Disregarding the Shoulders of the Giants: Inferences from Innovation Research*

David M. Reeb
NUS Business School
National University of Singapore, Singapore, 119245
Email: dmreeb@nus.edu.sg

Wanli Zhao
Hanqing Advanced Institute of Economics and Finance
The Renmin University, Beijing, China 100872
Email: zhaowl@msn.com

ABSTRACT

Using innovation research as a platform, we show that studies proposing new covariates rarely include previously identified determinants. Yet, only a sparse set of these hypothesized economic variables provide material, independent information about patents and citations. We then show that the inferences of previous studies can differ when including or excluding key economic determinants of innovation. Additional tests reveal that common solutions, including fixed effects and plausible shocks, do not mitigate the need to include previously identified innovation covariates. Rather than randomly selecting control variables, our analysis offers a framework for researchers to incorporate prior economic determinants of innovation.

JEL Classifications: G30; O30; G32; O34;
Keywords: variable selection; machine learning; lasso; innovation; patents; omitted variable

---

***** We benefited from discussions and comments from Sumit Agarwal, Chunrong Ai, Tara Bhandari, Ekkehart Boehmer, Simba Chang (discussant), Ben Charoenwong, Tony Cookson (discussant), Jin-Chuan Duan, Daniel Ferreira (discussant), Huasheng Gao, Vladimir Ivanov, Chao Jiang, Kathleen Kahle, Charles Lee, Ram Mudambi, Mattias Nilsson, George Papadakis, Ivan Png, Christopher Polk, Sungjune Pyun, Wenlan Qian, Srini Sankaraguruswamy, Johan Sulaeman, Parth Venkat, Matthew Wynter, Bernard Yeung, Mengxin Zhao, and seminar participants at 2018 Singapore Scholars Symposium, FMA 2018 conference, IESEG, the National University of Singapore, the National Cheng Chi University, Renmin University of China, the Securities and Exchange Commission, Southwestern University of Finance and Economics, and University of Oklahoma.

1. Introduction

In this study, we address a commonly encountered problem for empirical researchers, namely the large set of previously identified determinants on the same outcome variable. For instance, in the investment literature, if the predicative power of a covariate is subsumed by a benchmark, it provides no advantage to the investors. In empirical corporate finance research, we encounter similar situations where the selection of previously identified determinants to include in new studies seems ad hoc and random. Yet, there is little guidance on how to handle this problem nor is there a common agreement that a problem exists. Using innovation research as the platform, we aim to aggregate existing research and facilitate future studies by exploring three issues in this study. First, we show that among the numerous previously identified determinants of corporate innovation, only a small set of those covariates provide material and independent explanatory power. Innovation research provides a natural laboratory due to recent interest in this topic across finance, economics, and strategy journals. In addition, the public availability of the data ensures that the empirical results are strictly comparable between studies. Finally, the variable constructions and specifications are also similar among the studies in this field. Rather than using machine learning to data mine a wild host of potential variables, we use these techniques to evaluate and compare previously proposed economic determinants of corporate innovation.

Second, after identifying the set of the key variables from previously hypothesized economic determinants, we investigate how excluding these variables affects the robustness of previous studies. We find the key identified variables, if included, often invalidate the findings of the previous studies. We acknowledge that whether to include those key variables or not in the first place depends on the research question. We show that the inference or the interpretation of the

empirical findings could change drastically and thus needs careful attention when conducting new research.

Finally, to the extent that most empirical research is to identify marginal causal effect, one may argue that the omitted variable problem may not be a concern when an exogenous change in the independent variable can be exercised. The assumption of the exclusion condition however is not testable. As such, we show that common techniques such as fixed effects and plausible shocks do not mitigate the necessity of inclusion previous economic determinants of innovation. At a minimum, we suggest checking the exclusion condition with the key identified variables as they originate from prior studies with economic inferences.

We identify fifty-three recent articles in the most prominent finance and economics journals that propose new economically important patent covariates.[1] Strikingly, we observe that these studies rarely condition the analysis on a similar set of control variables, even though they all use the same dependent variable(s).[2] Clearly, it is imperative to aggregate this literature by identifying which characteristics provide material, independent information about patents and citations. Otherwise, uncovering new explanatory variables or features of corporate innovation becomes challenging without a clear understanding of the structure of the existing body of research.

We investigate how to aggregate the substantial amount of literature on the determinants of corporate innovation using a data-driven approach to test prior economic determinants with the

[1] Since 2010, the *Journal of Financial Economics* includes such nineteen such articles, *Management Science* ten, the *Review of Financial Studies* seven, the *Journal of Finance* six, the *Academy of Management Journal* five, the *American Economic Review* two, while the *Accounting Review*, *Strategic Management Journal*, *Quarterly Journal of Economics*, and *Econometrica* each contain one such article. A simple analysis of the 409 numerical variables in Compustat, shows that 70% are correlated with patent activity, which illustrates the difficulty in interpreting or assessing this body of research. Without a central theory of corporate innovation, identifying the relevant conditioning variables in empirical studies is challenging.

[2] These differing control variables are not a function of discipline heterogeneity. For instance, in six recent finance articles about patents (see Appendix A), twenty-five different control variables are included in the analyses. These six studies share only one common previously identified attribute of innovation, firm size, with some variables included quite frequently and others only listed a single time (e.g., board size).

goal of establishing a set of control variables that offer independent explanatory power on innovation. Not surprisingly, we find that only a small set of key variables identified in prior research survive the "horse race."

Our tests take advantage of machine learning techniques as they provide a systematic method for feature selection. Athey (2018) and Einav and Levin (2014) observe that natural experimental methods and recent advances in machine learning provide complementary tools for social science research, highlighting the benefits of these variable selection methods in identifying relevant control variables found in many proposed determinants. Machine learning techniques typically rely on semi-parametric algorithms, explicitly building on out-of-sample verification to compare different models from the in-sample analysis (Mullainathan and Spiess, 2017). Our base analysis uses the Adaptive Lasso technique to assess the explanatory power of previously proposed determinants of corporate innovation.[3] Similarities in outcomes from different regularization approaches (Elastic Net, Group Lasso, and Stepwise regressions) and different time periods, suggest the results stem from the underlying data generating process rather than weaknesses of any particular approach.

We assemble thirty-five potential determinants of corporate innovation based on the capital market literature regarding corporate innovation, with our main tests focusing on firms with patent applications (see Koh et al., 2016; Lerner and Seru, 2017). To better insure that our results stem from the true data generating process, rather than irrelevant covariates, we implement our machine learning tests by assessing the data within short (two-year) rolling windows and across the entire

[3] A central advantage of machine learning approaches is that it limits concerns about overfitting or mistaking random variations in the data as underlying trends (Caner and Fan, 2015). Adaptive Lasso provides an especially popular approach to variable selection in the natural sciences because it can give the same coefficient estimates as if one knew, with high probability or asymptotically, the true underlying model, i.e., the oracle property (Zou, 2006; Hui et al., 2015). Abadie and Kasy (2019) evaluate different regularization methods for empirical economics research by focusing on model selection criteria. Athey and Imbens (2017) discuss machine learning in econometrics. See Appendix B for a brief description of Adaptive Lasso.

sample period, relying on out-of-sample analysis for cross-validation. Performing the analysis across the entire time-period, in different time periods, or with different window choices emphasizes the relevance of these covariates as key conditioning variables (which explain over 90% of the explanatory power of the all-variable specification).

Our initial analysis focuses on the quantity of innovation, investigating whether previously identified covariates provide material, independent explanatory power for the number of patent applications.[4] The quantity of corporate innovation is a common measure of innovation in financial market research, especially in studies about managerial incentives to invest in innovation. Among patenting firms, we find four of the thirty-four (we exclude R&D stock in this test) variables explain patents in at least two thirds of the rolling windows of our sample. These four variables include stock liquidity (Bernstein, 2015; Fang et al., 2014), firm size, CEO reputation or centrality (Faleye et al., 2014), and industry citation intensity (Hall et al., 2001). Two additional variables are significant in the majority of windows, namely analyst following and industry patent intensity. Interestingly, we find that twenty of the candidate variables never survive the selection process in any single two-year rolling window.[5]

We separately repeat the analysis for innovation productivity, which includes R&D stock and gives thirty-five potential patent determinants. Unsurprisingly, R&D stock is an important variable in explaining patent activity and substantively influences the analysis. We find seven of the thirty-five variables explain patents in at least two thirds of the rolling windows of our sample.

[4] In the machine learning literature, due to the focus on pure predictive power over covariate identity, both non-linear forms of the variables and interaction terms are often included in the analysis. In contrast, our goal centers on aggregating the existing set of innovation covariates rather than increasing the predictive power from previously proposed determinants. Of additional concern in our setting is the difficulty in selecting the true innovation covariates with the inclusion of additional correlated covariates (i.e., the non-linear forms and interaction terms). Studies that identify non-linear effects (e.g., Im and Shon, 2019) could be affected by such concerns.

[5] Even though these results can not provide any causal inferences, they may still give some interesting interpretations. For instance, the short-term earnings pressure from analysts tend to curb innovation activity and this effect stands out when comparing with other internal governance mechnisms.

These seven variables include CEO centrality (Faleye et al., 2014), stock liquidity (Bernstein, 2015; Fang et al., 2014), R&D stock (Balsmeier et al., 2017), firm size, analyst following (He and Tian, 2013), and both industry citation and patent intensity (Hall et al., 2001). These results provide a potential standard set of control variables in studies that focus on innovation productivity, such as studies that argue governance improves R&D efficiency.

We undertake a similar analysis for patent citations to gauge innovation quality, starting with the same thirty-five potential explanatory variables. We find that two variables explain citations in at least two thirds of the rolling windows (stock liquidity and industry citation intensity). Performing the analysis across the entire sample period (without rolling windows) selects these same two variables and three additional variables (CEO centrality, R&D stock, and analyst following), which taken together explain over 90% of the explanatory power of the all-variable specification. As we get similar results across different regularization procedures, it suggests these findings stem from the data generating process rather than some peculiarity of a particular method.

After identifying the key variables, we next explore how the inferences of empirical studies may change if we include those key covariates, by looking at recent studies that propose new determinants of corporate innovation. Recent research suggests that a new managerial trait, CEOs with pilot licenses, is associated with corporate innovation. Using the same model specification as Sunder et al. (2017), we confirm the positive relation between pilot CEOs and patent activity.[6] After including the key covariates we identified from prior innovation literature, we find that the coefficient estimate on pilot CEOs is insignificantly different from zero. Using a similar research set-up, we confirm the positive relations they document between innovation and antitakeover devices (Chemmanur and Tian, 2017). After including the key covariates of innovation, we find

[6] We thank Stephen B. McKeon for providing us with the pilot CEO data.

that the coefficient estimate of antitakeover devices is no longer significantly different from zero. Yet, it is important to note that in other cases, adding the seven surviving innovation covariates strengthens the results about newly proposed innovation covariates (e.g., Mukherjee et al., 2017). That is, one potential benefit of including the key identified variables is improving the model fitness. Our empirical analysis reveals that the inference becomes statistically stronger by including the key variables.

Do we suggest that these studies not including the key variables are "wrong"? The answer is no. It clearly depends on the research question and the purpose of the study. For instance, suppose that CEO matters for innovation and the underlying mechanism is via risk tolerance. Since risk tolerance is hard to measure but presumably correlated with pilot CEOs, the interpretation of the pilot CEO study is that pilot CEOs have higher risk tolerance. However, pilot CEOs may also have better social connections which are also related to innovation. If the social connection measure is better captured than whether someone is a pilot, it clearly wins the horse race. This does not mean that risk tolerance is not related to innovation but simply due to poor correlation between pilot CEO and risk tolerance (conditioning on social connection). In this vein, our evidence suggests that the interpretation or inference of the focused covariate depends on the purpose of the research question and our proposed key variables offer some clue about the underlying mechanism of the effect.

In our last task, we explore whether commonly used alternative methods addressing omitted variable problem can mitigate the necessity of including the key variables. Specifically, as a parsimonious treatment to address the issue of omitted variables, studies often include industry or firm fixed-effects in the analysis to account for differences in patenting choices.[7] This leads us to

[7] Many R&D firms choose not to file patents. A common argument for using industry fixed effects is to mitigate the potential bias of including R&D firms without patent activity.

explore whether these key economic determinants provide additional explanatory power after including industry or firm fixed effects. Our analysis indicates that the seven surviving covariates of innovation typically retain substantial explanatory power after including industry fixed effects, or industry-year pairwise fixed effects, or firm fixed effects. Adding both firm and industry-year pairwise fixed effects, however, only leaves firm size as significant.[8]

Empirical research often focuses on providing causal marginal effects and relying on exogenous shocks is a popular approach to identify these causal relationships. This method arguably mitigates omitted variable concern in a more robust manner. However, this approach relies on satisfying the exclusion restriction, which requires the shock to be uncorrelated with other covariates of innovation, even when the shock is truly exogenous. Consequently, assessing a typical instrumental variable or natural experiment hinges on understanding whether the shock only influences innovation through the proposed dependent variable of interest or whether it also influences previously identified innovation covariates. Unfortunately, it is unclear which variables to use to evaluate the exclusion restriction in studies of corporate innovation. To assess the potential use of these variables in evaluating the exclusions restriction in patent-based studies, we replicate a study of institutional ownership and corporate innovation. The identification strategy evaluated relies on the inclusion of firms in the S&P 500. Our evidence shows that the S&P 500 shock also influences one of the previously identified covariates of innovation, namely that of stock liquidity. Consequently, the effect of the instrumental variable on innovation occurs beyond its impact on institutional ownership, indicating the exclusion condition is not satisfied.[9] This does

[8] In many studies, controlling for fixed effects may not be practical or meaningful, especially when the independent variable is sticky and lacks time-series variations (e.g., Bertoni and Tykvova, 2015). Consequently, including both firm and industry-year pairwise fixed effects is rare in practice.

[9] We confirm the positive relationship between institutional ownership and corporate innovation using the specification of Aghion et al. (2013). In further tests, we find that after the inclusion of prior innovation covariates, the coefficient estimates of institutional ownership and S&P 500 inclusion are negative in this IV analysis.

not imply that including the key variables offers a test of the exclusion condition. However, they do provide a better baseline for the analysis than random or ad hoc conditioning variable selection.

