	7,626
Pseudo R^2	0.1027
p-value of χ^2	0.0000

Panel B. Balance Tests

	First-years with Ph.D. CEOs n=38	First-years with non-Ph.D. CEOs n=133		
Variable	Mean	Mean	Diff.	p-value
log(1+patents)	1.85	1.37	0.47	0.16
Size	8.42	7.48	0.93	0.01***
Tangibility	0.32	0.26	0.06	0.14
Mktbk	2.04	2.08	-0.04	0.88
ROA	0.00	0.01	-0.01	0.69
Cash-to-assets	0.11	0.17	-0.06	0.07*
Annualized stock return	0.21	0.20	-0.07	0.95
log(1+CEO age)	3.99	3.92	-0.08	0.43
CEO founder	0.01	0.00	-0.09	0.59

Panel C. Univariate tests

	Firm-years with Ph.D. CEOs n=156	Firm-years with non-Ph.D. CEOs n=525		
Variable	Mean	Mean	Diff.	
log(1+patents)$_{t+1}$	1.87	1.09	0.77	***
log(1+weighted citations)$_{t+1}$	2.39	1.43	0.96	***
log(1+weighted citations per patent)$_{t+1}$	0.99	0.61	0.38	***

Panel D. Multivariate Tests

Dependent Variable:	log(1+patents)$_{t+1}$	log(1+weighted citations)$_{t+1}$	log(1+weighted citations per patent)$_{t+1}$
Model		OLS	
	(1)	(2)	(3)
Ph.D. CEO	0.530**	0.623**	0.267**
	(2.334)	(2.022)	(2.282)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	681	681	681
Adjusted R^2	0.4354	0.4517	0.4659

Table V. Endogenous Matching between CEOs and Firms—Subsample Analysis

This table presents the regression results on the relation between Ph.D. CEOs and innovation using a restricted sample of CEOs with tenure at least three years (Panel A), or a subsample containing firms founded or co-founded by a Ph.D. CEO (Panel B). CEO is the founder if BoardEx identifies the CEO as a founder or co-founder of the company. In each panel, the dependent variables are the natural logarithm of one plus the total number of patents (column 1), the natural logarithm of one plus total number of weighted citations (column 2), and the natural logarithm of one plus weighted, non-self citations per patent (column 3). All models use one-year forward dependent variables. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted) in a given year. The key variable of interest is *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. Control variables include *Firm Size, log(age), $[\log(age)]^2$, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts and controls are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. Excluding Observations with CEO Tenure Less Than Three Years

Dependent Variable:	$\log(1+\text{patents})_{t+1}$	$\log(1+\text{weighted citations})_{t+1}$	$\log(1+\text{weighted citations per patent})_{t+1}$
Model		Tobit	
	(1)	(2)	(3)
Ph.D. CEO	0.562**	1.198***	0.576***
	(2.540)	(2.863)	(2.939)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	5,805	5,805	5,805
Pseudo R^2	0.2439	0.2294	0.2532

Panel B. Subsample of Firms Founded by a Ph.D. CEO

Dependent Variable:	$\log(1+\text{patents})_{t+1}$	$\log(1+\text{weighted citations})_{t+1}$	$\log(1+\text{weighted citations per patent})_{t+1}$
Model		Tobit	
	(1)	(2)	(3)
Ph.D. CEO	0.369***	1.096***	0.583***
	(5.356)	(9.027)	(10.070)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	1,103	1,103	1,103
Pseudo R^2	0.3456	0.3246	0.3523

Table VI. Potential Explanation: Ph.D. CEOs and Innovation Ability

This table presents the regression results of the effect of CEO ability on the relation between Ph.D. CEOs and innovation, using full sample (Panel A), and subsample split by innovation ability (Panel B). In each panel, the dependent variables are the natural logarithm of one plus the total number of patents (columns 1-2), the natural logarithm of one plus total number of weighted citations (columns 3-4), and the natural logarithm of one plus weighted, non-self citations per patent (columns 5-6). All models use one-year forward dependent variables. *High latent ability* is a dummy variable equals one if the age at which the individual took the first CEO position falls in the bottom decile among all the CEOs in the sample, and zero otherwise. *General Ability* is the first factor of the principle components analysis of five proxies of general managerial ability (Custodio, Ferreira, and Matos, 2013). *High innovation ability* is a dummy variable equals one if the number of patents filed by the firm is greater than the average number of patents filed by firms in the same industry during the year, where the industry is defined at the 2-digit SIC code level. The key variable of interest is the *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted) in a given year. Control variables include *Firm Size, log(age),* $[log(age)]^2$ *, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts and controls are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. Ph.D. CEOs, Innate Talent, General Skills, and Innovation Ability