This study makes several important contributions to the escalating literature on the determinants of corporate innovation. We posit that applying on a data-driven approach for model selection adds rigor to the process, especially when there is no recognized theory to guide these conditioning variable choices. Our perusal of studies on corporate innovation using patents as the output metric indicates that this literature uses a wide variety of potential conditional variables. Seldom do any two of these studies include the same previously identified covariates of innovation, regardless of whether they focus on the quantity or the productivity of corporate innovation. More importantly, it is unclear from this research which previously identified factors should indeed be included as control variables in studies of corporate innovation.

Our analysis using machine learning technique reveals that a small set of previously identified covariates for innovation provide independent explanatory power on the quantity, productivity, and quality of corporate innovation. Specifically, our analysis suggests conditioning innovation studies by using:

Innovation Quantity:	*Stock liquidity, firm size, industry citation intensity, CEO centrality, analyst following, and industry patent intensity*
Innovation Productivity:	*Stock liquidity, R&D stock, firm size, analyst following, CEO centrality, and both industry citation and patent intensity*
Innovation Quality:	*Stock liquidity, industry citation intensity, CEO centrality, and analyst following*

Remarkably, relying on industry or firm fixed effects does not invalidate the need to control for the key determinants of patent activity in assessing any newly proposed determinants of

innovation. In this context, our analysis provides empirical support to the argument that machine learning methods and natural experiments provide complementary approaches in financial economics research.

Our next contribution is to show that without considering these key identified variables, the inference of the studies on innovation can be drastically different. Our study offers a framework or starting point for researchers to rethink the inference of their newly propose innovation covariates, by considering these key variables. Furthermore, instead of randomly choosing what omitted variables to include, we propose a small set of variables that are based on previous economic studies on innovation.

Innovation studies often rely on exogenous shocks to provide causal evidence. However, this approach depends on the applicability of the exclusion restriction. Studies seeking to provide causal evidence focus on an exogenous variation in some variable that influences the treatment group yet does not directly impact the outcome variable or influence it through other covariates. Identifying the potential covariates to analyze in testing the exclusion restriction is challenging; there is limited guidance in the innovation literature. Against this backdrop, our analysis provides some preliminary guidance on the covariates to analyze in testing the exclusion restriction. More specifically, our analysis suggests using the seven key explanatory variables to evaluate whether a potential shock influences previously identified covariates.

2. Data, Sample, and Variables

2.1. Data Sources and Sample

To capture the thirty-five previously identified determinants of innovation, we obtain data from multiple sources and our main sample is a cross-section of different databases. More

specifically, we use ExecuComp to capture compensation information about managers. We complement it with BoardEx data with information about other CEO characteristics such as age, gender, and centrality. To capture firm characteristics, we rely on the Compustat and The Center for Research in Security Prices (CRSP) databases. We acquire corporate governance practices of firms from the Investor Responsibility Research Center (IRRC) Risk Metrics. We obtain family firm status from Ron Anderson's website,[10] state marginal tax from the Department of Labor, and board-related information from BoardEx. We collect institutional ownership information from Thomson Reuters and analyst following information from the Institutional Brokers' Estimate System (I/B/E/S). The industry characteristics are based on Compustat and CRSP information. Finally, we obtain patent and citation information from the United States Patent and Trademark Office (USPTO) (Hall, 1990). We drop the financial (SIC 6000-6799) and utility (SIC 4900-4949) industries. Our main sample spans the years from 2001 to 2010 with 2,716 firm-year observations of 410 unique firms with patents. In parallel to the main sample which is the most restrictive due to data availability, we also present results using a larger sample from 1992 to 2010 of 5,955 observations of 832 unique firms.[11]

2.2. Variable Definitions

2.2.1. Dependent Variables

We use two commonly used metrics for innovation output, patents and citations, as the dependent variable. The patents are based on the patent applications and we focus on the application year rather than the grant year as the application year is closer to the actual time of

[10] http://www.ronandersonprofessionalpage.net/data-sets.html.

[11] The sample size in our main sample is largely restrained by data availability on CEO compensation and family firm data. In later sections, we loosen this data availability restriction to evaluate the candidate covariates across time and firms. Our first analysis, extends the sample to the 1990s and expands to 25,985 observations with 2,217 unique firms covering 24 variables. Our second analysis, extends the sample to include additional firms and gives 58,671 observations with 7,502 unique firms covering 19 variables.

innovation (Griliches, Pakes, and Hall, 1988). Specifically, we use the log of patents and log of (1 + citations), and our base tests only include firms with patents. In later tests we include non-patenting R&D firms, which allows us to incorporate firms with zero patents to provide insights on the determinants of the patenting choice.

2.2.2. Right-Hand-Side Variables

We include thirty-five potential determinants of corporate innovation from prior research that we classify into four categories. In the first group, we include seven variables of managerial characteristics: (log) CEO age, CEO gender, CEO total compensation, CEO delta, CEO vega, CEO confidence, and CEO centrality. The second group contains firm characteristics including firm size, R&D stock, Tobin's q, stock liquidity, the firm's headquarters' distance to nearest USPTO office, tangibility, a dummy variable indicating if the firm is in manufacturing, return on assets (ROA), sales growth, organizational capital, capital structure, and state marginal tax rate. In the third category, we include corporate governance variables, including six antitakeover provisions (staggered board, poison pill, golden parachutes, limits to shareholder bylaw amendments, and supermajority requirements for mergers and charter amendments), board size, board independence, institutional ownership, a blockholder dummy, analyst following, and family firm designation. Finally, we include industry characteristics in the fourth category, namely the industry patent intensity, industry citation intensity, industry average R&D, industry competition, and industry size. We provide detailed definitions of the variables in Appendix C.

2.3. Sample Summary Statistics

Table 1 presents the summary statistics of the variables for the main sample where we list the variables by categories. We note that these summary statistics are largely comparable to previous studies. First, we show that the average (log) patent is roughly 3, equivalent to 21 patents

on average. Turning to CEO characteristics, we find that the average total pay is roughly $4.4 million while the pay-stock-price sensitivity is approximately $80,000 for a $1 dollar increase in stock price. The average vega is equivalent to a $34,000 pay increase for a 1% increase in volatility. On average, a CEO is 55 years old and only 2% of the CEOs are female. Forty-five percent of them are classified as confidence CEOs. Next, we find that the average firm size is roughly $2.1 billion. On average, R&D Stock (accumulated over the prior 10 years with 15% amortization) is about 36.4% of total assets. The average Tobin's q is 1.27. Stock liquidity shows that on average 320 million shares are traded during the year. On average, firms are 1,200 miles from the nearest USPTO office. Roughly 19% of total assets are net property, plant & equipment. The average return on assets (ROA) is 12.8% and 64% of firms are in manufacturing industries (standard industrial codes (SICs) 3 and 4). The average sales growth rate is 25.4%. 17.3% of total assets are long-term debt.

Turning to governance factors, we observe that on average the firms have six analysts following the firm and 37% of the common equity is owned by institutional investors. Approximately 65% of firms have poison pills and 77% of firms have golden parachutes. Roughly 18% of the firms are family controlled. The average board has 12 members, and 80% of the board members are independent. There is a staggered board system in 55% of the firms; 22% and 48% of the firms have merger and charter amendment limitation clauses, and a bylaw amendment limit, respectively. Finally, we observe that the average industry patent intensity is 0.19, and the average industry citation intensity is 0.02. On average, the industry R&D is 9.1% of total assets and the industry total assets is roughly $2.1 trillion. In Panel B, we include firms with reported R&D, i.e., including firms with no patents but with non-missing R&D. We find similar characteristics in general.

3. Identifying Key Variables: A Machine Learning Approach

3.1. Machine Learning Method in Variable Selection: Adaptive Lasso

He and Tian (2017) indicate that in recent years corporate innovation research has drastically increased. Not surprisingly, multiple managerial, firm, and industrial characteristics have been identified as being associated with a firm's innovation performance. For instance, Aghion et al. (2013) document the effect of institutional investors, and Galasso and Simcoe (2011) posit that CEO confidence is positively associated with innovation. So far, more than thirty factors have been studied and found to be significantly related to innovation. We apply a "horse race" approach to the multiple factors that have been identified as important determinants of corporate innovation.

The traditional variable selection approach relies on stepwise regressions as it proves to be computationally tractable relative to all subset regressions. Ideally, one might select the best fit model via certain statistical criteria of model fitness (e.g., Akaike Information Criteria or Bayesian Information Criteria), among all the possible combinations of the variables (in our case, 2^{35} = 34,359,738,368). Stepwise regression overcomes the infeasibility of best subset model selection by drastically lowering the complexity or number of combinations to assess (there are 630 potential models in our case). However, as convenient as it is, stepwise regression has several shortcomings. For instance, the process only focuses on a subset of the potential models among the possible combinations of the thirty-five factors because the outcome is contingent upon the sequence of the variables in the regression. Specifically, when a stepwise process either drops or adds a variable one at a time, the sequence of the variables becomes important - and there is no clear treatment on this issue. Last but not least, stepwise procedure lacks validation of the outcome, i.e., there is no out-of-sample verification in the process. A benefit of these traditional methods, all subset or stepwise, is that they directly penalize the model for including additional coefficients.

Due to these shortcomings in traditional parametric approaches, we use a machine learning method, specifically the Adaptive Lasso procedure as our main approach. Machine learning methods divide the sample into training and validation subsamples. Adaptive Lasso (the least absolute shrinkage and selection operator) provides a strong and robust inference because it explicitly considers the "predictive" power of the selection outcome via a cross-validation process. Furthermore, Adaptive Lasso gives a differential weight for penalizing different coefficients instead of applying a common penalty factor to all coefficients. We apply Adaptive Lasso to our data using a ten-fold cross-validation and choosing the two tuning parameters (lambda and gamma) to minimize the mean square error in the out-of-sample testing (Hui et al., 2015). Using these ten random subsets, we fit the model on nine of the subsets and then test the model on the excluded or validation set. We repeat this approach for each of the excluded sets and select the model with the best out-of-sample performance across all ten subsets. In Appendix B, we provide a more detailed explanation of the Adaptive Lasso method (typically labeled as an L_1 penalty) as well as Group Lasso and Elastic Net methods.

For our main tests we apply Adaptive Lasso with a rolling window approach. Specifically, we apply the procedure for a two-year rolling window and show the frequency of each variable chosen among all the rolling windows across the entire sample period.[12] We have a ten-year sample period so we have nine rolling windows.[13] After the selection of variables via Adaptive Lasso, we then rank all the variables by their individual explanatory power. Finally, we show the incremental explanatory power loss when we drop the variables in reverse order from the least

[12] In tabulated results, we also try using a 1-year window and we find similar outcomes. Industry factors become weaker because they are the average value of firms so that single year window dramatically decreases the cross-sectional variation.

[13] As we use a rolling window approach the optimal penalty factor changes from sample to sample. To facilitate replicability of our analysis, we have based our tests on the average optimal gamma of four across our rolling windows. We find that using sample-specific gammas yields similar outcomes to the average optimal gamma across our rolling windows. In addition, we also present results using the entire sample without rolling windows.

important to the most important variable from the full specification. We present results using all four methods (Adaptive Lasso, Group Lasso, Elastic Lasso, and stepwise regression). The machine learning methods, Adaptive Lasso, Group Lasso, and Elastic Net regression, split the data into ten random subsets for testing, but the stepwise regressions rely on in-sample information criteria for variable selection.

3.2. Key Variables Identified

3.2.1. Patents as the Innovation Outcome Measure

Table 2 shows the variable selection results when we use patent as the output metric for innovation. In columns 1-4 we focus on innovation output quantity, and consequently we do not include R&D stock, the innovation input factor, as one of the candidate variables. We use a two-year rolling window approach and check the frequency with which each variable is chosen among the nine rolling windows. Specifically, for instance, in column 1 we find that CEO centrality is chosen for eight out of nine rolling windows but CEO age is never selected. In sum, the results in column 1 indicate that the Adaptive Lasso procedure yields five variables that are associated with patents in at least two thirds of the rolling windows. Specifically, we find that CEO centrality, firm size, stock liquidity, industry citation intensity, and industry patent intensity are chosen. More importantly, the Adaptive Lasso procedure applies the cross-validation process to yield results that have strong predictive power rather than only model fitness.

In column 2 we show the incremental explanatory power loss of each variable. Generally speaking, and not surprisingly, we show that the incremental explanatory power loss of the unchosen variables is rather small. On the other hand, the five variables identified together provide roughly 89% of the explanatory power of the full specification. Overall, we conclude that most of the variables are not significantly associated with patents after controlling for the five variables.

In columns 3-4 we repeat the same process with the larger sample, which spans years 1992 to 2010. We drop four variables due to data availability restrictions (CEO centrality, board size, board independence, and family firm). We find the same two overlapping variables as in the column 1 results, i.e., firm size and stock liquidity are key variables for patent output. Overall, columns 1-4 results indicate that five variables survive the horse race, namely firm size, stock liquidity, CEO centrality, industry patent intensity, and industry citation intensity.

Column 5 shows the results when we include R&D stock. Compared to the column 1 results, we find that two new variables are selected, i.e., R&D stock and analyst following, while industry patent intensity is dropped, suggesting that six variables are important factors for innovation productivity. In column 7, using the larger sample we find that the same four variables (firm size, stock liquidity, R&D stock, industry patent intensity) are selected while analyst following and industry citation intensity become weaker. Surprisingly, we find that stock liquidity provides the strongest explanatory power, consistent with the notion that access to the capital market is of crucial importance to corporate innovation.