Dependent Variable:	log(1+patents)$_{t+1}$			log(1+weighted citations)$_{t+1}$			log(1+weighted citations per patent)$_{t+1}$		
Model	Tobit			Tobit			Tobit		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High latent ability	-0.069			-0.252			-0.029		
	(-0.326)			(-0.657)			(-0.244)		
General ability		0.091			0.165			0.090	
		(1.393)			(1.298)			(1.503)	
High innovation ability			2.534***			4.654***			1.337***
			(27.845)			(26.560)			(20.226)
Ph.D. CEO	0.461**	0.505***	0.144	0.990***	1.069***	0.381*	0.311***	0.514***	0.147*
	(2.381)	(2.737)	(1.345)	(2.609)	(2.931)	(1.766)	(2.679)	(3.003)	(1.887)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	7,976	7,751	7,976	7,976	7,751	7,976	7,976	7,751	7,976
Pseudo R²	0.2373	0.2416	0.3577	0.2221	0.2247	0.3077	0.3004	0.2498	0.3735

Panel B. Subsample Split by CEO Innovation Ability

Dependent Variable:	log(1+patents)$_{t+1}$		log(1+weighted citations)$_{t+1}$		log(1+weighted citations per patent)$_{t+1}$	
Model	Tobit		Tobit		Tobit	
Innovation ability	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Ph.D. CEO	0.461**	0.144	0.990***	0.381*	0.311***	0.147*
	(2.381)	(1.345)	(2.609)	(1.766)	(2.679)	(1.887)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2,557	5,419	2,557	5,419	2,557	5,419
Pseudo R²	0.2060	0.2463	0.2139	0.2363	0.4282	0.2856

Table VII. Ph.D. CEOs and Innovative Efficiency

This table presents the regression results on the relation between Ph.D. CEOs and the effectiveness of innovation for given R&D expenditures, using full sample (Panel A), and subsample split by innovation ability (Panel B). The dependent variables are the natural logarithm of one plus the total number of patents (in column 1 of Panel A and columns 1-2 in Panel B), the natural logarithm of one plus total number of weighted citations (in column 2 of Panel A and columns 3-4 in Panel B), and the natural logarithm of one plus weighted, non-self citations per patent (in column 3 of Panel A and columns 5-6 in Panel B). All models use one-year forward dependent variables. A CEO has high innovation ability if the number of patents filed by the firm is greater than the average number of patents filed by firms in the same industry during the year, where the industry is defined at the 2-digit SIC code level. The key variable of interest is the *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted) in a given year. Only the tobit estimates on *Ph.D. CEO dummy* and *log(1+R&D-to-assets)* are reported. Control variables include *Firm Size, log(age),* $[log(age)]^2$ *, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts and controls are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A. Full Sample

Dependent Variable:	$log(1+patents)_{t+1}$	$log(1+weighted\ citations)_{t+1}$	$log(1+weighted\ citations\ per\ patent)_{t+1}$
Model		Tobit	
	(1)	(2)	(3)
Ph.D. CEO	0.448**	0.965***	0.478***
	(2.430)	(2.643)	(2.735)
log(1+R&D-to-assets)	11.310***	19.392***	8.171***
	(8.924)	(8.001)	(6.924)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	7,976	7,976	7,976
Pseudo R^2	0.2504	0.2308	0.2545

Panel B. Subsample Split by CEO Innovation Ability

Dependent Variable:	$log(1+patents)_{t+1}$		$log(1+weighted\ citations)_{t+1}$		$log(1+weighted\ citations\ per\ patent)_{t+1}$	
Model	Tobit		Tobit		Tobit	
Innovation ability	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Ph.D. CEO	0.227*	0.070	0.476**	0.106	0.182*	0.049
	(1.816)	(0.376)	(1.966)	(0.213)	(1.711)	(0.145)
log(1+R&D-to-assets)	6.368***	5.373***	8.322***	11.544***	1.870**	7.298***
	(6.231)	(4.273)	(4.705)	(3.373)	(2.408)	(3.060)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2,557	5,419	2,557	5,419	2,557	5,419
Pseudo R^2	0.2060	0.2463	0.2139	0.2363	0.4282	0.2856