3.2.2. Citations as the Innovation Quality Metric

Table 3 shows the results when we focus on citations as another commonly used metric for innovation outcome. Again, columns 1-4 present results without R&D stock and in columns 5-8 R&D stock is included. In column 1, the main sample result shows that only two variables are identified as key variables, namely stock liquidity and industry citation intensity, while firm size is chosen in five out of nine rolling windows. Together the two key variables provide 75% of the explanatory power of the full specification. Column 3 shows that when we use the larger sample, firm size is chosen in addition to the same two variables. Together, these three variables provide 86% explanatory power compared to the full specification. In column 5 we include R&D stock

and we repeat the process. We find that the same two variables are chosen as in column 1, namely stock liquidity and industry citation intensity. Column 7 with the larger sample shows the same findings as column 3 as firm size is chosen besides stock liquidity and industry citation intensity. Interestingly, R&D stock is not chosen as an important factor for patent citations. In short, Adaptive Lasso spots three variables that are significantly related to citations for at least two thirds of the rolling windows, namely, firm size, stock liquidity, and industry citation intensity. All in all, taken together, the findings so far show that only a handful factors are related to patents and citations while each of them provides significant incremental explanatory power.

3.2.3. Robustness Evidence with Alternative Methods

We present variable selection results using multiple alternative methods in Table 4. First, rather than using rolling windows we present the results using the full sample as a whole. Again, we use a ten-fold cross-validation approach in our Adaptive Lasso, Group Lasso, and Elastic Net analyses. Specifically, in column 1 the results show that the same seven variables are chosen when we apply Adaptive Lasso to the entire sample. Figure 1 illustrates these findings visually and shows the order of variable selection in terms of each variable's relative explanatory power. Each line represents a variable, and the farther right the variable is, the more influential the variable. The vertical line shows the threshold at which the variables are selected or retained. The figure shows again that there are seven variables (plus a year dummy variable) that are selected by the process. Figure 2 shows the fraction of deviance (similar to R^2) explained by the seven variables. The figure shows that roughly 55% of model deviance is explained by the seven variables while the full specification explains about 58%. In column 2 of Table 4 we use the Group Lasso methodology and we allow individual variables being selected within the group. We find the same six key variables identified in the Adaptive Lasso analysis and three additional variables, namely vega,

institutional ownership, and industry size.[14] In column 3 we use an Elastic Net procedure to further address whether the results are driven by the fundamentals or by a specific machine learning specification. We find that the results show the same set of nine covariates providing significant explanatory power for innovation.

While the traditional all subset variable selection method remains unfeasible with thirty-five potential covariates, the commonly used alternative centers on stepwise regressions. The backward stepwise procedure starts with all the variables and takes an elimination approach; that is, if the variable incurs the least amount of model fitness loss, the variable is dropped. The process is repeated until no more variables can be deleted without a statistically significant loss of model fitness. This variable selection approach is often labeled as "greedy" because of concerns with overfitting. To further corroborate the machine learning results, we use a backward stepwise regression and show the results in column 4. The order of the variables is determined by the individual R^2 as we include each variable one at a time, separately. For instance, the last variable is CEO gender, as it has the lowest R^2 in a regression when it is included as the only explanatory variable. The results show that eight variables are chosen via stepwise regressions.[15] In sum, we find that across the four methods, seven variables, i.e., CEO centrality, firm size, R&D stock, stock liquidity, analyst following, industry patent intensity, and citation intensity matter for patents, are identified as the key variables. This outcome is largely similar to what we obtain with the rolling-window approach in Table 2.

[14] Recall that in the two-year rolling window analysis, we find that these three variables are not selected. This highlights that among a group of highly correlated variables (such as in a group), Adaptive Lasso tends to choose one variable and ignore the others.

[15] Internet Appendix Table IA1 provides the results of the stepwise regressions with rolling windows. The results show that four variables are identified as a key variable. The differences between full-sample and rolling-window results imply the instability of stepwise regressions compared to machine learning methodologies.

Turning to columns 5-8 in which we assess citations, we observe that across the four methodologies, six variables are identified as the key variables, namely, CEO centrality, firm size, R&D stock, stock liquidity, analyst following, and industry citation intensity.[16] No other CEO characteristics are chosen and none of the corporate governance factors are identified. In conclusion, we find that the identification of key variables is robust across four different methodological approaches and they yield similar sets of key variables.

3.2.4. Correlation between Variables

A potential concern in using machine learning methods centers on the correlation among the 35 proposed variables. Our main approach for assessing this issue is the group lasso analysis. Yet, explicitly investigating the correlation among the various variables provides another layer of robustness. The correlation matrix is provided in Table IA2. Casual observation suggests that most the correlations are rather low. More specifically, the average absolute value of correlation is 0.11 and the median is 0.07. The bottom (upper) quartile is 0.03 (0.13) and the 5th (95th) percentile is 0.01 (0.34). To further assess the consistency of our main results, we repeat the analysis after excluding the variables with highest correlations. For instance, institutional ownership and analyst following have a correlation of 0.92. When we take out institutional ownership, the results remain the same. When we take out analyst following, the results also remain the same. We repeat this exercise with variables that exhibit correlations higher than 0.5 and we obtain the same findings. Coupled with the results in the group lasso analysis, we conclude that the correlations between the 35 variables do not appear to provide spurious findings that would invalidate their use as control variables.

3.2.5. Innovation Regimes

[16] Rolling window results in Internet Appendix Table IA1 yield three overlapping variables: firm size, stock liquidity, and industry citation intensity.

Our selection procedure is applied to a sample that is restricted by data availability. How does the sampling process influence the outcome? Do the same variables show up across different innovation regimes? This question is important because it sheds light on the usefulness of the variables for future research. In Table 5, we use two different ways to check on this issue.

First, in Panel A, we show the results using Adaptive Lasso but with different samples. In Panel A, column 1, we drop the CEO characteristics to loosen the sample restriction from the ExecuComp database. The sample enlarges to 25,985 observations. Again, we find that except for industry size, the same set of key variables are selected by the Adaptive Lasso procedure, namely firm size, R&D stock, analyst following, stock liquidity, and industry patent intensity. In column 2, we further drop the antitakeover provisions from the IRRC database, thus expanding the sample to 58,671 observations. Once again, we find the same five key variables are identified by the Adaptive Lasso procedure. In columns 3 and 4, we show that key variables for the citation test are also robust to different samples. In Panel B, we show the results after focusing on an earlier time, 1990-2000, that does not overlap with the main test sample time frame. We find fairly similar results across the full-time period and within various subsets of the data.

3.2.6. Firms Without Patents

A common approach to dealing with the firms without patents is to include them in the sample while denoting the patent as zero. Internet Appendix Table IA3 provides the results; we repeat the variable selection process, but we include non-patenting firms and denote their patents and citations as zero. We apply the Adaptive Lasso technique and find that eight variables are chosen for the patent test and three variables are selected for the citation test. These results are almost identical to our main findings, suggesting that again the key variables are robust to different sampling treatment.

3.2.7. Using R&D to Capture Innovation

Recent corporate innovation research often focuses on the output measures for innovation, i.e., patents and citations. Still, R&D expenditure offers another metric as an important input for innovation. In Internet Appendix Table IA4, we apply the variable selection process using R&D as the measure of innovation. Interestingly, we find a much smaller set of factors that are important in explaining R&D spending. In contrast to the seven variables for patents, we find only four factors are chosen in at least two thirds of the rolling windows, namely firm size, stock liquidity, tangibility, and industry patent intensity, regardless of the procedure we use. Taken together these results suggest that R&D spending and patent citations share similar explanatory variables, while patents exhibit a larger set of covariates.

3.2.8. A Horse-Race Between Variables

In Internet Appendix Table IA5 we show the results of two specifications. The first specification only includes each of the thirty-five previously identified determinants one at a time separately along with year fixed effects in columns 1 and 2. The second specification includes all the variables simultaneously in columns 3 and 4. In columns 1 and 2, not surprisingly, we find that most of the thirty-five factors are significantly related to corporate innovation. In columns 3 and 4 we find that 12 variables are significant at the 5% level or better for patents and 9 variables are significant at the 5% level or better for citations. These findings provide a baseline for the identification of key variables as they lack predication power via cross-validation. The key variables identified are indeed a subset of these surviving factors.

4. An Reexamination of Inferences of Previous Studies

4.1. Three Studies Revisited

In this section, we assess the importance and usefulness of these key covariates in recent studies of corporate innovation. Our purpose is to demonstrate the influence of omitting the key identified variables on previous studies that examine innovation determinants. We do not intend to invalidate the findings of the previous studies. Instead, we focus on how the inference or the interpretation of the results could be viewed differently.

The first example is a recent study shows the importance of a managerial trait, CEO with pilot license, in corporate innovation. Sunder et al. (2017) suggest that firms with pilot CEOs exhibit more successful innovation and show that their companies generate more patents than comparable firms with non-pilot CEOs. In Table 6, Panel A, column 1, we repeat their original specification, documenting that pilot CEOs are positively associated with patents (t-statistics > 1.97) but that it adds limited explanatory power to the model (accounting for 0.002 of R^2).[17] In column 2, we add the key explanatory variables that we have identified from the original specification to assess the importance of these key variables when including industry fixed effects.[18] The results of this test reveal that pilot CEOs are not significantly related to innovation (t-statistics < 1.18). In contrast, the key variables add substantially to the model fit in this analysis, increasing adjusted R^2 by 0.072 (13.5%) relative to the baseline analysis in column 1. In column 3, we drop R&D stock from the key variables; we find that pilot CEO remains insignificant (t-statistics < 1.62). In sum, we find that this newly documented trait, CEOs with a pilot license, is not associated with innovation quantity or productivity after controlling for previously identified determinants of corporate innovation.

[17] Our pilot data may not be identical to that in Sunder et al. (2017) as they were unable to provide their data for verification purposes. Stephen McKeon kindly supplied the CEO pilot data used in their study of CEO pilot risk-taking (Cain and McKeon, 2016). The descriptions of the data collection processes in both studies appear similar but could still differ in the underlying data obtained.

[18] We do not include CEO centrality to minimize influence on sampling.

The second recent study we examine is presented in Table 6 Panel B that uses firm fixed effects to study antitakeover devices and corporate innovation. More specifically, Chemmanur and Tian (2017) show that a firm's corporate governance strength, via antitakeover provisions, is associated with future innovation output. In Panel B, column 1, we replicate their specification using future patents as the dependent variable and incorporating firm fixed effects. Similar to their result, we find that antitakeover provisions are positively related to the firm's patents (t-statistics > 2.37). In column 2, we document that after including analyst following, R&D stock, stock liquidity, and industry patent and citation intensity, the antitakeover effect becomes insignificant (t-statistics < 1.06). In column 3, without R&D stock, we still find that the antitakeover effect is insignificant. We also note that except for stock liquidity, the other five key variables are significant even with firm fixed effects.

In a recent study, Mukherjee et al. (2017) rely on staggered changes in state-level corporate tax rates as an identification strategy and show that a tax rate increase is negatively associated with future patents. In Table 6 Panel C, column 1, using the original specification, we show that the dummy variable indicating state tax increase is negatively associated with the change in patents. In column 2 we include the key variables, and we show that the effect in column 1 remains robust. As a matter of fact, we find that the magnitude of the effect remains similar while the statistical significance increases, suggesting that including the key variables improve model fit.

4.2. Discussions

It is imperative that we discuss these findings. Maybe not surprisingly, the first two studies above show evidence consistent with the conventional wisdom about omitted variable problem, that is, omitted variables often result in weaker significance in the focused covariate. Does that mean that these studies are wrong? The answer is "not really". Clearly, it depends on the purpose

of the research question. For instance, the pilot CEO study is consistent with the notion that risk tolerance is related to corporate innovation. However, it becomes insignificant after including the key variables. The explanation could be that one or more of the key variables capture risk tolerance better than pilot CEO. As such, it does not mean that pilot CEO is not related to corporate innovation. It does, however, change the inference of pilot CEO on corporate innovation, after the key variables are included as controls. In sum, whether or not to include the key variables as controls depends on the nature of specific research question. Nevertheless, it is important to consider these key controls because at a minimum, they could change the inference or the interpretation of the newly proposed covariate for innovation. Furthermore, the third study indicates that because the key variables identified provide significant incremental explanatory power, one benefit of including them is the improvement of model fitness, which results in a lower model's standard error. Overall, we suggest that including the key identified variables could provide significant benefits for new research on corporate innovation.

5. Fixed Effects and Exclusion Condition Revisited

To the extent that the purpose of the majority of empirical studies is to understand causal marginal effects rather than claiming some best predicative model, the control variables are to isolate alternative explanations for the covariate to help adjudicate the validity of the proposed hypothesis in the context of the research question. Two common approaches to achieve that goal are including fixed effects in the regression analysis or relying on exogenous shock to the covariate to establish identification. We explore the usefulness of our key identified variables in the context of these two methods.

We first assess whether controlling for industry and/or firm fixed effects makes the key variables identified above redundant. Studies in innovation often include industry fixed-effects to address two empirical issues. One effect is the non-random distribution of firms with and without patents across different industries. In other words, in some industries many of their member firms choose to apply for patents while in other industries the opposite is the common practice. Including industry fixed effects is often argued as being a means to mitigate such self-selection bias. Yet another issue is that of omitted industry characteristics. Industry fixed effects only address this problem if the omitted industry factors are time-invariant. Likewise, firm-level fixed effects are commonly used in innovation studies to address firm-level time-invariant omitted variable concerns.[19]

In Table 7 we present the results with the seven key identified variables and a mix of industry and firm fixed effects. In Panel A, we show that in column 1, the base case, we only include the year fixed effects, the same as in the variable selection process. Not surprisingly, all the key variables remain significant. In column 2 we add the industry fixed effects, and the results show that except for industry citation patent intensity, other key variables remain significant. This suggests that controlling for industry fixed effects does not make the key variables redundant. In column 3 we control for firm fixed effects. The results suggest that stock liquidity and industry patent intensity are still significant. Prior studies emphasize that liquidity proxies for ease of financing, with cross-sectional tests revealing that innovation is increasing in liquidity. Strikingly, when adding a firm fixed effect and no longer differentiating across firms, the sign on the liquidity coefficient estimate changes from positive to negative. This change in the sign arises because

[19] Even though fixed effects are meant to mitigate industry or firm-level time-invariant variables and our key variables are time-variant, the necessity of controlling for the key variables is still relevant due to the correlations between the time-invariant fixed effects and the key variables.

liquidity here is measured contemporaneously with innovation. Presumably, studies focusing on changes in liquidity and innovation would likely use lagged liquidity instead.