Table VIII. Ph.D. CEOs, Innovation Productivity, and CEO Turnover

This table presents the logit regression results of CEO turnover on Ph.D. CEOs and innovation productivity. The dependent variable is a dummy equal to one if there is a CEO turnover in a given company. *Ph.D. CEO* is a dummy variable equal to one if a CEO has doctorate degrees, and zero otherwise. *log(1+patents)* is the natural logarithm of one plus the total number of patents. *log(1+weighted citations)*is the natural logarithm of one plus total number of weighted non-self citations. *log(1+weighted citations per patent)* is the natural logarithm of one plus weighted, non-self citations per patent. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted) in a given year. Only the logit estimates on *Ph.D. CEO dummy* and the interaction terms are reported. Control variables include *Firm Size, log(age),* $[\log(age)]^2$ *, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts and controls are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	CEO Turnover Dummy			
Model	Logit			
	(1)	(2)	(3)	(4)
Ph.D. CEO	-0.728***	-0.405*	-0.396**	-0.335*
	(-4.736)	(-1.959)	(-2.112)	(-1.869)
Ph.D. CEO * log(1+patents)		-0.211*		
		(-1.903)		
Ph.D. CEO * log(1+weighted citations)			-0.157**	
			(-2.378)	
Ph.D. CEO * log(1+weighted citations per patent)				-0.443***
				(-3.228)
log(1+patents)		0.037		
		(1.272)		
log(1+weighted citations)			0.009	
			(0.491)	
log(1+weighted citations per patent)				-0.007
				(-0.169)
F-statistic (Innovation Measure Ph.D. CEO + Ph.D. CEO)*		15.61***	11.78***	20.71***
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	8,972	8,972	8,972	8,972
Pseudo R^2	0.1031	0.1016	0.1018	0.1025

Table IX. Alternative Explanation 1: Ph.D. CEOs and Risk-Taking

This table presents the OLS regression results on the relation between Ph.D. CEOs and firm risk. The dependent variables are stock return volatility (columns 1-2), and the idiosyncratic risk (columns 2). *Stock return volatility* is calculated as the standard deviation of a firm's daily stock returns over the fiscal year. *Idiosyncratic risk* is measured as the standard deviation of the error terms estimated from regressing daily stock returns on Fama and French three factors, averaged over the fiscal year. The key variable of interest is *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. Control variables include *Firm Size, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), and log(1+ CEO vega).* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	Stock Return Volatility		Idiosyncratic Risk	
Model	OLS		OLS	
	(1)	(2)	(3)	(4)
Ph.D. CEO	0.001	0.001	0.001	0.001
	(1.043)	(0.993)	(0.922)	(0.903)
Size	-0.002***	-0.002***	-0.002***	-0.002***
	(-13.207)	(-7.136)	(-13.836)	(-7.122)
Tangibility	-0.002	-0.002	-0.002	-0.002
	(-1.063)	(-0.999)	(-1.240)	(-1.210)
Mktbk	0.001***	0.001***	0.001***	0.001***
	(5.133)	(5.874)	(5.114)	(5.983)
Sales growth	0.003***	0.004***	0.003***	0.004***
	(6.502)	(6.958)	(6.566)	(7.051)
ROA	-0.032***	-0.032***	-0.031***	-0.031***
	(-16.835)	(-16.937)	(-16.968)	(-16.994)
Leverage	0.002	0.003	0.002	0.003*
	(1.355)	(1.633)	(1.403)	(1.656)
Cash-to-assets	0.014***	0.014***	0.013***	0.013***
	(7.775)	(7.813)	(7.447)	(7.504)
Annualized stock return	0.002**	0.001	0.002**	0.001
	(2.514)	(1.537)	(2.511)	(1.524)
log(1+CEO tenure)	0.001***	0.001***	0.001***	0.001***
	(3.600)	(3.531)	(3.369)	(3.354)
log(1+CEO age)	-0.008***	-0.011***	-0.008***	-0.010***
	(-4.060)	(-6.903)	(-3.946)	(-6.779)
CEO MBA degree	-0.000	-0.000	-0.000	-0.000
	(-0.419)	(-0.974)	(-0.392)	(-0.903)
log(1+CEO delta)		-0.000		-0.000
		(-0.307)		(-0.493)
log(1+CEO vega)		-0.000**		-0.000***
		(-2.377)		(-2.596)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	8,588	7,976	8,588	7,976
Adjusted R^2	0.4531	0.4729	0.4401	0.4586

Table X. Alternative Explanation 2: Are Ph.D. CEOs Paid More?

This table presents the OLS regression results on the relation between Ph.D. CEO and total compensation. The dependent variable is the natural logarithm of one plus total value of a CEO's compensation package including option grants, converted into 2006 dollars. The key variable is *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. *G-index* is the corporate governance index in Gompers, Ishii, and Metrick (2003). *Institutional ownership* is the fraction of shares held by institutions over the fiscal year. *CEO chairman* is an indicator variable equals one if the CEO also serves as board chair, and zero otherwise. All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	log(1+Total compensation)			
Model	OLS			
	(1)	(2)	(3)	(4)
Ph.D. CEO	0.065	0.170*	-0.047	0.114
	(0.839)	(1.846)	(-0.263)	(1.333)
High G-index		0.085**		
		(2.376)		
High G-index * Ph.D. CEO		-0.246**		
		(-2.142)		
Institutional ownership			0.215***	
			(3.591)	
Institutional ownership * Ph.D. CEO			0.216	
			(0.862)	
CEO chairman				0.025
				(0.683)
CEO chairman * Ph.D. CEO				-0.126
				(-0.744)
Size	0.510***	0.508***	0.508***	0.509***
	(35.957)	(32.333)	(35.831)	(35.616)
Mktbk	0.121***	0.129***	0.118***	0.121***
	(6.855)	(6.389)	(6.711)	(6.876)
Previous year stock return	0.094***	0.084***	0.088***	0.094***
	(4.528)	(3.513)	(4.235)	(4.532)
Previous two years stock return	0.067***	0.064***	0.064***	0.067***
	(3.958)	(3.223)	(3.818)	(3.972)
Idiosyncratic risk	2.624*	2.799*	3.506**	2.608*
	(1.933)	(1.778)	(2.514)	(1.926)
log(1+CEO tenure)	0.089*	0.083	0.084*	0.089*
	(1.749)	(1.508)	(1.660)	(1.757)
[log(1+CEO tenure)]^2	-0.041**	-0.036**	-0.040**	-0.042***
	(-2.567)	(-2.033)	(-2.522)	(-2.591)
*F-statistic (Governance Measure * Ph.D. CEO + Ph.D. CEO)*		0.29	2.86*	0.01
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	7,976	6,629	7,976	7,976
Adjusted R^2	0.4983	0.4999	0.5020	0.4984

Table XI. Ph.D. CEOs and Innovation: Robustness Tests

This table presents the robustness of regression results on the relation between Ph.D. CEOs and innovation. In each panel, the dependent variables are the natural logarithm of one plus the total number of patents, the natural logarithm of one plus total number of weighted citations, and the natural logarithm of one plus weighted, non-self citations per patent. Panel A displays the estimates using subsample of firms not in high-tech industries. Panel B shows the results when restricting the time period leading up to the Dot.com bubble. Panel C reports the results after including dummies indicating whether the Ph.D. is in technical field. Panel D presents the estimates after controlling for CEO's technical background. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted) in a given year. The key variable of interest is *Ph.D. CEO dummy*, which equals one if a CEO has doctor degrees, and zero otherwise. Control variables include *Firm Size, log(age), [log(age)]^2 , Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts and controls are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	log(1+patents)$_{t+1}$	log(1+weighted citations)$_{t+1}$	log(1+weighted citations per patent)$_{t+1}$
Model		Tobit	
	(1)	(2)	(3)
Panel A. Excluding high-tech industries			
Ph.D. CEO	0.648**	1.500**	0.802***
	(2.211)	(2.476)	(2.582)
N	5,682	5,682	5,682
Pseudo R²	0.263	0.2424	0.2648
Panel B. Before 1997 (the Dot.com Bubble)			
Ph.D. CEO	0.590*	1.243**	0.707**
	(1.795)	(2.107)	(2.342)
N	2,328	2,328	2,328
Pseudo R²	0.2286	0.1908	0.2042
Panel C. Controlling for Ph.D. in Technical field			
Ph.D. CEO	0.446**	0.975**	0.489***
	(2.165)	(2.456)	(2.640)
Ph.D. in Engineering &	1.186**	1.828*	0.670
Science	(2.440)	(1.774)	(1.233)
Ph.D. in Medicine	-0.501	-1.626	-0.736**
	(-0.594)	(-1.538)	(-1.973)
N	7,976	7,976	7,976
Pseudo R²	0.2386	0.2229	0.2481
Panel D. Controlling for CEO's technical background			
Ph.D. CEO	0.374*	0.705*	0.335*
	(1.788)	(1.706)	(1.715)
Engineering & Science	0.429**	0.993***	0.488***
Degree	(2.360)	(2.920)	(3.194)
Medical Degree	-0.472	-1.104	-0.536
	(-0.626)	(-1.096)	(-1.389)
N	7,976	7,976	7,976
Pseudo R²	0.2390	0.2237	0.2492
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