In column 4 we use industry-year pairwise fixed effects, which is more restrictive than the industry fixed effects shown in column 2. We show that, similar to the findings in column 2, five out of seven variables (two industry variables are dropped automatically) still survive the fixed effects. Finally, in column 5, we show that adding firm fixed effects on the top of industry-year fixed effects leaves only firm size significant. Overall, these results suggest that the commonly used fixed effects approach does not rule out the necessity of using the key variables. However, the specification of all three types of fixed effects (as in column 5) may not be feasible in practice, especially when the focused variable has low time-series variations (e.g., board size, corporate bylaw, antitakeover practice).

In Panel B, the citation test based on four key variables yields similar inference as the patent test. That is, controlling for industry or firm fixed effects is not sufficient to provide the explanatory power of if the key variables are omitted. Notably, in columns 4 and 5 even after controlling for industry-year pairwise and/or firm fixed effects, stock liquidity is still highly significant, suggesting that it provides important incremental explanatory power, which is not captured by time-invariant industry or firm factors.

A common and more prevalent approach to allow for causal inferences about corporate innovation is the use of instrument variable with an exogenous event. In their study, Aghion et al. (2013) document a positive relationship between institutional holdings and corporate innovation. Furthermore, by using S&P 500 membership as an exogenous event that increases institutional ownership in the firm, and is thus a valid instrument variable for institutional ownership, they show that S&P 500 inclusion is followed by increased innovation output.

In Panel A of Table 8, we replicate the instrument variable approach in the original study in which the authors use S&P 500 membership as the instrument variable for institutional ownership. We first show the original results in columns 1-3. Column 1 shows the baseline Poisson regression results. Columns 2 and 3 show the first stage and second stage instrument variable regression results. In column 4, we repeat column 2 with three additional key variables (the other key variables are already in the original specifications). First, we find that the S&P 500 variable is still positive and highly significant at 1%, suggesting that it is a strong instrument for institutional ownership. Second, in column 5, we show the second stage IV regression results with the key variables. We find that the institutional ownership effect flips its sign and becomes both negative and significant, indicating that the omitted key variables have significant influence on IV regression outcome.

We further explore the validity of the instrument variable. To be a valid instrument variable, the candidate variable must satisfy both relevance and exclusion condition.[20] In this case, the results suggest that the exclusion condition is likely to be violated because, presumably, the S&P 500 index inclusion changes a firm's stock liquidity, which prior research indicates is strongly related to innovation. If that is indeed the case, the effect of the instrument variable on innovation is also manifested via other channel besides institutional ownership, which invalidates the instrument variable. We show in Table 8, Panel B that across different windows over the S&P 500 index shock there is a clear change in stock liquidity among the treatment firms. The evidence suggests that the instrument variable does not satisfy exclusion condition.

[20] Roberts and Whited (2012) suggests that corporate finance research tends to emphasize the relevance of the instrument but often gives limited attention to the exclusion restriction. One impediment in testing the exclusion restriction is the identification of the relevant potential covariates to consider in the evaluation.

6. Concluding Remarks

In this study, we accomplish three tasks. First, using innovation research as a platform, we apply machine learning methodology to identify key variables that provide independent explanatory power among 35 previously identified determinants. Both parametric and non-parametric variable selection methods identify a few key variables that provide incremental explanatory power for patent activity. Notably, six of the seven variables identified in the Adaptive Lasso, Group Lasso, Elastic Net, and the stepwise analysis are the same. Undertaking a similar analysis for citations, we document that four previously identified covariates represent a tractable set of relevant of control variables. Fundamental firm and industry characteristics, such as firm size, R&D stock, stock liquidity, and industry innovation intensity, provide independent explanatory power for patents and their citations. Most surprisingly, we find that stock liquidity provides by far the strongest explanatory power among the variables, consistent with the notion that access to capital market is of crucial importance to corporate innovation.

Second, besides the seemingly-random choice of control variables across disciplines in innovation research, we show how the inferences of previous studies may change after including the key identified variables. We do not advocate that new research should *always* condition their empirical analysis on the key explanatory variables for corporate innovation. Nor do we propose that our key variables are the only set of variables to consider. Whether to include certain control variables depends on the nature of the research question. Still, our analysis offers two benefits along this line. One is that these key identified variables help with the interpretation of the underlying mechanism of the newly proposed covariate. In addition, these variables provide a starting point for studies seeking to model the selection process when developing a first stage model for the self-selection of firms' patenting choices.

Lastly, our study shows that the popular methods of using fixed effects to treat omitted variable bias does not invalidate the need to include these key determinants of patent activity in assessing newly proposed determinants of innovation. Furthermore, we demonstrate that even exogenous shock approach does not mitigate the omitted variable problem in studies providing causal evidence on corporate innovation.

References

Abadie, Alberto, and Maximilian Kasy, 2019. Choosing among regularized estimators in empirical economics: The risk of machine learning. *Review of Economics and Statistics* 101 (5), 743-762.

Acharya, Viral V., Ramin P. Baghai, and Krishnamurthy V. Subramanian, 2014. Wrongful discharge laws and innovation. *Review of Financial Studies* 27 (1), 301-346.

Aghion, Philippe, John Van Reenen, and Luigi Zingales, 2013. Innovation and institutional ownership. *American Economic Review* 103 (1), 277-304.

Anderson, Ronald C., David M. Reeb, and Wanli Zhao, 2012. Family-controlled firms and informed trading: Evidence from short sales. *Journal of Finance* 67 (1), 351-385.

Atanassov, Julian, 2013. Do hostile takeovers stifle innovation? Evidence from antitakeover legislation and corporate patenting. *Journal of Finance* 68 (3), 1097-1131.

Athey, Susan, 2018. The impact of machine learning on economics, Stanford working paper.

Athey, Susan, and Guido Imbens, 2017. The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives* 31 (2), 3-32.

Balsmeier, Benjamin, Lee Fleming, and Gustavo Manso, 2017. Independent boards and innovation. *Journal of Financial Economics* 123 (3), 536-557.

Bernstein, Shai, 2015. Does going public affect innovation? *Journal of Finance* 70 (4), 1365-1403.

Bertoni, Fabio, and Tereza Tykova, 2015. Does governmental venture capital spur invention and innovation. *Research Policy* 44 (4), 925-935.

Brown, James, Steven Fazzari, and Bruce Petersen, 2009. Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *Journal of Finance* 64 (1), 151-185.

Caner, Mehmet, and Qinliang Fan, 2015. Hyrbrid generalized empirical likelihood estimators: Instrument selection with adaptive lasso. *Journal of Econometrics* 187 (1), 256-274.

Cain, Mathew, and Stephen McKeon, 2016. CEO personal risk-taking and corporate policies. *Journal of Financial and Quantitative Analysis* 51 (1), 139-164.

Chemmanur, Thomas J., Elena Loutskina, and Xuan Tian, 2014. Corporate venture capital, value creation, and innovation. *Review of Financial Studies* 27 (8), 2434-2473.

Einav, Liran, and Jonathon Levin, 2014. Economics in the age of big data. *Science* 346, 1243089.

Eisfeldt, Andrea L., and Dimitris Papanikolaou, 2013. Organization capital and the cross-section of expected returns. *Journal of Finance* 68 (4), 1365-1406.

Faleye, Olubunmi, Tunde Kovacs, and Anand Venkateswaran, 2014. Do better-connected CEOs innovate more? *Journal of Financial and Quantitative Analysis* 49 (5-6), 1201-1225.

Fang, Vivian W., Xuan Tian, and Sheri Tice, 2014. Does stock liquidity enhance or impede firm innovation? *Journal of Finance* 69 (5), 2085-2125.

Galasso, Alberto, and Timothy S. Simcoe, 2011. CEO overconfidence and innovation. *Management Science* 57 (8), 1469-1484.

Gompers, Paul A., Joy L. Ishii, and Andrew Metrick, 2003. Corporate governance and equity prices. *Quarterly Journal of Economics* 118 (1), 107-155.

Griliches, Zvi, Ariel Pakes, and Bronwyn H. Hall, 1988. The value of patents as indicator of inventive activity. NBER Working Paper No. 2083.

Hall, Bronwyn, 1990. The manufacturing sector master file: 1959-1987. NBER Working Paper No. 3366.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, 2001. The NBER patent citations data file: Lessons, insights, and methodological tools. NBER Working Paper No. 8498.

He, Jie (Jack), and Xuan Tian, 2013. The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics* 109 (3), 856-878.

He, Jie (Jack), and Xuan Tian, 2017. Finance and corporate innovation: A survey. Available at SSRN: https://ssrn.com/abstract=3062863.

Hui, Francis K. C., David I. Warton, and Scott D. Foster, 2015. Tuning parameter selection for the adaptive Lasso using ERIC. *Journal of the American Statistical Association* 110 (509), 262-269.

Im, Hyun Joon, and Janhoon Shon, 2019. The effect of technological imitation on corporate innovation: Evidence from US patent data, *Research Policy* 48 (9), 103802.

Intintoli, Vincent, Kathleen Kahle, and Wanli Zhao, 2018. Director connectedness: Monitoring efficacy and career prospects. *Journal of Financial and Quantitatively Analysis* 53 (1), 65-108.

Kaplan, Steven N., and Luigi Zingales, 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics* 112 (1), 169-215.

Koh, Ping-Sheng, David Reeb, Elvira Sojli, and Wing Wah Tham, 2016. Measuring innovation around the world. HKUST Institutional Repository.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017. Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics* 132 (2), 665-712.

Lerner, Josh, and Amit Seru, 2017. The use and misuse of patent data: Issues for corporate finance and beyond. Available at SSRN: https://ssrn.com/abstract=3071750

Malmendier, Ulrike, and Geoffrey A. Tate, 2005. CEO overconfidence and corporate investment. *Journal of Finance* 60 (6), 2661-2700.

Mukherjee, Abhiroop, Manpreet Singh, and Alminas Zaldokas, 2017. Do corporate taxes hinder innovation? *Journal of Financial Economics* 124 (1), 195-221.

Mulliainathan, Sendhil, and Jann Spiess, 2017. Machine learning: An applied econometric approach. *Journal of Economic Perspectives* 31(2), 87-106.

Roberts, Michael R., and Toni M. Whited, 2012. Endogeneity in empirical corporate finance. Available at SSRN: https://ssrn.com/abstract=1748604.

Sunder, Jayanthi, Shyam V. Sunder, and Jingjing Zhang, 2017. Pilot CEOs and corporate innovation. *Journal of Financial Economics* 123 (1), 209-224.

Taddy, Matt, 2017. One-step estimator paths for concave regularization. *Journal of Computational and Graphical Statistics* 26 (3), 525-536.

Zou, Hui, 2006. The adaptive Lasso and its oracle properties. *Journal of American Statistical Association* 101 (476), 1418-1429.

Appendix A: Typical Control Variable Selection

Control Variables	Acharya et al. 2014; RFS	Chemmanur et al. 2014; RFS	Sunder et al. 2017; JFE	Balsmeier et al. 2017; JFE	Atanassov 2013; JF	Fang et al. 2014; JF	Common Usage
Firm Size	**Significant**	**Significant**	**Significant**	**Significant**	**Significant**	**Significant**	6/6
Tangible Assets	Absent	Insignificant	Insignificant	Absent	Insignificant	**Significant**	4/6
Stock Returns	Absent	Absent	Insignificant	Absent	Absent	Absent	1/6
Growth	Insignificant	**Significant**	**Significant**	**Significant**	Absent	Insignificant	5/6
Institutional Ownership	Absent	Absent	**Significant**	Absent	Absent	Absent	1/6
CEO Tenure	Absent	Absent	Insignificant	Absent	Absent	Absent	1/6
CEO Delta	Absent	Absent	Insignificant	Absent	Absent	Absent	1/6
CEO Vega	Absent	Absent	**Significant**	Absent	Absent	Absent	1/6
Military CEO	Absent	Absent	Insignificant	Absent	Absent	Absent	1/6
R&D Spending	**Significant**	**Significant**	Absent	**Significant**	Absent	**Significant**	4/6
Firm Age	Absent	Insignificant	Absent	**Significant**	**Significant**	**Significant**	4/6
Capital Expenditures	Absent	Insignificant	Absent	Absent	**Significant**	**Significant**	3/6
Board Size	Absent	Absent	Absent	Insignificant	Absent	Absent	1/6
Capital Structure	Absent	**Significant**	Absent	**Significant**	Absent	Absent	2/6
Competition	**Significant**	Insignificant	Absent	Absent	Insignificant	Insignificant	4/6
Value added	Insignificant	Absent	Absent	Absent	Absent	Absent	1/6
Colleges	Insignificant	Absent	Absent	Absent	Absent	Absent	1/6
Enrollment	Insignificant	Absent	Absent	Absent	Absent	Absent	1/6
State GDP	Insignificant	Absent	Absent	Absent	Absent	Absent	1/6
Unemployment Insurance	Insignificant	Absent	Absent	Absent	Absent	Absent	1/6
Population	Insignificant	Absent	Absent	Absent	Absent	Absent	1/6
ROA	Absent	Insignificant	Absent	Absent	**Significant**	Insignificant	3/6
Financial Constraints	Absent	Insignificant	Absent	Absent	Absent	**Significant**	2/6
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	6/6
Firm Fixed Effects	Yes	No	No	Yes	Yes	Yes	4/6
Industry Fixed Effects	No	Yes	Yes	Yes	No	No	3/6
Reported R^2	0.178	0.330	0.497	0.207	0.847	0.839	-

Appendix B: Machine Learning and Variable Selection

Machine learning encompasses a variety of techniques for identifying patterns and relationships in the data, and it is commonly used in forecasting and to simplify model selection processes. In the realm of machine learning, our interest lies in finding an econometric model that maps the set of variables that potentially explain an output, in this case patents or citations. Among the multiple methods available, such as subset selection, least squares, generalized additive models, trees, support vector machines, bagging and boosting, Lasso regression offers a balanced trade-off between interpretability and flexibility. The more flexible the method, the lower its bias has, since it can better approximate the true relationship existing in the data. But increased flexibility increases the variance of the method, since it attempts to fit not only true data points but also the unavoidable noise present in the data set.