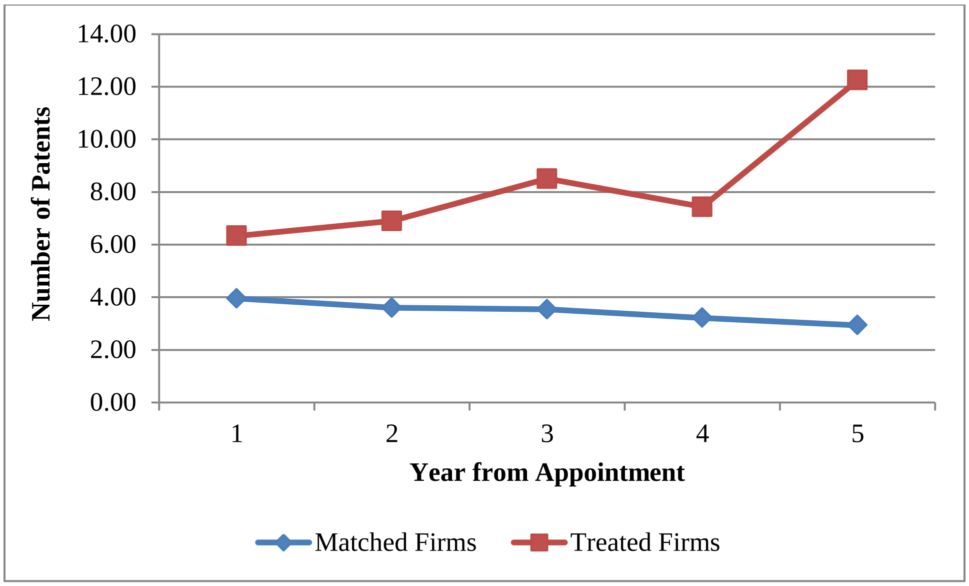
Table XII. Instrumental Variables Approach

This table presents the results on the relation between Ph.D. CEOs and firm innovation using an instrumental variables approach. Column (1) presents first stage regression results, and columns (2)-(4) present second-stage estimates from a Tobit model. The dependent variable in column (1) is the *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. *No. of degrees* is the total number of degrees that a CEO has obtained. *CEO MBA degree* is an indicator variable equals to one if a CEO holds an MBA degree, and zero otherwise. *No. of CEO positions to date* is the number of firms the individual worked as CEO. In the next three columns, the dependent variables are the natural logarithm of one plus the total number of patents (column 2), the natural logarithm of one plus total number of weighted citations (column 3), and the natural logarithm of one plus weighted, non-self citations per patent (column 4). In the second stage regressions, the Ph.D. CEO dummy is the predicted value from the first stage regression. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted). All Tobit regressions incorporate the full set of controls as specified in Equation (1). Control variables include *Firm Size, log(age),* $[log(age)]^2$ *, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	First-stage	Second-stage		
Dependent Variable:	Ph.D. CEO	$\log(1+\text{patents})_{t+1}$	$\log(1+\text{weighted citations})_{t+1}$	$\log(1+\text{weighted citations per patent})_{t+1}$
Model	OLS		Tobit	
	(1)	(2)	(3)	(4)
Ph.D. CEO		2.230***	4.380***	2.200***
		(3.050)	(3.082)	(3.226)
log(No. of degrees)	0.166***			
	(22.561)			
CEO MBA degree	-0.125***			
	(-18.941)			
CEO positions to date	-0.006*			
	(-1.785)			
Controls	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	7,976	7,976	7,976	7,976
Adjusted R^2	0.1419			
F-statistic/Wald $\chi2$	36.49	5,461.85	4,825.94	4,099.04
Test of weak instruments (p-value)	772.31(0.000)			
Hansen J test (p-value)		4.573 (0.102)	2.632 (0.268)	4.758 (0.093)

Figure I. Number of Patents after CEO Appointments

This figure shows the total number of patents for the sample of treated firms (firms with Ph.D. CEOs) and matched firms (firms with non-Ph.D. CEOs) since the first year of CEO appointment. Each treated firm is matched to firms with non-Ph.D. CEOs using propensity score matching (with replacement and a radius of 1%). Firms are matched in the first year of CEO appointment by *log(1+ patents), Firm Size, Tangibility, Mktbk, ROA, Cash-to-assets, Annualized stock return, log(1+ CEO age)* and *CEO founder*. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. The sample consists of 156 treated firm-years and 525 matched firm-years. The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP.



Appendix: Variable Definition

I. Dependent Variables:

Patent application dummy: an indicator equals one if a firm applies for patents in that year, and zero otherwise.

No. of patents: total number of patents applied for (and eventually granted) during the year.

Citation (raw) count: total number of non-self citations summed across all patents filed (and eventually granted) during the year.

Citation (weighted) count: total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed (and eventually granted) during the year.

Citations per patent: total number of non-self citations received on the firm's patents, scaled by the number of patents filed (and eventually granted) during the year.

Citations (weighted) per patent: total number of weighted, non-self citations , scaled by the number of patents filed (and eventually granted) during the year.

II. CEO Characteristics:

Ph.D. CEO: an indicator variable equals one if a CEO has doctorate degrees, and zero otherwise.

Age to CEO: age at which the individual took the first CEO position.

No. of CEO positions to date: the number of firms the individual worked as CEO.

No. of degrees: the total number of degrees a CEO has obtained.

CEO MBA degree: an indicator variable equals one if a CEO holds an MBA degree, and zero otherwise.

CEO age: the age of the current CEO during the year, measured in years.

CEO tenure: years the CEO has been in the firm.

CEO founder: a dummy equals one if the CEO is also the founder or co-founder of the company.