Lasso (Least Absolute Shrinkage Selection Operator) is a shrinkage method that reduces (or shrinks) the values of the coefficients to zero compared with ordinary least squares. The advantage of shrinkage methods is that the estimated model exhibits lower variance than those of least squares estimates. We compare Lasso with least squares estimation as follows:

$$\text{Least squares: } \frac{min}{\beta_0,\beta_j}\sum_{i=1}^{n}(y_i - \beta_0 - \sum_1^p \beta_j\, x_{ij})^2$$

$$\text{Lasso regression: } \frac{min}{\beta_0,\beta_j}\sum_{i=1}^{n}(y_i - \beta_0 - \sum_1^p \beta_j\, x_{ij})^2 \text{ subject to } ||\beta||_1 \leq t$$

where y is the vector of observations of the dependent variable, x denotes the independent variables, β are the corresponding coefficients, $||\ ||_1$ and $||\ ||_2$ are the L1 and L2 norms respectively, and t is a user-specified parameter. The Lagrange formulation of the Lasso regression is:

$$\text{Lasso regression: } \frac{min}{\beta_0,\beta_j}\sum_{i=1}^{n}(y_i - \beta_0 - \sum_1^p \beta_j\, x_{ij})^2 + \lambda||\beta||_1$$

The least squares estimation corresponds to an unconstrained minimization problem; the Lasso regression imposes a convex but non-smooth l_1 constraint. Least squares analysis rewards including as many covariates as possible, since additional right-hand side variables help to reduce the sum of the squares. However, Lasso regression imposes a penalty factor on the coefficients that helps to reduce the value of the coefficients or reduces the number of factors included in the model. As such, Lasso regressions appear well-suited to addressing the model selection challenge when developing forecast models that work well with out-of-sample data. In addition, the Lasso regression performs

both the variable selection and the parameter estimation simultaneously. Lasso exhibits both low variability and limited computational costs, especially in high dimensional problems.

Lasso regression solutions coincide with the least squares solution if the penalty parameter is set sufficiently small. In a Lasso regression, the penalty parameter controls both the size and the number of coefficients, with higher values leading to a lower number of covariates to be included in the linear model. This increases the flexibility of the model and reduces its variance but at the cost of a higher model bias. Lasso regression utilizes cross-validation, a resampling technique, to facilitate finding a parameter value that ensures a proper balance between bias and variance (or flexibility and interpretability), as the one that minimizes the estimated test error rate of the estimator. In cross-validation a subset of the data observations, the training set, is used to estimate (or train) the model; the remaining observations are held to serve as test set or validation set. The selected test sets serve to provide an estimate of the test error rate. Typically, the measure of the test error is the mean square error (MSE).

The K-fold cross-validation method divides the data set randomly into K different subsets, in which we set K = 10. Keeping one of the subsets as the validation set, the model is estimated over the remaining K-1 sets for a range of values of the penalty parameter. We repeat this process using each of the K subsets as a validation set, yielding K estimates of the MSE for each parameter value; its K-fold estimate is simply the average value of the K estimates. The best parameter value is the one yielding the lowest K-fold estimate, which we denote as λ-min in the Lagrange formula. This parameter estimate is the one-standard error rule parameter, λ-1se.

Adaptive Lasso has been developed from Lasso to address the issue that Lasso does not possess the oracle property. Note that an estimator that is consistent in variable selection is not necessarily consistent in parameter estimation. An oracle estimator must be consistent in both. Adaptive Lasso performs a different regularization for each coefficient, adjusting the penalty factor differently for each coefficient, avoiding overfitting by penalizing large coefficients. Consequently, Adaptive Lasso penalizes more of the coefficients with lower initial estimates. This adjustment, as compared to Lasso, helps to achieve the oracle property of the estimators (Zou, 2006).

We also use two variants of Lasso, Group Lasso and Elastic Net regression. Group Lasso takes into consideration that the variables within predetermined categories are meant to be selected or unselected together. However, for our purpose, we loosen the “togetherness” restriction on

selections at the group level while we use a methodology that selects at both the group and individual variable levels. In addition, Lasso's penalty function is a linear combination of the coefficients, which tends to select one variable from a group and ignore the others if there is a group of highly correlated variables. In order to mitigate this concern, we also use the Elastic Net method, which adds a quadratic part to the L1 penalty function that used alone is the ridge regression penalty form. In other words, the elastic net is another regularized regression method that linearly combines the L1 and L2 penalties of the lasso and ridge methods.

In contrast to these machine learning approaches, traditional statistical methods for variable selection focus on either the best subset or stepwise regressions. The best subset approach chooses the single best model from all possible combinations of the potential covariates. The best model is chosen based on some sort of pre-specified information criteria, which typically penalizes the number of non-zero parameters. Two potential issues are the computational difficulty with large p and that it relies on within-sample analysis, leading to concerns of fitting noise in the data. Stepwise regressions follow the same general approach as best subset selection but limit the number of models that need to be analyzed. For instance, with 25 potential covariates, best subset analysis requires estimating 33,554,432 regressions, while stepwise regression only requires 325 regressions. Stepwise variable selection is computationally feasible, but the results are sensitive to the sequencing of the variables. As best subset selection and stepwise variable selection typically rely on fitting the data within the sample, these methods are typically labeled as "greedy" selection methods.

Appendix C: Variable Definitions

Analyst Following: the number of financial analysts that follow the firm during the year;

Antitakeover: the corporate governance index developed in Gompers et al. (2005);

Blockholder: a dummy variable that equals 1 if at least one institutional investor holds more than 5% of the common equity;

Board Size: the log of the number of directors;

Board Independence: the proportion of independent directors;

Bylaw Amendments Limit: a dummy variable indicating whether the firm has a policy limiting shareholders' ability through majority vote to amend the corporate bylaws;

Capital Structure: the long-term debt divided by total assets;

CEO Age: the log of CEO age;

CEO Centrality: a factor score based on the factor analysis of four metrics of the CEO's social connectedness. The data is drawn from BoardEx; we calculate the four metrics following Intintoli et al. (2018);

CEO Confidence: a dummy variable equals to 1 if the CEO's in-the-money exercisable options exhibit greater than 67 percent moneyness and it happens twice during the sample period (Malmendier and Tate, 2005);

CEO Delta: measures CEO wealth change in dollars to one percent change in stock price;

CEO Gender: a dummy variable indicating whether the CEO is female;

CEO Total Pay: the log of CEO annual total compensation (tdc1);

CEO Vega: measures CEO wealth change in dollars to one annualized standard deviation of stock return;

Citation: the log of 1 plus the number of citations the firm obtains during the year;

CITES: number of a firm's patents weighted by the number of future citations;

Distance to USPTO: the log of distance in miles between the firm headquarters and the USPTO office that oversees the firm's state;

Family Firm: a dummy variable indicating whether the firm is a family firm as defined in Anderson et al. (2012);

Financial Distress: the firm's KZ index (Kaplan and Zingales, 1997), calculated as -1.002 × Cash flow + 0.28 × Tobin's q + 3.18 × Leverage - 39.368 × Dividends - 1.315 × Cash holdings;

Firm Age: the log number of years that the firm appears in Compustat;

Golden Parachutes: a dummy variable indicting whether the firm has a severance agreement that provides benefits to management/board members in the event of firing, demotion, or resignation following a change in corporate control;

Industry Patent Intensity: the total industry patents divided by total industry assets; the industry is SIC two-digit;

Industry Citation Intensity: the total industry citations divided by total industry assets; the industry is a SIC two-digit code;

Industry R&D: the average R&D for each two-digit SIC industry;

Industry Competition (HHI): the Herfindahl index based on sales for each two-digit SIC industry;

Industry Size: log of total assets of the firms in each two-digit SIC industry;

Institutional Ownership: institutional ownership of a common equity;

Manufacturing: a dummy variable equals to 1 if the firm is in one-digit SIC code 3 or 4;

Market-to-book: the market value of an equity divided by the book value of equity;

Mergers and Charter Amendments: a dummy variable indicating if the firm has a provision limiting shareholders' ability through majority vote to amend the corporate charter or requires more than a majority of shareholders to approve a merger;

Organizational Capital: the stock of Selling, General & Administrative Expense (SG&A) as in Eisfeldt and Papanikolaou (2013);

Patent: the log of 1 plus the number of patents applied;

Patenting: a dummy variable indicating the firm has patent application during the year;

Poison Pill: a shareholder right that is triggered in the event of an unauthorized change in control that typically renders the target company financially unattractive or dilutes the voting power of the acquirer;

PPE/EMP: the net property, plant, and equipment value divided by the number of employees;

R&D: R&D expenditures scaled by total assets;

R&D Stock: the cumulative R&D over the previous 10 years, assuming a depreciation rate of 15%;

ROA: income before extraordinary items divided by total assets;

S&P 500: a dummy variable equal to 1 if the firm is included in the S&P 500 index during the year;

Sales Growth: the average growth rate of sales over the prior three years;

Size: the log of total assets;

Staggered Board: a dummy variable indicating if the board is staggered, i.e., the directors are divided into separate classes with each class being elected to overlapping terms;

State tax: the marginal tax rate of the state;

Stock Liquidity: the log of stock daily trading volume aggregated over the year;

Tangibility: net property, plant and equipment scaled by total assets;

Tobin's q: the market value of equity plus book value of debt scaled by book value of total assets;

Figure 1: Adaptive Lasso

This figure shows how the log (lambda) corresponds to the number of non-zero variables. It shows that seven variables are chosen to the right of the vertical line of optimal log (lambda).

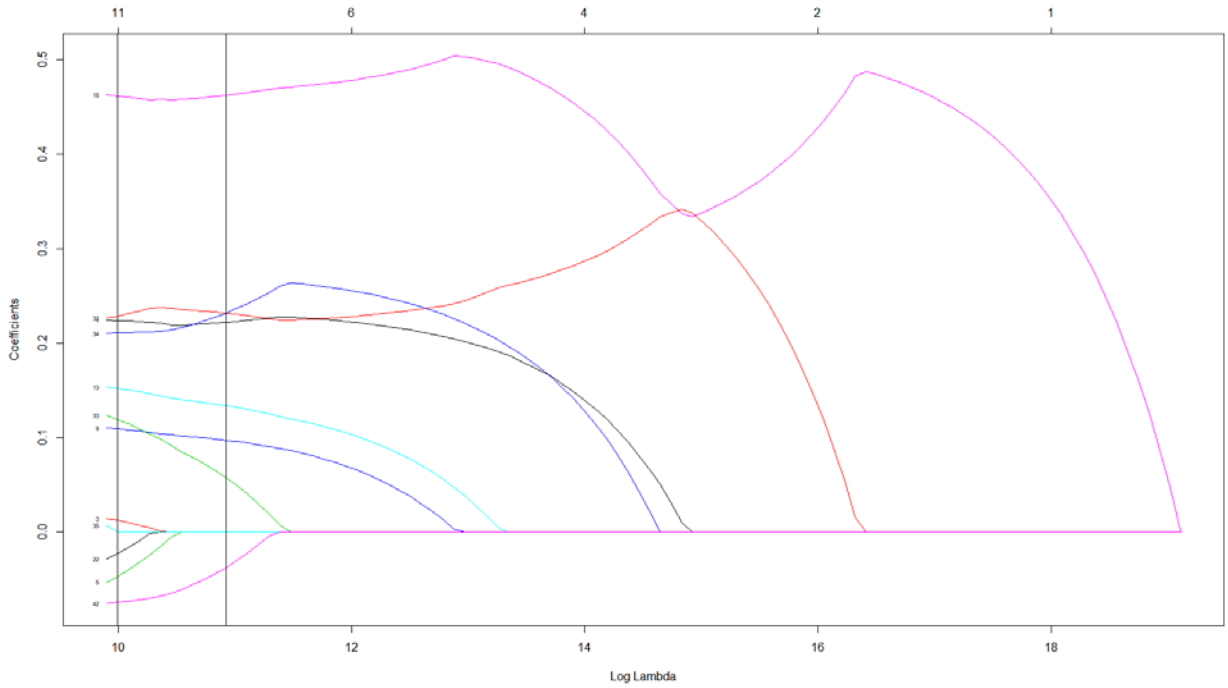

Figure 2: Fraction of Deviance Explained

This figure shows the fraction of model deviance explained by the number of variables. It shows that the 7 variables chosen roughly explains 55% of the deviance while the whole model explains 58%.

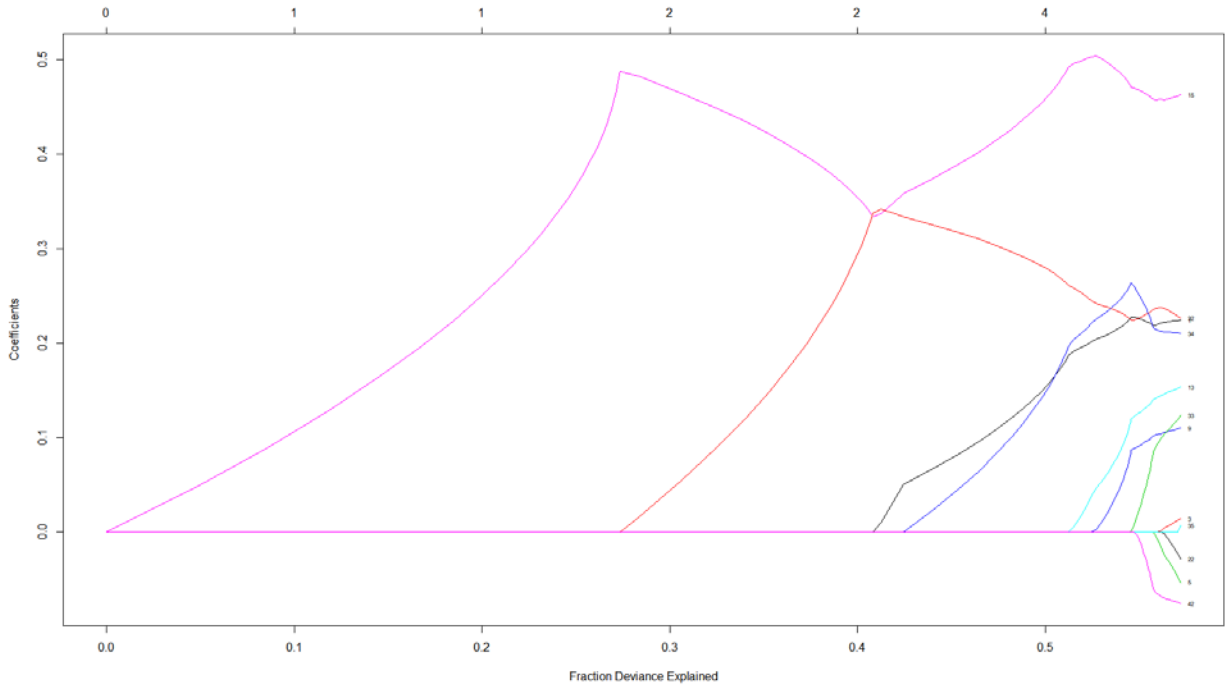

Table 1 Summary Statistics and Univariate Analysis

This table shows the summary statistics of the sample firms with patents with firm-year observations from 2001 to 2010. All variables are defined in Appendix C.