CEO chairman: a dummy equals one if the CEO also serves as board chair, and zero otherwise.

CEO male: an indicator variable equals one if the CEO is a man.

Total compensation: total value of a CEO's compensation package including option grants for the year in ExecuComp, converted into 2006 dollars (in thousands)

Salary and bonus: value of a CEO's total salary and bonus for the year in ExecuComp, converted into 2006 dollars (in thousands).

CEO delta: the dollar increase in the CEO's wealth for a 1% increase in stock price, converted into 2006 dollars (in thousands).

CEO vega: the dollar increase in the CEO's option holdings for a 1% increase in stock return volatility, converted into 2006 dollars (in thousands).

CEO stock ownership: the ratio of the CEO's holdings of common stocks to the total shares outstanding.

Engineering & Science degree: An indicator variable takes a value of one if the CEO has a degree in Engineering, Science, Physics, Chemistry, or Biology.

Medical degree: An indicator variable equals one if the CEO has a degree in Medicine, Optometry, Osteopathy, or Pharmacy.

Ph.D. in Engineering & Science: An indicator variable takes a value of one if the CEO has a doctorate degree in Engineering or Science.

Ph.D. in Medicine: An indicator variable equals one if the CEO has a doctorate degree in Medicine (e.g., M.D., Doctor of Osteopathic Medicine, and Doctor of medical Physics).

High latent ability: A dummy variable equals one if the age at which the individual took the first CEO position is in the bottom decile among all the CEOs in the sample, and zero otherwise.

General ability: The General Ability Index is the first factor of the principle components analysis of five variables: 1) the number of different positions in the past, 2) the number of firms where a CEO has worked, 3) the number of industries where a CEO has worked, 4) a dummy variable indicating whether a CEO held CEO positions before, and 5) a dummy indicating whether a CEO worked for a conglomerate firm (Custodio, Ferreira, and Matos, 2013, 2014).

High innovation ability: A dummy variable equals one if the number of patents filed by the firm is greater than the average number of patents filed by firms in the same industry during the year, where the industry is defined at the two-digit SIC code level.

III. Firm Characteristics:

Assets: the book value of assets, converted into 2006 dollars (in millions).

Size: Natural logarithm of book assets.

Tangibility: The ratio of net property, plant, and equipment to book assets.

Mktbk (Market-to-Book): Total assets minus book value of common equity plus the market value of common equity, divided by book value of assets.

Sales Growth: the sales growth rate from year *t-1* to year *t*.

ROA: The ratio of net income to total book assets.

Industry-adjusted ROA: industry demeaned ROA, calculated as the firm's ROA minus the mean ROA of firms in the same industry every year, where industry is defined at the two-digit SIC code level.

Leverage: The ratio of long-term debt plus debt in current liabilities to total assets.

Cash-to-assets: Cash plus marketable securities, divided by book value of total assets.

R&D-to-assets: The ratio of R&D expenditures to book assets; missing values are set to zero.

Institutional ownership: the fraction of shares held by institutions over fiscal year, calculated as the arithmetic mean of the four quarterly institutional holdings.

Age: the number of years the firm is listed with a non-missing stock price on Compustat.

Stock return volatility: the standard deviation of a firm's daily stock returns over the fiscal year.

Idiosyncratic risk: the standard deviation of the error terms estimated from regressing daily stock returns on Fama and French three factors, averaged over the fiscal year.

Annualized stock return: the firm's stock return averaged over the fiscal year.

G-index: the corporate governance index in Gompers, Ishii, and Metrick (2003), obtained from IRRC RiskMetrics Governance database.

High-tech Industries are defined following Brown, Fazzari, and Petersen (2009) as SIC codes 283, 357, 366, 367, 382, 737, 384.

Online Appendices for

“Are CEOs with a Ph.D. More Innovative?”