	Mean	Median	Std. Dev.	Lower Quartile	Upper Quartile
Innovation Characteristics:					
Patent	3.041	2.833	1.598	1.792	4.060
Citation	0.513	0.000	0.900	0.000	0.693
CEO Characteristics:					
CEO Total Pay	8.381	8.441	1.031	7.725	9.098
CEO Delta	4.390	4.370	1.513	3.435	5.397
CEO Vega	3.518	3.596	1.554	2.551	4.590
CEO Centrality	0.326	-0.042	1.305	-0.334	0.445
CEO Age	4.003	4.007	0.121	3.932	4.094
CEO Gender	0.019	0.000	0.137	0.000	0.000
CEO Confidence	0.450	0.000	0.498	0.000	1.000
Firm Characteristics:					
Size	7.644	7.561	1.491	6.594	8.788
R&D Stock	0.364	0.262	0.328	0.110	0.555
Tobin's q	1.273	0.879	1.467	0.494	1.508
Stock Liquidity	19.587	19.499	1.540	18.486	20.679
Distance to USPTO	7.081	7.283	1.071	6.462	7.970
Tangibility	0.193	0.161	0.134	0.095	0.259
State Tax	0.294	0.338	0.093	0.298	0.348
ROA	0.128	0.129	0.092	0.087	0.177
Manufacturing	0.637	1.000	0.481	0.000	1.000
Sales Growth	0.254	0.189	0.700	-0.013	0.417
Organization Capital	-0.023	-0.047	0.130	-0.066	-0.018
Capital Structure	0.173	0.155	0.158	0.030	0.264
Governance Characteristics:					
Analyst Following	1.813	1.386	1.872	0.000	3.738
Institutional Ownership	0.373	0.340	0.382	0.000	0.757
Poison Pill	0.654	1.000	0.476	0.000	1.000
Golden Parachutes	0.769	1.000	0.422	1.000	1.000
Family Firm	0.175	0.000	0.380	0.000	0.000
Blockholder	0.905	1.000	0.293	1.000	1.000
Board Size	2.496	2.398	0.566	2.079	2.708
Mergers and Charter Amendments	0.217	0.000	0.412	0.000	0.000
Staggered Board	0.552	1.000	0.497	0.000	1.000
Board Independence	0.793	0.818	0.118	0.750	0.889
Bylaw Amendments Limit	0.476	0.000	0.500	0.000	1.000
Industry Characteristics:					
Industry Patent Intensity	0.190	0.222	0.073	0.129	0.243
Industry Citation Intensity	0.020	0.017	0.013	0.010	0.030
Industry Competition	0.099	0.062	0.097	0.050	0.120
Industry R&D	0.091	0.097	0.058	0.042	0.114
Industry Size	14.555	14.694	0.928	14.268	15.185

Table 2 Patents: A Rolling Window Approach with Adaptive Lasso

This table presents the variable selection results on multiple managerial, governance, firm, and industry factors on patents among patenting firms. We apply the Adaptive Lasso process on a two-year rolling window basis. All variables are defined in Appendix C. The incremental R^2 loss is computed by comparing OLS regressions after dropping each variable in the order suggested by the Adaptive Lasso variable selection. "-" denotes variables that are not included in the selection process due to design or data availability.

Sample:	*2001-2010*		*1992-2010*		*2001-2010*		*1992-2010*	
	Frequency Chosen	*Incremental R^2 Loss*	*Frequency Chosen*	*Incremental R^2 Loss*	*Frequency Chosen*	*Incremental R^2 Loss*	*Frequency Chosen*	*Incremental R^2 Loss*
	Without R&D Stock				*With R&D Stock*			
Variables	**Patent**							
Managerial								
CEO	**8/9**	**0.0196**	-	-	**9/9**	**0.0098**	-	-
CEO	0/9	0.0079	1/18	0.0030	1/9	0.0059	0/18	0.0003
CEO Vega	1/9	0.0011	4/18	0.0034	3/9	0.0010	0/18	0.0030
CEO Delta	1/9	0.0014	1/18	0.0032	1/9	0.0031	0/18	0.0012
CEO Total	1/9	0.0001	0/18	0.0011	1/9	0.0001	1/18	0.0017
CEO Age	0/9	0.0000	0/18	0.0022	0/9	0.0000	0/18	0.0006
CEO Gender	0/9	0.0001	0/18	0.0001	0/9	0.0001	0/18	0.0002
Firm								
Size	**9/9**	**0.0646**	**18/18**	**0.1011**	**9/9**	**0.0629**	**18/18**	**0.1011**
R&D Stock	-	-	-	-	**9/9**	**0.0580**	**15/18**	**0.0361**
Tobin's q	0/9	0.0015	0/18	0.0000	0/9	0.0000	0/18	0.0007
Stock	**9/9**	**0.3696**	**17/18**	**0.3320**	**9/9**	**0.3696**	**14/18**	**0.3320**
Distance to	1/9	0.0073	1/18	0.0052	1/9	0.0073	1/18	0.0024
Tangibility	0/9	0.0005	0/18	0.0004	0/9	0.0003	1/18	0.0024
Manufacturing	1/9	0.0009	1/18	0.0011	4/9	0.0015	4/18	0.0069
State Tax	0/9	0.0000	0/18	0.0000	0/9	0.0000	1/18	0.0006
ROA	0/9	0.0001	2/18	0.0000	0/9	0.0001	1/18	0.0009
Capital	0/9	0.0010	0/18	0.0002	0/9	0.0010	0/18	0.0000
Sales Growth	1/9	0.0001	1/18	0.0016	2/9	0.0002	0/18	0.0013
Organizational	0/9	0.0012	0/18	0.0006	1/9	0.0000	0/18	0.0000
Corporate								
Analyst	5/9	0.0042	8/18	0.0041	**6/9**	**0.0039**	9/18	0.0084
Poison Pill	0/9	0.0011	1/18	0.0023	1/9	0.0012	2/18	0.0009
Blockholder	0/9	0.0021	1/18	0.0029	0/9	0.0029	1/18	0.0016
Institutional	0/9	0.0020	1/18	0.0022	0/9	0.0020	0/18	0.0015
Board Size	0/9	0.0000	-	-	1/9	0.0000	-	-
Golden	0/9	0.0007	0/18	0.0004	0/9	0.0002	0/18	0.0019
Board	0/9	0.0008	-	-	0/9	0.0002	-	-
Mergers and	0/9	0.0008	0/18	0.0005	0/9	0.0008	0/18	0.0013
Staggered	0/9	0.0000	0/18	0.0002	0/9	0.0000	0/18	0.0003
Bylaw	0/9	0.0001	2/18	0.0008	0/9	0.0001	1/18	0.0001
Family Firm	0/9	0.0022	-	-	0/9	0.0023	-	-
Industry								
Industry	**8/9**	**0.0580**	9/18	0.0212	**9/9**	**0.0346**	9/18	0.0147
Industry	**6/9**	**0.0201**	11/18	0.0267	3/9	0.0180	**12/18**	**0.0267**
Industry Size	1/9	0.0000	5/18	0.0035	3/9	0.0008	6/18	0.0022
Industry R&D	1/9	0.0001	12/18	0.0065	2/9	0.0001	6/18	0.0010
Industry	0/9	0.0009	0/18	0.0011	0/9	0.0013	0/18	0.0001

Table 3 Citation: A Rolling Window Approach with Adaptive Lasso

This table presents the variable selection results on multiple managerial, governance, firm, and industry factors on patent citations among patenting firms. We apply the Adaptive Lasso process on a two-year rolling window basis. All variables are defined in Appendix C. The incremental R^2 loss is computed by comparing OLS regressions after dropping each variable in the order suggested by the Adaptive Lasso variable selection. "-" denotes variables that are not included in the selection process due to design or data availability.

Sample:	*2001-2010*		*1992-2010*		*2001-2010*		*1992-2010*	
	Frequency Chosen	*Incremental R^2 Loss*	*Frequency Chosen*	*Incremental R^2 Loss*	*Frequency Chosen*	*Incremental R^2 Loss*	*Frequency Chosen*	*Incremental R^2 Loss*
	Without R&D Stock				*With R&D Stock*			
Variables	**Citation**							
Managerial								
CEO	3/9	0.0301	-	-	3/9	0.0258	-	-
CEO	0/9	0.0040	1/18	0.0037	0/9	0.0037	0/18	0.0026
CEO Vega	1/9	0.0003	0/18	0.0004	0/9	0.0003	0/18	0.0001
CEO Delta	0/9	0.0008	0/18	0.0004	0/9	0.0004	2/18	0.0003
CEO Total	0/9	0.0009	0/18	0.0017	0/9	0.0009	0/18	0.0013
CEO Age	0/9	0.0018	0/18	0.0010	0/9	0.0011	0/18	0.0003
CEO Gender	0/9	0.0003	0/18	0.0003	0/9	0.0004	1/18	0.0013
Firm								
Size	5/9	0.0264	**14/18**	**0.0270**	5/9	0.0264	**16/18**	**0.0270**
R&D Stock	-	-	-	-	2/9	0.0120	4/18	0.0090
Tobin's q	0/9	0.0031	0/18	0.0036	0/9	0.0013	1/18	0.0010
Stock	**9/9**	**0.2797**	**15/18**	**0.2394**	**9/9**	**0.2797**	**15/18**	**0.2394**
Distance to	0/9	0.0044	0/18	0.0046	0/9	0.0006	0/18	0.0005
Tangibility	0/9	0.0001	0/18	0.0002	0/9	0.0001	0/18	0.0001
Manufacturing	0/9	0.0022	0/18	0.0025	0/9	0.0019	2/18	0.0011
State Tax	0/9	0.0002	0/18	0.0004	0/9	0.0002	0/18	0.0000
ROA	0/9	0.0015	0/18	0.0015	0/9	0.0004	0/18	0.0018
Capital	0/9	0.0009	0/18	0.0010	0/9	0.0009	0/18	0.0006
Sales Growth	0/9	0.0001	0/18	0.0001	0/9	0.0001	0/18	0.0010
Organizational	0/9	0.0011	0/18	0.0010	0/9	0.0010	0/18	0.0003
Corporate								
Analyst	0/9	0.0020	3/18	0.0029	0/9	0.0020	3/18	0.0017
Poison Pill	0/9	0.0028	0/18	0.0028	0/9	0.0033	2/18	0.0046
Blockholder	2/9	0.0019	1/18	0.0014	2/9	0.0018	1/18	0.0008
Institutional	0/9	0.0055	0/18	0.0075	0/9	0.0055	1/18	0.0082
Board Size	0/9	0.0000	-	-	0/9	0.0000	-	-
Golden	0/9	0.0001	0/18	0.0001	0/9	0.0001	0/18	0.0024
Board	0/9	0.0001	-	-	0/9	0.0001	-	-
Mergers and	0/9	0.0009	0/18	0.0010	0/9	0.0004	0/18	0.0007
Staggered	0/9	0.0007	0/18	0.0007	0/9	0.0005	1/18	0.0005
Bylaw	0/9	0.0012	0/18	0.0000	0/9	0.0015	2/18	0.0014
Family Firm	0/9	0.0000	-	-	0/9	0.0000	-	-
Industry								
Industry	**9/9**	**0.0350**	**12/18**	**0.0340**	**9/9**	**0.0350**	**14/18**	**0.0340**
Industry	1/9	0.0007	1/18	0.0005	1/9	0.0008	1/18	0.0094
Industry Size	0/9	0.0051	1/18	0.0058	0/9	0.0070	1/18	0.0083
Industry R&D	1/9	0.0016	3/18	0.0006	1/9	0.0019	4/18	0.0019
Industry	0/9	0.0019	0/18	0.0017	0/9	0.0000	0/18	0.0025

Table 4 Full Sample Results Using Alternative Methods

This table shows the Adaptive Lasso, Group Lasso, Elastic Net regression, and Stepwise results using the 2001-2010 sample without rolling windows. The variables are defined in Appendix C.