Not Intended for Publication Unless Requested

Zhaozhao He

University of Kansas School of Business

zzhe@ku.edu

September, 2014

Table A.I: The Determinants of Successor CEO with Ph.D. Degrees

This table presents the logit estimates examining the likelihood of successor being a Ph.D. CEO as the former CEO during two years surrounding the CEO turnover. Turnover events are identified using ExecuComp database over the period of 1992-2004. Turnover years are excluded to ensure the analysis of two different individuals. The dependent variable is *Ph.D. CEO dummy (in year t)*, which equals one if a CEO has doctorate degrees in a given year, and zero otherwise. The key variable of interest is *Ph.D. CEO_{t-2}*. All control variables are lagged by two years. Controls include *Firm Size, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, log(1+ CEO tenure), and log(1+ CEO age).* Industries are defined at the 2-digit SIC code level. The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	Ph.D. CEO			
Model	Logit			
	(1)	(2)	(3)	(4)
Ph.D. CEO $_{t-2}$	3.391***	3.271***	3.091***	3.071***
	(18.373)	(15.951)	(12.569)	(12.253)
Size $_{t-2}$		0.207***	0.248***	0.234***
		(2.885)	(3.339)	(3.147)
Tangibility $_{t-2}$		-0.001	0.589	0.569
		(-0.003)	(0.767)	(0.742)
Mktbk $_{t-2}$		-0.064	-0.131*	-0.124
		(-0.970)	(-1.682)	(-1.565)
Sales growth $_{t-2}$		0.449***	0.491***	0.501***
		(2.777)	(3.078)	(3.175)
ROA $_{t-2}$		-0.528	-0.120	-0.289
		(-0.780)	(-0.162)	(-0.406)
Leverage $_{t-2}$		-0.365	-0.237	-0.292
		(-0.498)	(-0.281)	(-0.358)
Cash-to-assets $_{t-2}$		0.921	0.633	0.731
		(1.260)	(0.810)	(0.948)
Annualized stock return $_{t-2}$		0.163	0.252	0.247
		(1.124)	(1.586)	(1.570)
log(1+CEO tenure) $_{t-2}$				1.310
				(1.337)
log(1+CEO age) $_{t-2}$				-0.118
				(-0.932)
Industry FE	No	No	Yes	Yes
Year FE	No	No	Yes	Yes
N	2,342	2,277	1,700	1,700
Pseudo R^2	0.2296	0.2419	0.2690	0.2712

Table A.II: Ph.D. CEOs, Operating Performance, and CEO Turnover

This table presents the logit regression results of CEO turnover on Ph.D. CEOs and firm operating performance. The dependent variable is a dummy equal to one if there is a CEO turnover in a given company. *Ph.D. CEO* is a dummy variable equal to one if a CEO has doctorate degrees, and zero otherwise. *Industry-adjusted ROA* is calculated as the firm's ROA minus the mean ROA of firms in the same industry every year. *General Ability* is the first factor of the principle components analysis of five proxies of general managerial ability (Custodio, Ferreira, and Matos, 2013). Only the logit estimates on *Ph.D. CEO dummy, General ability,* and the interaction terms are reported. Control variables include *Firm Size, log(age), $[log(age)]^2$, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts and controls are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	CEO Turnover Dummy				
Model	Logit				
	(1)	(2)	(3)	(4)	(5)
Ph.D. CEO	-0.584***		-0.583***	-0.454***	-0.445***
	(-4.340)		(-4.161)	(-3.000)	(-2.812)
Ph.D. CEO * ROA $_{t-1}$	-0.596		-0.542		
	(-0.910)		(-0.781)		
General ability		0.189***	0.187***		0.155***
		(4.777)	(4.678)		(3.175)
General ability * ROA $_{t-1}$		0.295	0.298		
		(1.084)	(1.044)		
Ph.D. CEO * Industry-adjusted ROA $_{t-1}$				-3.002	-2.908
				(-1.175)	(-1.080)
General ability * Industry-adjusted ROA $_{t-1}$					1.155
					(1.623)
ROA $_{t-1}$	-0.696**	-0.734***	-0.714**		
	(-2.401)	(-2.643)	(-2.340)		
Industry-adjusted ROA $_{t-1}$				-0.919	-1.089
				(-0.802)	(-0.945)
Controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	11,018	10,684	10,684	10,814	10,598
Pseudo R^2	0.0902	0.0920	0.0945	0.0867	0.0923