	Patent				***Citation***			
	Adaptive Lasso	*Group Lasso*	*Elastic Net*	*Stepwise*	*Adaptive Lasso*	*Group Lasso*	*Elastic Net*	*Stepwise*
Managerial Characteristics:								
CEO Centrality	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	No
CEO Confidence	No	No	No	No	No	No	No	No
CEO Vega	No	**Yes**	**Yes**	No	No	No	No	No
CEO Delta	No	No	No	No	No	No	No	No
CEO Total Pay	No	No	No	No	No	No	No	No
CEO Age	No	No	No	No	No	No	No	No
CEO Gender	No	No	No	No	No	No	No	No
Firm Characteristics:								
Size	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
R&D Stock	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Tobin's q	No	No	No	No	No	No	No	No
Stock Liquidity	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Distance to USPTO	No	No	No	**Yes**	No	No	No	No
Tangibility	No	No	No	No	No	No	No	No
Manufacturing	No	No	No	No	No	No	No	No
State Tax	No	No	No	No	No	No	No	No
ROA	No	No	No	No	No	No	No	No
Capital Structure	No	No	No	No	No	No	No	No
Sales Growth	No	No	No	No	No	No	No	No
Organizational Capital	No	No	No	No	No	No	No	No
Corporate Governance:								
Analyst Following	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	No
Poison Pill	No	No	No	No	No	No	No	No
Blockholder	No	No	No	No	No	No	No	No
Institutional Ownership	No	**Yes**	No	No	No	**Yes**	No	No
Board Size	No	No	No	No	No	No	No	No
Golden Parachutes	No	No	No	No	No	No	No	No
Board Independence	No	No	No	No	No	No	No	No
Mergers and Charter	No	No	No	No	No	No	No	No
Staggered Board	No	No	No	No	No	No	No	No
Bylaw Amendments Limit	No	No	No	No	No	No	No	No
Family Firm	No	No	No	No	No	No	No	No
Industry Characteristics:								
Industry Citation	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Industry Patent Intensity	**Yes**	No	**Yes**	**Yes**	No	**Yes**	No	No
Industry Size	No	**Yes**	**Yes**	No	No	**Yes**	No	No
Industry R&D	No	No	No	No	No	**Yes**	No	No
Industry Competition	No	No	No	No	No	No	No	No

Table 5 Variable Selection and Regime Stability

This table presents the results of different tests checking the regime stability of the selected variables. Panel A shows the variable selection outcome via Adaptive Lasso using different samples as we expand the sample size by data availability. Columns 1 and 3 drop CEO characteristics. Columns 2 and 4 drop antitakeover provisions. Panel B shows the variable selection outcome with Adaptive Lasso when we check the sub-period of 1990-2000.

Panel A: Variable Selection Across Regimes

	(1)	*(2)*	*(3)*	*(4)*
Sample:	*1990-2010*			
	Patent		**Citation**	
Variable Chosen:	*Firm Size* *R&D Stock* *Analyst Following* *Stock Liquidity* *Industry Size* *Industry Patent*	*Firm Size* *R&D Stock* *Analyst Following* *Stock Liquidity* *Industry Patent Intensity*	*Firm Size* *R&D Stock* *Analyst Following* *Stock Liquidity* *Industry Citation Intensity*	*Firm Size* *Analyst Following* *Stock Liquidity* *Industry Citation Intensity*
Observations	25,985	58,671	25,985	58,671

Panel B: Variable Selection for 1990-2000

	(1)	*(2)*	*(3)*	*(4)*
Sample:	*1990-2000*			
	Patent		**Citation**	
Variable Chosen:	*Firm Size* *R&D Stock* *Analyst Following* *Stock Liquidity* *Industry Size* *Industry Patent*	*Firm Size* *R&D Stock* *Analyst Following* *Stock Liquidity* *Industry Patent Intensity*	*Firm Size* *R&D Stock* *Analyst Following* *Stock Liquidity* *Industry Citation Intensity*	*Firm Size* *Analyst Following* *Stock Liquidity* *Industry Citation Intensity*
Observations	13,024	30,208	13,024	30,208

Table 6 Evaluating Previous Studies

This table compares the results of including and excluding the key identified variables. We replicate the results in Sunder et al. (2017) in Panel A, column 1 (2/3) without (with) the key identified variables. In Panel B, we replicate Chemmanur and Tian's (2017) work without (with) key identified variables in column 1 (2/3). All variables are defined in Appendix C. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

Panel A: Pilot CEOs and Innovation

	(1)	(2)	(3)
	Without Additional Controls	*With Additional Controls*	
Dependent Variable:	**Patent**		
Constant	2.990***	3.132***	3.307***
	(6.41)	(7.26)	(8.40)
Pilot CEO	**0.350****	**0.198**	**0.275**
	(1.97)	**(1.17)**	**(1.61)**
Size	1.003***	0.935***	0.696***
	(13.46)	(12.49)	(9.11)
Log(PPE/EMP)	0.315***	0.231***	0.196***
	(4.99)	(4.12)	(3.37)
Stock Return	0.046**	0.045**	0.083***
	(2.21)	(2.37)	(4.16)
Tobin's q	-0.091**	-0.104**	-0.181***
	(-2.05)	(-2.56)	(-4.32)
Institutional Ownership	0.054	-0.126	-0.200*
	(1.13)	(-1.24)	(-1.93)
CEO Tenure	0.008	0.014	0.018
	(0.25)	(0.49)	(0.60)
Delta	-0.018	-0.024	-0.042
	(-0.40)	(-0.54)	(-0.88)
Vega	0.204***	0.106**	0.131**
	(3.59)	(2.00)	(2.28)
CEO Age	-0.010	0.000	0.002
	(-0.24)	(0.01)	(0.05)
CEO Confidence	-0.036	-0.060	-0.042
	(-0.51)	(-0.90)	(-0.60)
Stock Liquidity	-	0.339***	0.523***
		(5.34)	(7.74)
Analyst Following	-	0.233**	0.319***
		(2.12)	(2.84)
R&D Stock	-	0.486***	-
		(9.16)	
Industry Patent Intensity	-	0.543***	0.620***
		(4.75)	(5.44)
Industry Citation Intensity	-	0.058	0.025
		(1.07)	(0.46)
Industry and Year Fixed Effects	*Yes*	*Yes*	Yes
Observations	4,426	4,426	4,426
Adjusted R^2	0.532	0.604	0.573

Panel B: Firm Fixed Effects

	(1)	(2)	(3)
	Without Additional Controls	*With Additional Controls*	
Dependent Variable:		**Patent$_{t+1}$**	
Constant	-0.908***	-1.496***	-1.136***
	(-3.55)	(-4.73)	(-5.73)
Antitakeover	**0.096****	**0.072**	**0.069**
	(2.37)	**(1.05)**	**(1.00)**
Size	0.269***	0.250***	0.271***
	(8.18)	(7.63)	(8.12)
ROA	-0.303**	-0.179	-0.199
	(-2.05)	(-1.26)	(-1.46)
Leverage	-0.214***	-0.188**	-0.231***
	(-2.82)	(-2.50)	(-2.61)
HHI	-3.647***	-2.373***	-2.165***
	(-4.40)	(-2.99)	(-2.87)
HHI^2	7.546***	4.920***	5.105***
	(4.81)	(3.26)	(3.14)
R&D	0.605	-0.249	-0.286
	(1.05)	(-0.45)	(-0.58)
PPE/Assets	0.279**	0.306***	0.316***
	(2.44)	(2.77)	(2.66)
CAPX/Assets	-0.138	-0.092	-0.100
	(-0.90)	(-0.59)	(-0.31)
Tobin's q	-0.006	-0.007	-0.009
	(-1.00)	(-1.28)	(-1.04)
Financial Distress	-0.004	-0.008	-0.007
	(-0.30)	(-0.58)	(-0.59)
Institutional Ownership	-0.177***	-0.237***	-0.220***
	(-2.96)	(-3.67)	(-3.70)
Analyst Following	-	0.022*	0.019*
		(1.75)	(1.75)
R&D Stock	-	0.663***	-
		(3.02)	
Stock Liquidity	-	0.016	0.019
		(1.33)	(1.43)
Industry Patent Intensity	-	3.890***	3.945***
		(4.52)	(4.49)
Industry Citation Intensity	-	6.281**	5.837**
		(2.26)	(2.49)
Firm Fixed Effects	*Yes*	*Yes*	*Yes*
Year Fixed Effects	*Yes*	*Yes*	*Yes*
Observations	19,274	19,274	19,274
Adjusted R^2	0.915	0.917	0.909

Panel C: State Tax and Innovation

This table presents the results of the effect of a tax rate change on patents. We replicate the study of Mukherjee et al. (2017). All the variables are defined in Appendix C.

	(1)	**(2)**
	Without Additional Controls	*With Additional Controls*
Dependent Variable:	**ΔLog(1+#Patent)**	
Tax Increase	-0.027***	-0.029***
	(-2.66)	(-3.08)
Tax Decrease	0.007	0.007
	(0.76)	(0.81)
Controls	Yes	Yes
Year Fixed Effects	Yes	Yes
Observations	34,752	34,752

Table 7 Industry and Firm Fixed Effects

This table checks the robustness of the key factors that we identify when we include industry-level and firm-level fixed effects. OLS regression based on the main sample is applied. Statistical significance at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively. "-" denotes variables missed in the specification because they are absorbed by the fixed effect.

Panel A: Patents

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:			**Patent**		
Stock Liquidity	0.252***	0.218***	-0.092**	0.224***	-0.064
	(6.22)	(5.28)	(-2.36)	(5.07)	(-1.41)
Industry Citation Intensity	0.185***	0.011	0.025	-	-
	(4.95)	(0.30)	(0.90)		
Size	0.449***	0.516***	0.461***	0.513***	0.446***
	(9.38)	(10.18)	(6.85)	(9.50)	(6.36)
R&D Stock	0.213***	0.237***	0.069	0.238***	0.107
	(4.79)	(5.46)	(0.84)	(5.21)	(1.30)
Industry Patent Intensity	0.137***	0.172***	0.235***	-	-
	(3.10)	(3.20)	(5.09)		
CEO Centrality	0.154***	0.158***	0.034	0.157***	0.032
	(3.18)	(3.27)	(1.29)	(3.03)	(1.22)
Analyst Following	0.112***	0.118***	-0.014	0.121***	-0.018
	(2.97)	(3.02)	(-0.16)	(2.95)	(-0.19)
Firm Fixed Effects	*No*	*No*	*Yes*	*No*	*Yes*
Industry Fixed Effects	*No*	*Yes*	*No*	*No*	*No*
Year Fixed Effects	*Yes*	*Yes*	*Yes*	*No*	*No*
Industry-Year Fixed Effects	*No*	*No*	*No*	*Yes*	*Yes*
Observations	*2,716*	*2,716*	*2,716*	*2,716*	*2,716*
Adjusted R^2	*0.562*	*0.609*	*0.907*	*0.592*	*0.910*

Panel B: Citations

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:			**Citation**		
Stock Liquidity	0.311***	0.272***	-0.088	0.271***	-0.111*
	(6.17)	(5.35)	(-1.60)	(4.90)	(-1.66)
Industry Citation Intensity	0.244***	0.198***	0.242***	-	-
	(6.88)	(4.21)	(5.07)		
Size	0.304***	0.384***	0.298***	0.406***	0.332***
	(5.71)	(5.89)	(3.20)	(5.77)	(2.97)
R&D Stock	0.140***	0.157***	0.087	0.160***	0.071
	(3.21)	(3.54)	(0.92)	(3.30)	(0.73)
Firm Fixed Effects	*No*	*No*	*Yes*	*No*	*Yes*
Industry Fixed Effects	*No*	*Yes*	*No*	*No*	*No*
Year Fixed Effects	*Yes*	*Yes*	*Yes*	*No*	*No*
Industry-Year Fixed Effects	*No*	*No*	*No*	*Yes*	*Yes*
Observations	*2,716*	*2,716*	*2,716*	*2,716*	*2,716*
Adjusted R^2	*0.372*	*0.408*	*0.759*	*0.377*	*0.753*

Table 8 Exclusion Restriction Test

This table presents the results of examining the exclusion condition for the instrument variable in Aghion et al. (2013). We demonstrate these tests using the effect of institutional ownership on innovation after their inclusion in the S&P 500. All the variables are defined in Appendix C. Statistical significance at 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

Panel A: Instrument Variable Regression

	(1)	(2)	(3)	(4)	(5)
	Original Results			*With Additional Controls*	
	Poisson	**OLS (first-stage)**	**OLS (second-stage)**	**OLS (first-stage)**	**OLS (second-stage)**
Dependent Variable:	**CITES**	**Institutional Ownership**	**CITES**	**Institutional Ownership**	**CITES**
Institutional Ownership	**0.007*** **	-	**0.029** **	-	**-0.036***
	(2.97)		**(2.16)**		**(-1.73)**
S&P 500	-	8.872***	-	5.493***	-
		(3.77)		(2.68)	
Stock Liquidity	-	-	-	3.559***	0.391***
				(7.20)	(3.84)
Industry Patent Intensity	-	-	-	13.200	-0.288
				(0.45)	(-0.22)
Industry Citation Intensity	-	-	-	95.594	9.986
				(0.93)	(1.61)
Industry Fixed-Effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Year Fixed-Effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	6,208	6,208	6,208	6,208	6,208

Panel B: Exclusion Condition Check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1-year window			*2-year window*			*3-year window*		
	Pre	Post	*t-test*	Pre	Post	*t-test*	Pre	Post	*t-test*
Stock Liquidity	17.92	17.70	0.85	17.81	17.56	1.25	18.03	17.61	2.35**

Internet Appendix
Not for Publication

Table IA1 Variable Selection: A Rolling Window Approach via Stepwise Procedure

This table presents the results of multiple managerial, governance, firm, and industry factors on innovation among patenting firms. We apply the stepwise process on a two-year rolling window basis. All the variables are defined in Appendix C. The Huber-White Sandwich estimator is clustered at the firm level.

Sample:	*2001-2010*	*1992-2010*	*2001-2010*	*1992-2010*
	Frequency Chosen			
Variables:	**Patent**		**Citation**	
Managerial Characteristics:				
CEO Centrality	4/9	-	0/9	-
CEO Vega	2/9	2/18	1/9	2/18
CEO Confidence	0/9	3/18	1/9	2/18
CEO Total Pay	0/9	0/18	0/9	0/18
CEO Delta	2/9	1/18	0/9	1/18
CEO Age	0/9	0/18	0/9	0/18
CEO Gender	0/9	4/18	0/9	1/18
Firm Characteristics:				
Size	**9/9**	**18/18**	**6/9**	**18/18**
R&D Stock	**8/9**	**17/18**	3/9	9/18
Tobin's q	0/9	0/18	0/9	0/19
Stock Liquidity	**9/9**	**12/18**	**9/9**	**14/18**
Distance to USPTO	3/9	1/18	0/9	1/18
Tangibility	0/9	4/18	0/9	0/18
State Tax	0/9	2/18	0/9	1/18
ROA	0/9	3/18	0/9	2/18
Manufacturing	1/9	7/18	0/9	2/18
Sales Growth	1/9	1/18	0/9	0/18
Organizational Capital	0/9	0/18	0/9	0/18
Capital Structure	0/9	0/18	0/9	2/18
Corporate Governance:				
Analyst Following	4/9	6/18	0/9	2/18
Institutional Ownership	0/9	3/18	0/9	2/18
Poison Pill	2/9	3/18	2/9	5/18
Golden Parachutes	0/9	0/18	0/9	0/18
Family Firm	0/9	-	0/9	-
Blockholder	0/9	1/18	0/9	2/18
Board Size	0/9	-	0/9	-
Mergers and Charter Amendments	0/9	0/18	0/9	2/18
Staggered Board	0/9	0/18	0/9	3/18
Board Independence	0/9	-	0/9	-
Bylaw Amendments Limit	0/9	0/18	1/9	5/18
Industry Characteristics:				
Industry Citation Intensity	**8/9**	**10/18**	**9/9**	**13/18**
Industry Size	5/9	8/18	0/9	4/18
Industry R&D	4/9	10/18	0/9	4/18
Industry Patent Intensity	4/9	6/18	0/9	2/18
Industry Competition	0/9	0/18	0/9	0/18

Table IA2 Correlation Table

This table shows the correlation matrix between the 35 variables based on the sample of 2001-2010.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	CEO Total Pay	1.00																	
2	CEO Delta	0.42	1.00																
3	CEO Vega	0.46	0.78	1.00															
4	CEO Centrality	0.19	0.18	0.19	1.00														
5	CEO Age	-0.01	0.02	-0.01	-0.09	1.00													
6	CEO Gender	0.04	0.03	0.01	0.06	-0.08	1.00												
7	CEO Confidence	-0.03	0.17	0.08	-0.06	0.19	-0.08	1.00											
8	Size	0.62	0.43	0.42	0.22	0.06	0.03	-0.17	1.00										
9	R&D Stock	-0.11	-0.02	0.01	0.03	-0.13	0.00	0.08	-0.40	1.00									
10	Tobin's q	0.12	0.38	0.21	0.02	-0.06	0.02	0.22	-0.01	0.17	1.00								
11	Stock Liquidity	0.49	0.38	0.38	0.28	-0.12	0.05	-0.09	0.54	0.21	0.26	1.00							
12	Distance to USPTO	-0.06	-0.08	-0.09	-0.03	0.06	0.01	-0.04	0.07	-0.22	-0.11	-0.27	1.00						
13	Tangibility	-0.09	-0.13	-0.09	-0.08	0.12	-0.06	-0.13	0.11	-0.34	-0.27	-0.20	0.15	1.00					
14	State Tax	0.24	0.30	0.26	0.02	0.06	0.02	0.13	0.39	-0.34	0.17	-0.03	0.10	0.00	1.00				
15	ROA	0.23	0.30	0.23	0.02	0.02	-0.04	0.10	0.40	-0.29	0.30	0.06	0.10	0.17	0.55	1.00			
16	Manufacturing	-0.16	-0.13	-0.10	-0.12	0.04	-0.06	0.11	-0.16	0.04	-0.05	-0.08	-0.01	-0.17	-0.09	-0.12	1.00		
17	Sales Growth	-0.03	0.10	0.07	-0.02	-0.02	0.02	0.08	-0.19	0.19	0.22	-0.01	-0.05	-0.09	-0.08	-0.22	-0.05	1.00	
18	Organization Capital	-0.12	-0.14	-0.14	0.02	-0.01	0.06	-0.03	-0.12	0.05	-0.11	-0.12	0.00	0.05	-0.20	-0.05	-0.01	-0.07	1.00
19	Capital Structure	-0.01	-0.07	-0.06	-0.04	0.11	-0.03	-0.07	0.08	-0.18	-0.13	-0.07	0.11	0.14	-0.15	-0.08	-0.09	0.05	0.09

		19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
19	Capital Structure	1.00															
20	Analyst Following	0.10	1.00														
21	Institutional Ownership	0.10	0.92	1.00													
22	Poison Pill	-0.02	-0.02	0.05	1.00												
23	Golden Parachutes	0.13	0.07	0.14	0.23	1.00											
24	Family Firm	-0.08	-0.14	-0.13	-0.18	-0.27	1.00										
25	Blockholder	0.02	-0.11	0.01	0.07	0.16	0.00	1.00									
26	Board Size	-0.02	0.14	0.05	-0.04	-0.01	-0.01	-0.10	1.00								
27	Mergers and Charter Amendments	0.02	0.11	0.11	0.04	0.08	0.02	0.04	-0.01	1.00							
28	Staggered Board	0.02	0.04	0.09	0.25	0.18	-0.08	0.12	0.02	0.19	1.00						
29	Board Independence	0.03	0.17	0.20	0.06	0.25	-0.25	0.05	0.03	0.06	0.03	1.00					
30	Bylaw Amendments Limit	0.00	0.04	0.05	-0.05	0.14	-0.07	0.10	0.04	0.22	0.10	0.20	1.00				
31	Industry Patent Intensity	-0.12	-0.11	-0.09	0.10	0.03	-0.03	0.00	-0.09	-0.01	0.01	-0.06	-0.10	1.00			
32	Industry Citation Intensity	-0.14	-0.12	-0.13	0.05	-0.08	-0.01	-0.04	-0.08	-0.08	-0.10	-0.10	-0.13	0.65	1.00		
33	Industry Competition	0.11	0.14	0.14	-0.07	0.01	-0.06	-0.03	0.00	-0.01	-0.06	0.03	0.03	-0.22	-0.13	1.00	
34	Industry R&D	-0.09	-0.08	-0.06	0.09	0.03	0.00	-0.03	-0.01	-0.01	0.03	0.00	0.04	0.60	0.13	-0.34	1.00
35	Industry Size	-0.12	0.02	-0.02	0.07	-0.03	-0.07	-0.09	0.08	-0.08	-0.03	0.08	0.06	0.31	0.29	-0.18	0.53

Table IA3 Including Non-Patenting Firms

This table presents Adaptive Lasso variable selection results when we include firms with zero patents.

	Frequency Chosen	
Variables	**Patent**	**Citation**
Managerial Characteristics:		
CEO Centrality	**9/9**	4/9
CEO Delta	0/9	0/9
CEO Confidence	0/9	0/9
CEO Vega	1/9	0/9
CEO Total Pay	1/9	0/9
CEO Age	0/9	0/9
CEO Gender	0/9	0/9
Firm Characteristics:		
Size	**9/9**	**6/9**
R&D Stock	**9/9**	2/9
Tobin's q	0/9	0/9
Stock Liquidity	**9/9**	**9/9**
Distance to USPTO	**7/9**	0/9
Tangibility	0/9	0/9
State Tax	0/9	0/9
ROA	0/9	0/9
Manufacturing	0/9	0/9
Sales Growth	0/9	0/9
Organizational Capital	0/9	0/9
Capital Structure	0/9	0/9
Corporate Governance:		
Analyst Following	**7/9**	0/9
Blockholder	1/9	1/9
Bylaw Amendments Limit	2/9	0/9
Institutional Ownership	0/9	0/9
Poison Pill	1/9	0/9
Board Size	0/9	0/9
Golden Parachutes	0/9	0/9
Family Firm	0/9	0/9
Mergers and Charter Amendments	0/9	0/9
Staggered Board	0/9	0/9
Board Independence	0/9	0/9
Industry Characteristics:		
Industry Citation Intensity	**9/9**	**9/9**
Industry Patent Intensity	**7/9**	0/9
Industry Size	4/9	0/9
Industry R&D	0/9	0/9
Industry Competition	0/9	0/9

Table IA4 Variable Selection: R&D Expenditures

This table presents the results of variable selection in two-year rolling window using R&D as the dependent variable. All the variables are defined in Appendix C.

	Frequency Chosen	
	Stepwise	**Adaptive Lasso**
Variables	**R&D**	
Managerial Characteristics:		
CEO Centrality	1/9	0/9
CEO Confidence	0/9	0/9
CEO Gender	0/9	0/9
CEO Vega	4/9	0/9
CEO Total Pay	0/9	0/9
CEO Delta	1/9	0/9
CEO Age	0/9	0/9
Firm Characteristics:		
Size	**9/9**	**9/9**
Tobin's q	1/9	0/9
Stock Liquidity	**9/9**	**9/9**
Distance to USPTO	0/9	0/9
Tangibility	**9/9**	1/9
Manufacturing	1/9	0/9
State Tax	4/9	3/9
ROA	0/9	0/9
Capital Structure	1/9	0/9
Sales Growth	0/9	1/9
Organizational Capital	0/9	0/9
Corporate Governance:		
Analyst Following	0/9	0/9
Poison Pill	0/9	0/9
Blockholder	0/9	0/9
Institutional Ownership	0/9	0/9
Board Size	0/9	0/9
Golden Parachutes	0/9	0/9
Board Independence	0/9	0/9
Mergers and Charter Amendments	0/9	0/9
Staggered Board	0/9	0/9
Bylaw Amendments Limit	0/9	0/9
Family Firm	0/9	0/9
Industry Characteristics:		
Industry Citation Intensity	0/9	0/9
Industry Patent Intensity	**8/9**	**8/9**
Industry Size	0/9	0/9
Industry R&D	1/9	1/9
Industry Competition	0/9	0/9

Table IA5 Factors on Innovation: An Inclusive Test

This table presents the results of multiple managerial and firm factors on innovation among patenting firms. All the variables are defined in Appendix C. The Huber-White Sandwich estimator is clustered at the firm level. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)
	Only Include One Variable		*Include All*	
Dependent Variable:	*Patent*	*Citation*	*Patent*	*Citation*
CEO Total Pay	0.482***	0.179***	-0.043	-0.047
	(5.56)	(4.68)	(-1.30)	(-1.05)
CEO Delta	0.322***	0.132***	-	-0.041
	(11.02)	(7.72)	(-2.71)	(-0.98)
CEO Vega	0.414***	0.154***	0.126***	0.068
	(13.57)	(8.53)	(3.01)	(1.48)
CEO Centrality	0.343***	0.175***	0.156***	0.139*
	(5.86)	(4.39)	(3.53)	(1.79)
CEO Age	0.563**	-0.031	0.047*	0.063**
	(2.10)	(-0.22)	(1.87)	(2.41)
CEO Gender	0.201	0.269	0.009	0.029
	(0.50)	(1.14)	(0.30)	(0.75)
CEO Confidence	-0.158**	-0.042	-0.059**	-0.062**
	(-2.14)	(-1.20)	(-2.15)	(-2.06)
Size	0.351***	0.097***	0.489***	0.316***
	(25.27)	(12.15)	(8.37)	(4.63)
Tobin's q	0.059	0.069	-0.001	-0.005
	(1.45)	(1.52)	(-0.04)	(-0.16)
Stock Liquidity	0.614***	0.534***	0.195***	0.211***
	(15.77)	(9.81)	(4.14)	(3.43)
R&D Stock	-0.189***	-0.059***	0.235***	0.137***
	(-6.61)	(-6.25)	(5.29)	(3.28)
Manufacturing industry (1-digit SIC = 3, 4)	0.029	0.032	0.064	0.053
	(0.44)	(1.21)	(1.53)	(1.18)
Distance to USPTO	-0.085***	-0.032***	-	-0.041
	(-2.93)	(-2.74)	(-2.81)	(-1.11)
Tangibility	-0.036	-0.097**	0.017	-0.026
	(-0.76)	(-2.36)	(0.49)	(-0.72)
Analyst Following	0.358***	0.101***	0.228***	0.240*
	(14.11)	(7.42)	(2.82)	(1.92)
Institutional Ownership	1.396***	0.341	-0.126*	-0.188*
	(11.69)	(6.49)	(-1.73)	(-1.87)
Blockholder	1.028***	0.335***	-0.045**	-0.039
	(13.51)	(8.29)	(-2.02)	(-1.27)
Family Firm	-0.480***	-0.173**	-0.017	-0.008
	(-3.32)	(-2.30)	(-0.51)	(-0.25)
ROA	1.136***	0.319***	-0.013	-0.022
	(14.53)	(10.27)	(-0.46)	(-0.66)
Capital Structure	0.717***	0.066*	-0.017	-0.039
	(7.02)	(1.75)	(-0.66)	(-1.55)
Sales Growth	-0.033***	-0.014***	-0.004	-0.006
	(-5.17)	(-6.75)	(-0.20)	(-0.29)
Board Size	0.270***	0.253	0.004	0.009
	(5.15)	(0.81)	(0.16)	(0.33)
Board Independence	1.594***	0.428**	0.007	-0.005
	(4.38)	(2.36)	(0.27)	(-0.17)
Staggered Board	-0.117	-0.135***	0.031	-0.027

	(-1.26)	(-3.13)	(1.03)	(-0.83)
Bylaw Amendments Limit	-0.119	-0.079**	-0.000	-0.053
	(-1.29)	(-2.38)	(-0.00)	(-1.60)
Poisson Pill	0.084	-0.065	-0.056*	-0.056
	(0.93)	(-1.49)	(-1.87)	(-1.58)
Golden Parachutes	-0.106	-0.082**	-0.012	0.003
	(-1.26)	(-2.25)	(-0.43)	(0.08)
Mergers and Charter Amendments	-0.256***	-0.152***	-0.019	-0.024
	(-2.77)	(-4.51)	(-0.79)	(-1.04)
Organizational Capital	-0.514***	-0.108***	0.023	0.030
	(-9.95)	(-6.10)	(0.84)	(0.80)
State Tax	3.113***	0.787***	0.003	-0.012
	(15.74)	(9.73)	(0.09)	(-0.40)
Industry Patent Intensity	1.737***	0.048	0.082	-0.141**
	(4.29)	(0.27)	(1.50)	(-2.11)
Industry Citation Intensity	22.389***	8.814***	0.214***	0.317***
	(9.21)	(7.93)	(4.97)	(6.58)
Industry Competition	1.123**	0.303	0.037	0.002
	(2.39)	(1.35)	(1.50)	(0.08)
Industry R&D	-0.993*	-0.573***	0.086*	0.112**
	(1.69)	(-3.04)	(1.93)	(2.02)
Industry Size	0.170***	0.057***	-	-
	(5.74)	(5.90)	(-3.07)	(-2.60)
Year Dummy	*Yes*	*Yes*	*Yes*	*Yes*
Adjusted R^2	-	-	*0.600*	*0.424*