Table A.III: Ph.D. CEOs, Firm Value, and Operating Performance

This table presents the OLS regression results on the relation between Ph.D. CEOs and Tobin's Q (columns 1-3), and the relation between Ph.D. CEOs and industry-adjusted ROA (columns 4-6). For each investigation, results are shown for the full sample (columns 1 and 4), and for subsamples split by innovation ability (columns 2-3 and 5-6). A CEO has high innovation ability if the number of patents filed by the firm is greater than the average number of patents filed by firms in the same industry during the year, where the industry is defined at the 2-digit SIC code level. The key variable of interest is the *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. Control variables include *Firm Size, Sales growth, industry-adjusted ROA(for the regressions of Tobin's Q), R&D-to-assets, Annualized stock return, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable:	Tobin's Q			Industry-adjusted ROA		
Model	OLS			OLS		
Innovation ability	All	High	Low	All	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Ph.D. CEO	0.207	0.330	0.125	-0.013	-0.013	-0.017
	(1.492)	(1.614)	(0.858)	(-1.309)	(-1.036)	(-1.383)
Size	-0.072***	0.036	-0.153***	0.001	0.003	-0.002
	(-2.632)	(0.777)	(-5.301)	(0.519)	(1.217)	(-0.876)
Sales growth	0.001	0.004	0.001	0.000	-0.001	0.000
	(1.051)	(0.455)	(1.060)	(0.402)	(-1.308)	(1.212)
Industry-adjusted ROA	3.943***	4.566***	3.502***			
	(10.417)	(7.499)	(8.362)			
R&D-to-assets	9.164***	8.001***	9.270***	-0.427***	-0.317***	-0.561***
	(11.508)	(6.867)	(9.869)	(-4.900)	(-3.241)	(-4.343)
Annualized stock return	0.390***	0.499***	0.336***	0.005*	0.009*	0.004
	(10.678)	(7.137)	(8.529)	(1.835)	(1.863)	(1.046)
log(1+CEO tenure)	0.067**	0.096*	0.064**	0.010***	0.009**	0.009***
	(2.215)	(1.709)	(2.123)	(4.307)	(2.407)	(3.641)
log(1+CEO age)	-0.678***	-1.533***	-0.266	0.007	0.020	0.005
	(-2.692)	(-3.622)	(-1.138)	(0.433)	(0.650)	(0.316)
CEO MBA degree	0.028	0.130	-0.033	0.003	0.010	-0.000
	(0.435)	(1.035)	(-0.485)	(0.764)	(1.292)	(-0.061)
CEO stock ownership	1.088*	3.902***	0.179	0.020	0.151**	-0.013
	(1.822)	(3.006)	(0.389)	(0.617)	(2.299)	(-0.416)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	8,507	2,679	5,828	8,510	2,680	5,830
Adjusted R^2	0.3183	0.3487	0.3247	0.0426	0.0344	0.0534

Table A.IV: Other Robustness Checks

This table presents the robustness of the results on the relation between Ph.D. CEOs and innovation. In each panel, the dependent variables are the natural logarithm of one plus the total number of patents (column 1), the natural logarithm of one plus total number of weighted citations (column 2), and the natural logarithm of one plus weighted, non-self citations per patent (column 3). Panel A shows the estimates using a Negative Binomial model. Panel B reports the OLS estimates. Panel C shows the estimates when using two-year forward dependent variables. Panel D presents estimates after controlling for CEO power. *CEO chair* is an indicator equals to one if the CEO also serves as board chair, and zero otherwise. *CEO founder* is an indicator equals to one if the CEO is also the founder or co-founder of the company. *Patent count* is the total number of patents applied for (and eventually granted) in a given year. *Weighted citation count* is the total number of non-self citations, weighted by the weighting index from Hall, Jaffe, and Trajtenberg (2001, 2005), and then summed across all patents filed in a given year. *Weighted citations per patent* is the total number of weighted, non-self citations, scaled by the number of patents filed (and eventually granted) in a given year. The key variable of interest is *Ph.D. CEO dummy*, which equals one if a CEO has doctorate degrees, and zero otherwise. Control variables include *Firm Size, log(age),* $[log(age)]^2$ *, Tangibility, Mktbk, Sales growth, ROA, Leverage, Cash-to-assets, Annualized stock return, Institutional ownership, log(1+ CEO tenure), log(1+ CEO age), CEO MBA degree, log(1+ CEO delta), log(1+ CEO vega), and CEO stock ownership.* All the regressions control for year and industry fixed effects (at 2-digit SIC code level). The full sample comprises all non-financial and non-utility U.S. industrial firms in ExecuComp from 1992 to 2004, with data on educational background of the CEOs in BoardEx, and data coverage in Compustat, NBER patent database, and CRSP. All variables are defined in Appendix A. Intercepts and controls are not shown. The t-statistics in parentheses are based on robust standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable: Model	log(1+patents)$_{t+1}$	log(1+weighted citations)$_{t+1}$ Tobit	log(1+weighted citations per patent)$_{t+1}$
	(1)	(2)	(3)
Panel A. Negative Binomial			
Ph.D. CEO	0.467***	0.598***	0.283**
	(2.832)	(2.763)	(2.079)
N	7,976	7,976	7,976
Panel B. OLS			
Ph.D. CEO	0.239*	0.389**	0.189**
	(1.886)	(2.016)	(2.288)
N	7,976	7,976	7,976
Adjusted R^2	0.4400	0.4247	0.3937
Panel C. Using 2-year forward dependent variables			
Ph.D. CEO	0.555***	1.157***	0.527***
	(2.719)	(2.992)	(3.025)
N	6,277	6,277	6,277
Pseudo R^2	0.2437	0.2281	0.255
Panel D. Controlling for CEO Power			
Ph.D. CEO	0.463**	0.994***	0.499***
	(2.350)	(2.615)	(2.800)
CEO chair	0.128	0.092	-0.022
	(1.033)	(0.380)	(-0.194)
CEO founder	0.402**	0.517	0.143
	(2.031)	(1.361)	(0.797)
N	7,976	7,976	7,976
Pseudo R^2	0.2389	0.2226	0.2477
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes