Does Stock Market Boost Firm Innovation? Evidence from Chinese firms*

Hui He†, Hanya Li‡ and Jinfan Zhang§

Abstract

We investigate the effect of stock market on firm innovation via the lens of initial public offering (IPO) using a uniquely matched Chinese firm-level data. We find that IPO leads to an increase in both quantity and quality of firm innovation activity. In addition, IPO expands a firm's scope of innovation beyond its core business. The impact of IPO on firm innovation varies across corporate governance structure, financial constraint, and ownership. Our results further illustrate that IPO makes a firm better at retaining internal inventors and hiring external inventors after the IPO. Finally, we show that the enhanced innovation activity increases a firm's Tobin's Q in the long run.

** First draft: February 2016. We thank Tong Hui, Jianhuan Xu, and Colin Xu for helpful suggestions. Hui He thanks research support sponsored by Shanghai Pujiang Program (No. 2013140026) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (No. 2013140034). The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.
† International Monetary Fund (IMF) and Shanghai University of Finance and Economics (SHUFE). E-mail: hhe@imf.org;
‡SHUFE. E-mail: lihnyaa@gmail.com;
§ Corresponding author. IMF. E-mail: JZhang2@imf.org.

1. Introduction

Does stock market affect a firm's innovation activity? This is an important question because, at the macro level, technological innovation is vital for a country's economic growth (Schumpeter 1943, Solow 1957, Aghion and Howitt 1992), while at the micro level, the innovation capacity defines a firm's long term competitive advantage (Porter 1992).

On the one hand, stock market provides firms an important access to raise capital. In a frictionless world, the way of financing should not affect a firm's innovation. The existence of financial constraint, however, makes a firm's innovation activity sensitive to credit constraint. IPO can thus help increase firms' innovation activity by providing firms with the access to low cost funding source. Moreover, as Holmstrong (1989) points out, since the nature of innovation is long term and idiosyncratic, the payoff of innovation is heavily skewed and risky. As a result, debt could not be an efficient way of financing innovation as compared to equity. For these reasons, IPO could lead to more innovation activities by providing firms access to low cost equity financing.

On the other hand, the corporate finance literature has widely documented that agency problems weaken the efficiency of a firm's business operations after IPO, e.g. investment and M&A (Berle and Means 1932, Jensen and Meckling 1976). Similar analysis also applies to innovation activities. Innovation often is risky and hence could cost a manager her job. Career concern thus may force mangers to choose less risky and therefore less innovative projects to pursue after the IPO, which would eventually weaken a firms' innovation activity. In addition, innovation is a difficult and time-consuming task which might require a manager's extra time and effort. An entrenched manager therefore could put less effort on innovation and enjoys her “quiet life”, because she faces weaker investor monitoring and feel her job is more secure after the IPO. These two incentive channels would lead to less innovation after IPO.

With all these trade-offs, whether IPO would increase or decrease a firm’s innovation remains an empirical question. Using a uniquely matched Chinese firm-level data, our paper aims to tackle this empirical question. In this paper, we use the annual patent applications data from China's State Intellectual Property Office (SIPO) to capture the firm innovation activities. First, we use the number of patents firms apply each year to measure the *quantity* of innovation activity. Second, China’s patent system divides patents into three categories: invention, design and utility. The invention patents are taken as major innovation and would go through rigorous and lengthy examination of substance similar to the practice in the US. The design and utility patents are taken as minor innovation, thereby receiving much loose examination of substance and weaker protection as compared to the invention patents. Prior literature, which mainly focuses on the US and the EU patents data, uses the patent citation number as a proxy for the quality of innovation (e.g. He and Tian, 2013). Since the citation data is not available for the Chinese patents. We, instead, take advantage of the unique characteristics of the Chinese patent system and use the number of invention patent applications (the proportion of invention patents, and per-patent inventor number) to measure the patent *quality*. Finally, the detailed information on patents’ technology class allows us to explore a firm’s innovation *scope* after IPO.

We match patent application data with the Chinese Industrial Survey (CIS) data, which contains the detailed balance sheet information for *all* industrial firms (both listed and non-listed) that are either state-owned, or are non-state owned firms with annual sales ≥RMB 500 million (about 800000 US$), via firm names (See Section 3.1 for details of matching the two datasets). To identify an IPO firm and its year of IPO, we further rely on the Chinese Stock Market & Accounting Research Database (CSMAR).

With all these uniquely matched patent-firm-IPO datasets, we first run a baseline panel data ordinary least squares (OLS) regression. Our results suggest that firm IPOs are positively associated with both innovation quantity and quality. To be more specific, public listing is associated with an average of 37.5% annual increase in patent applications in the subsequent three years. Similar increase also holds for the invention patent applications, indicating that IPO boosts both quantity and quality of firm innovation.

An important concern of the OLS analysis is that public listing could be endogenous and therefore our results could be subject to selection bias. For example, some factors could be correlated with both firm IPO decision and its innovation activity simultaneously. As Jain and Kini (1994) pointed out that firms choose to go public at a specific stage of its life cycle, which could correspond to certain level of firm innovation activities.

To mitigate these endogeneity concerns, we employ a difference-in-difference (DiD) approach that compares the innovation output of firms that are either IPO firms or subsidiaries of IPO firms (which were established before IPO and keep being subsidiary firms over the three years after IPO) to that of firms that are neither IPO firms nor their subsidiaries. To validate the DiD approach and make sure that IPO firms are compared to similar unlisted firms, we carefully match each IPO firm (and subsidiary firm) in the treatment group with a firm in the same industry but outside of the treatment group using the *propensity score matching (PSM) algorithm* that includes a broad range of firm characteristics. After undertaking a series of diagnostic tests to ensure the removal of observable differences between these two groups of firms, we find that treatment group firms experience an average of 22.9% annual patent application increase due to the IPO. The quality of the patent, as indicated by the invention patent applications also increases by 12.9%. Other robustness tests, such as invention patent ratio and per-patent inventor number also indicate an increase in innovation quality.

We further examine the scope change of firms' innovation after their IPOs in the DiD framework by dividing a firm' patent applications into two categories, *related* and *unrelated*, based on whether a patent's technology area is related to the applying firm's core business or not. Theoretically, capital raised in IPO could be used to either help a firm strengthen its competitive advantage in its core business, or facilitate a firm expanding into new business. We provide empirical evidence showing that IPO firms increase their innovation activity in both core business and new business.

The DiD estimation approach helps us mitigate the endogeneity problem in the OLS regression and shows that IPO indeed has positive impact on innovation activity. To better understand what factor causes the positive impact of IPO on innovation, we further apply a difference-in-difference-in-difference (DiDiD) framework to explore how the IPO impact varies across firms' corporate governance structure, financial constraint and ownership.

Our results first show that corporate governance structure plays an important role on IPO's positive impact on innovation (including both quantity and quality). We find that IPO has a stronger impact on firms granting stocks to their managers. By aligning the interest of managers and shareholders, stocks provide a stronger incentive for managers to innovate, so as to create long-term value for the firm. In addition, we find that firms with the CEO and president being the same person tend to innovate more after IPO. The career concern theory argues that managers may not want to engage in innovation to endanger their job given the high failure risk associated with innovation. According to this theory, CEOs, who are also the presidents of the firm, are entrenched (likely are founders or even controlling shareholders) thereby more willing to take risk. Our results lend support to the career concern theory (but against "quiet life" hypothesis).

We further show that financial constraint and ownership structure can make a difference on firms' innovation performance after IPO. Following the literature (e.g., Kiyotaki and Moore 1997), we use the ratio of fixed asset in a firm's total assets as a proxy for the credit constraint. A high ratio implies a higher collateral value and hence a less stringent credit constraint. We find evidence showing that an IPO firm facing a stricter financial constraint before IPO tends to improve innovation more after IPO (both in quantity and quality). Moreover, we discover that privately-owned enterprises (POEs) witness higher innovation increase than state-owned enterprises (SOEs) after IPO. Since the POEs are better at dealing with the agency problems than SOEs and face more stringent financial constraint[**], this result further corroborates the importance of agency problem and financial constraint in affecting the innovation activity after IPO.

Given the substantive increase in firm innovation quantity, quality and the expansion of innovation scope after IPO, it is natural to raise the question: How does it happen? Leveraging the inventor information in our patent data, we show that hiring new inventors and retaining existing inventors might be the reason behind the enhanced innovation activity after a firm's IPO. Out DiD test indicates that IPO firms tend to increase their inventors more than non-IPO firms after IPO, implying that the accumulation of human capital is one mechanism through which firms can increase their patent applications. In addition, our analysis of inventors further demonstrates that IPO firms have much lower inventor departure rate than the non-IPO firms, indicating that IPO helps a firm better retain its innovators. Finally, we don't find evidence that the increase in patent application is due to the productivity growth of inventors after IPO. Therefore, the enhanced innovation after IPO seems to be associated with changes in the *extensive* margin rather than *intensive* margin of the inventors.

In the last part, we explore the "bottom-line" question: whether the enhanced innovation activity after IPO helps creating firm value in the long run? Our answer is yes. This question is important because IPO firms may simply increase their patent applications for the purpose of "window dressing". We have two evidences against the "window dressing" hypothesis. First, as mentioned above, the quality of the patent application (measured by both invention patent counts, the share of invention patent in the total patent, and the number of inventors for each patent) increases after IPO. Second, we find that firms, which apply more patents during the three years after IPO, see a larger Tobin's Q increase in the subsequent three years, indicating that an IPO firm's patent application is associated with a long-term value creation. Moreover, we find that invention patents have a stronger impact on Tobin's Q increase than other patent types. All the above results lend support to the argument that the increased innovation activity after IPO leads to firm value creation in the long run, rather than just window dressing activity to boost short term valuation.

The rest of the paper is organized as follows. Section 2 discusses the related literature. Section 3 describes data sample and variable construction, and reports summary statistics. Section 4 presents the baseline results and addresses identification issues on innovation quantity, quality and scope. Section 5 explores the relationship among IPO, inventors and innovation. Section 6 investigates the consequence of enhanced innovation after IPO on a firm's value. Section 7 concludes.

2. Relation to the existing literature

Our paper contributes to three strands of literature. First, the literature on real impact of IPO literature has documented a wide range of post-IPO firm performance change, such as the profitability decline documented in Jain and Kini (1994), Pagano, Panetta, and Zingales (1998), and Pastor, Taylor, and Veronesi (2009); and productivity reduction shown in Chemmanur, He, and Nandy (2009). This paper contributes to the literature by leveraging on our data and proposing a new DiD strategy to estimate the IPO effect on firms' innovative activities in a developing country context.

This paper is closely related to the contemporaneous research by Bernstein (2015), which examines a similar question in the context of US IPOs. Our paper differs from Bernstein (2015) in two-folds. First, we have implemented a Difference-in-Difference methodology to tackle the identification problem leveraging our unique datasets, which include the balance sheet information of *both* listed and non-listed manufacturing firms in China. In contrast, Bernstein (2015) relies on an instrumental variable methodology using the US IPO firm data.[††] Second, we find that IPO can substantially enhance firm innovation activity in quantity, quality and scope, which is in sharp contrast to the findings of Bernstein (2015) that IPO has no impact on firm innovation quantity and has significantly negative impact on quality. Our paper complements the Bernstein (2015) by emphasizing on the positive impact of IPO on innovation in the *developing economy* context, where firms face different financial constraint and corporate governance environment compared to their counterparts in developed countries. Our finding suggests that the benefit of IPO outweighs the cost in affecting firm innovation in *developing* countries. This could be due to two reasons. First, firms in developing countries often face much

[††] Since IPO is not exogenous event, to control potential selection bias, Bernstein (2015) uses NASDAQ fluctuations during the book-building phase as an instrumental variable (IV) for IPO completion in the estimation.

more severe financial constraint than firms in developed countries, which makes the access to equity market much more beneficiary for innovation in developing economy (e.g. we find that firms in China are better at retaining their inventors and attract more outside inventors after IPO). Second, firms in developing countries usually do not stay at the forefront of technological innovation as their counterparties in developed countries do. Therefore, their innovation is less risky and therefore less likely being affected by agency problem, which according to Bernstein (2015) is the source of the major negative impact on innovation brought by IPO in the US.

Second, the paper is also related to a broader literature on equity market and innovation. The paper is related to a growing literature that compares the behavior of public and private firms along various dimensions such as investment sensitivity (e.g., Asker, Farre-Mensa, and Ljungqvist, 2010; and Sheen 2009), debt financing and borrowing costs (Saunders and Steffen, 2009; and Brav, 2009), dividend payouts (Michaely and Roberts, 2007), and CEO compensation (Gao, Lemmon, and Li, 2010). Atanassov, Nanda, and Seru (2007) argue that arm's length financing (equity and public debt) is positively related to innovation while relationship-based bank financing is negatively related to innovation. Our paper advances this line of inquiry by providing new evidence showing the positive role arm's length financing can play in innovation.

Third, this paper also contributes to a growing theoretical and empirical literature that explores the role of governance, capital structure, and ownership concentration on corporate innovation. For instance, a larger institutional ownership (Aghion, Van Reenen, and Zingales, 2013), corporate venture capital (Chemmanur et al., 2014) and "hot" rather than "cold" markets (Nanda and Rhodes-Krpof,2013) all alter managerial incentives and hence motivate managers to focus more on long-term innovation activities. Our paper provides new evidence showing that career concern and management incentive can play an important role in affecting firm innovation after IPO.

3. Data description

3.1 Data and sample construction

A. The firm-level data

Our first dataset is the annual Chinese Industrial Survey (CIS), conducted by the National Bureau of Statistics of China for the period from 1998 to 2007. This data provides the most comprehensive coverage of the Chinese manufacture firms, including all state-owned and non-state-owned with sales more than five million Renminbi (about US $800000). The number of firms included in this database various from over 165000 in 1998 to 337000 in 2007. Firms in CIS span all over the country and across all manufacture industries. CIS provides detailed information on firm registration (e.g., name, location, industry, age) and balance sheet variables such as capital, debt, ownership structure (POE vs SOE), employment, sales, interest payment, profit and etc. CIS provides a rare opportunity to get access to the high quality firm-level data for non-listed firms, which makes it widely used in the recent literature (e.g., Hsieh and Klenow 2009, Brandt, Biesebroeck and Zhang 2012, Aghion et al. 2014, Hsieh and Song 2015).

B. The patent filing data

We use a firm's patent application as a proxy for the firm's innovation activity. Our patent filings data are from the State Intellectual Property Office of China (SIPO). The patent application dataset contains detailed information on each patent, such as the application identifier, the title of the patent, the application date, the type of the patent (i.e. whether the patent is an invention patent, a utility patent, or a design patent), the technological class of the patent, the names of the inventors, and patent's application institution. The dataset covers all the patent applications from 1985 to 2011, with annual application quantity varies from 575 in 1985 to 1,108,534 in 2011.

In China, patents are classified into three classes: invention, utility and design. Invention patents are granted for major technological innovation as compared to utility and design patents. Invention patents undergo much more rigorous scrutiny by the patent officers. Inventors must submit a clear and comprehensive description of the invention and reference materials for patent application and accept the "substantive examination" (including novelty, inventiveness and industrial applicability). It usually takes 2-3 years for an invention patent application to go through the entire process. Once granted, invention patents enjoy a much stronger legal protection than utility and design patents. For all these reasons, the invention patents are the highest quality among all three types of patents. Therefore, we use the quantity of invention patent applications as a measure of innovation quality.

We match the patent application data with firm information from CIS data using the firm name. The accuracy of the match is checked by comparing the location information of the patent filing firms in the patent data to the location information of the matching firms in the CIS data. In our sample, 6.01% of firms have applied for at least one patent over the period from 1998 to 2007. According to a report by the National Bureau of Statistics of China, about 8.8% of manufacturing firms with annual sales above 5 million Renminbi (RMB) have applied for patents during 2004-2006. Given the long term trend that Chinese firms are increasing their patent applications during the sample period, the quality of our matching is reasonably good. We also only keep firms that have at least five years of observations in CIS data to minimize the noise of short lived firms and abnormal observations. After further filtering out firms with key missing variables and obvious data error following the literature such as Cai and Liu (2009), we eventually obtain a final sample of 999,114 firm-year observations from 136,762 firms.

C. IPO data

Our third dataset is the Chinese Stock Market & Accounting Research Database (CSMAR), through which we can identify the IPO firms and their subordinate firms. Among the 620 IPOs from 1999 to 2006, we can directly match 148 manufacture IPO firms in the CIS data and further identify another 208 firms subsidiary to IPO firms[‡‡], which were established at least one year before the IPO and are subsidiary to the IPO firms over the subsequent three years after IPO. Among these 356 firms affected by the IPO, 190 firms have at least one patent application over the period from 1998 to 2007.

In summary, by linking the CIS, patent filing, and CSMAR data together, we have generated a comprehensive dataset, which includes firm balance sheets, patent filings and IPO information from 1998-2007.

3.2 Variable construction

We use the patent filing counts to quantify the innovation activity. The distribution of patent applications in our final sample is right skewed, similar to U.S. patent data, with its median at zero as shown in Table 1. Following the literature (e.g. He and Tian 2013, Gu, Mao and Tian 2014), we winsorize these variables at the 99th percentiles. We then use the natural logarithm of one plus patent counts, LnPat, as the main measure of innovation activity. As illustrated before, since the invention patent is widely regarded as of higher quality compared to

[‡‡] We require that the matched firms have to have at least one year observations before and after the IPO in CIS data.

the design and utility patents, we also used the natural logarithm of one plus invention patent applications (LnInv) as the measure of the quality of a firm's innovation.

In our regression analysis, we control for a set of firm and industry characteristics that might affect a firm's future innovation output following the innovation literature. In the baseline regression, the control variables are firm size measured by the natural logarithm of total assets, Ln(AT); FixAT_AT, measured by net fixed assets divided by total assets; Admin_AT, measured by administration expenditure divided by total assets; Interest_AT, measured by interest expenditure divided by total assets; Leverage, measured by total liability divided by total assets; Liquidity, measured by the ratio of the difference between current assets and liability, and total assets; ROA, measured by net income divided by total assets; Ln(Age), measured by natural logarithm of one plus the number of years after establishment; HI, measured by the Herfindahl index based on annual sales; SOE, a state-owned enterprise dummy which equals one for SOEs; EX, an exporter dummy which equals one for exporters; Stock, patent stock measured by total patents applications of the firm prior to current year. To circumvent potential non-linear effects of product market competition (Aghion et al., 2005), we include the squared Herfindahl index in our baseline regressions. All variables are computed for firm *i* over its fiscal year *t*.

3.3 Summary Statistics

Following the literature, we winsorize all variables at the 1^{st} and 99^{th} percentile of their distribution to mitigate the effect of outliers. Table 1 presents summary statistics of main variables used in the analysis based on our final sample. On average, a firm in our sample files 0.18 patents per year and among them 0.03 are invention patents. In comparison, the IPO firms and their subsidiaries are more active in innovation. These IPO related firms file 2.21 patents per year and 0.68 are invention patents on average. In our merged sample, an average firm has a total asset of RMB 9.99 million (1.5 million US dollars), a fixed asset to total asset ratio FixAT_AT of 0.33, and a leverage ratio of 0.59. The firms in our sample are on average 12.64 years old.

4. Empirical results

4.1 Time series pattern and the panel regression results

In Figure 1, we plot the average number of industry and year adjusted innovation output variable around an IPO within a 7-year window around the IPO event. More specifically, we adjust total patent (invention patent) filing number by subtracting the average value of total patent (invention patent) filing of all firms in the same industry and year. The time-series plots in both Panel A and Panel B demonstrate that total patent (including invention, utility and design patents) and invention patent filings do not significantly increase over the three years leading up to firms' IPOs. In contrast, there is sharp jump in the number of both total patent filing and invention patent filing starting from the IPO year. The increase lasts through the three years following the IPO.

Inspired by the time-series raw pattern of innovation output surrounding the IPO as demonstrated in Figure 1, we use the IPO firms' firm-year observations for a seven-year window centered on the IPO event and estimate an OLS regression:

$$Pat_{it}(Inv_{it}) = \alpha + \beta_1 before_{it}^{-3} + \beta_2 before_{it}^{-2} + \beta_3 current_{it} + \beta_4 after_{it}^{1} + \beta_5 after_{it}^{2} + \beta_6 after_{it}^{3} + \varepsilon_{it} \quad (1)$$

where *i* indexes firm and *t* indexes year. The dependent variable is firm *i*' total patent filing (invention patent filing) adjusted for the industry-year average in year t. $before_{it}^{-3}$ is a dummy that equals one if the firm-year observation is three years before the IPO event, and zero otherwise. Similarly, $before_{it}^{-2}$ is a dummy that equals one if the firm-year observation is 2 years before the IPO event and zero otherwise. $current_{it}$ is a dummy that equals one if the firm-year observation is exactly in the IPO year and zero others. Dummy variables $after_{it}^{1}, after_{it}^{2}$, $after_{it}^{3}$ equal to one when the observations are one year, two years and three years after the IPO event respectively, and zero otherwise. As can be seen, the benchmark group is the observations one year before the IPO.

The results are reported in Table 2. We find that the coefficient estimates of $before_{it}^{-2}$ and $before_{it}^{-3}$ are statistically insignificant for both total patents and invention patents, showing that firm innovation output stays at the same level before the IPO. In contrast, the coefficients of $current, after_{it}^{1}, after_{it}^{2}$ and $after_{it}^{3}$ are all positive and statistically significant, indicating a jump in innovation activity, both quantity and quality, after the IPO. These findings confirm the time-series trend shown in Figure 1 Panel A and Panel B

Next, to take advantage of our rich data set, we explore the impact of IPO on firm innovation using a panel data regression approach. We estimate the following regression

$$\mathrm{LnPat}_{i,t+\mathrm{n}}(\mathrm{Ln}Inv_{i,t+n}) = \alpha + \beta list_{it} + \gamma X_{it} + year_{t} + industry_{j} + \varepsilon_{it} \qquad (2)$$

where $\mathrm{LnPat}_{i,t+\mathrm{n}}(\mathrm{Ln}Inv_{i,t+n})$ is the natural logarithm of one plus the total patent (invention patent) filing in year *t*+*n* for firm *i*, where *n* is equal to 1,2,3. The variable of interest $list_{it}$ is equal to one if firm *i* or firm *i*'s parent firm undergoes IPO in year *t*, and zero else. X_{it} is a group of firm-level control variables, summarized in Table 1, which according to literature (e.g. Gu, Miao and Tian, 2013) are relevant to a firm's innovation activities. $year_{t}$ and $industry_{j}$ control for the year and industry-level fixed effects.

The results of the panel data regression are reported in Table 2 panel B. The coefficient estimate of $list_{it}$ is positive and statistically significant at the 1% level for n=1,2,3. Based on the coefficients estimated from columns (1)-(3), going public is associated with an average annual patent filings increase by 37.5%, indicating an economically significant magnitude. Similar results also hold for the invention patent applications according to columns (4)-(6). Various robustness tests have been conducted to corroborate our baseline OLS regression results. For instance, we have tried the quantile regressions with various specifications, and the coefficients on the list dummy variable are always positive and significant.

4.2 Difference-in-difference Test

A legitimate concern of the panel data regression results is the endogeneity problem. It is possible that both the IPO decision and firm innovation are affected by firms' unobservable characteristics. For instance, a rapidly growing firm is more likely to go public, while, at the same time, it has reached to a growth stage to increase its innovation activity. Our identification strategy to alleviate the endogeneity concern is based on a difference-in-difference (DiD) approach. The CIS data, which provides detailed balance sheet information for non-publicly listed manufacture firms, allow us to match the IPO firms with non-publicly listed firms that have similar firm characteristics with the IPO firms. We then compare the innovation output of the IPO firms to the innovation output of the matching firms, which do not go public, but are otherwise comparable, before and after the IPO.

The DiD approach has important advantages. First, it can rule out omitted trends that are correlated with IPO and innovation in both treatment and control group. For instance, the prosperity of certain industry can simultaneously increase the likelihood of IPO and future innovation. The DiD approach rules out the possibility that the industry prosperity, rather than the IPO itself, is driving the changing in innovation activity. Second, the DiD approach controls for constant unobservable differences between the treatment and control group.

We construct the treatment and control group of firms using the propensity score matching algorithm. Specifically, the treatment group includes all IPO firms and their subsidiary firms for 1999-2006, which have non-missing matching variables in the pre-IPO year (t-1) and the post-IPO year (t+1). We further require that the subsidiary firms must maintain the subsidiary status from IPO year t to year t+3 and they must be established no later than t-3.

We use a propensity score matching algorithm to identify matches between IPO firms and firms that do not go public. In our propensity score matching, we first estimate a probit model with 356 firms associated with IPO and 772,734 firm-year observations without IPO. The dependent variable equals one if the observation is from IPO firms and zero otherwise. In the probit model, we include all control variables from the baseline specification in equation (2) that are measured in the year immediately preceding the IPO. We also include the patent growth during the three years before IPO, g_pn, in the probit model to help ensure the parallel trend assumption, the key requirement for the DiD approach.

The probit regression results indicate that our model captures a significant amount of variation in the dependent variable. We report the probit model results in column (1) of Table 3 Panel A. As can be seen, the Pseudo R^2 reaches to 16.5% and a p-value from the Chi-square test of the overall model fitness is well below 0.0001, indicating that our model specification works well. We then use the predicted probabilities, or propensity scores to conduct a nearest-neighbor propensity score matching procedure. Specifically, we match each IPO firm-year observation (i.e. treatment group) with a non-IPO firm-year observation, which has the closest propensity score among the observations in that IPO year and belongs to the same industry as the IPO firm, to form the control group. We eventually generate 355 pairs of matched firms.[§§]

Before moving forward to the DiD test, we conduct a number of diagnostic tests to check the validity of the parallel trend assumption which is essential for DiD approach. First, we rerun the probit model restricting the observations to only the 710 matched firms after the pairing. The results are shown in column (2) of Table 3 Panel A. As can be seen, none of the coefficients of the independent variables are statistically significant. In particular, the coefficient of the pre-IPO patent application growth variable g_pn is insignificant, indicating that there is no observable difference in innovation trend between the IPO and non-IPO firms before the IPO event. Moreover, the Pseudo R^2 has reduced drastically from 16.52% to 0.66% after the matching and the P-value of Chi-square test is close to 1, indicating that the null hypothesis that all the coefficient estimates are zero cannot be rejected.

Furthermore, we conduct a series of univariate variable test of firm characteristics between the treatment and control group firms one year before the IPO. The results are reported in Table 3 Panel B, with the p-value of those tests shown in the last column of the table. As can be seen, none of the differences in these firm characteristics between the treatment and control group are statistically significant. In particular, the insignificance of the pre-IPO patent application growth variable (g_pn) provides a further evidence that the parallel trend assumption is likely hold.

In summary, the above diagnostic tests provide a strong evidence that our propensity score matching process has removed significant observable differences (other than the

[§§] Among the 356 firms affected by the IPO that we matched between CSMAR and CIS, only one firm cannot be matched to a similar firm using the propensity score.

difference in IPO) between the treatment and control group firms, which makes the changing innovation activity more likely driven by IPO in the DiD test.

To visually demonstrate the trend of patent application around IPO for both the treatment and control group firms, we have plotted the time-series of variable LnPat (LnInv) during a 7-year window around the IPO in Figure 2. As can be seen, the trend lines of treatment and control groups move closely in parallel during years leading up to IPO, which provides another piece of evidence for the parallel trend assumption. However, after the IPO, we can see a significant increase in patent applications for the line of treatment group. In sharp contrast, the line of control group keeps flat after the IPO.

To corroborate the visual evidence, we have conducted the formal DiD tests and reported the results in Table 3 Panel C and Panel D. Panel C shows that the change of average annual total patent (invention patent) applications around IPO. Specifically, the first column in Panel C reports the difference of average annual total patent (invention patent) applications during three years before and after IPO for the treatment group. Similarly, we calculate the change of average annual total patent (invention patent) application for the control group in the second column. These results indicate that the changes of both total patent and invention patent applications are positive and statistically significant for the treatment group. The changes, however, are insignificant for the control group. The third and fourth columns further report the DiD estimate results between the treatment and control group. We find that the DiD tests of the total patent and invention patent filings are both positive and statistically significant at the 1% level. The DiD test results are not only statistically significant, but also economically significant. On average, IPO results in an increase of 1.56 (0.77) total patent (invention patent) applications annually over the three years after the IPO.

Table 3 Panel D further demonstrates the innovation dynamics of the DiD results in a regression framework. The firm-year observations are gathered in a seven-year window centered on the IPO year. The regression model is as follows:

$$\ln Pat_{it}\ (\ln Inv_{it}) = \alpha + \beta_1 list_i * before_{it}^{-2\&-3} + \beta_2 list_i * current_{it} + \beta_3 list_i * after_{it}^{1} + \beta_4 list_i * after_{it}^{2} + \beta_5 list_i * after_{it}^{3} + before_{it}^{-2} + before_{it}^{-1} + current_{it} + after_{it}^{1} + after_{it}^{2} + after_{it}^{3} + year_t + firm_i + \varepsilon_{it} \tag{3}$$

where *i* indexes firm and *t* indexes time. $list_i$ is a dummy, which equals one if firm i belongs to the treatment group (IPO firms), and zero otherwise (non-IPO firms). $before_{it}^{-2\&-3}$, $current_{it}, after_{it}^{1}, after_{it}^{2}$ and $after_{it}^{3}$ are all dummy variables that equal one if the year t is three or two years before IPO, IPO year, one year, two years or three years after IPO, respectively, and zero otherwise. Therefore, the omitted or benchmark group is the observations one year before the IPO. We also include the year and firm fixed effects in the regression.

The coefficient estimates of interest are $\beta_{1,}\beta_2, \beta_3, \beta_4, \beta_5$. If the reverse causality holds, that is, the increase in innovation activities help improve the quality of the firm, which then leads to IPO, we should observe significant and negative coefficient estimates of β_1 However, in contrast, we find statistically insignificant coefficient estimates of $\beta_{1,}$, which indicates that there is not a systemically different trend in firm innovation activity between IPO firms and non-IPO firms before the IPO. The coefficient estimates of β_3, β_4 and β_5 are all positive and statistically significant, suggesting that compared to the non-IPO firms, IPO firms experience a significant increase in their patent application activity after IPO. Overall, our findings suggest that IPO helps increase firm innovation activity and the reverse causality does not hold.

To check the robustness of our results, we also repeat the propensity score matching algorithm and construct the control group samples by choosing five (instead of one) most closely matched non-IPO firms for each IPO firm in the treatment group. In our unreported results, the DiD tests on the treatment group and the augmented control group generate quantitatively similar results as in Table 3.

4.2 Further tests on innovation quality

To corroborate our findings on patent quality, we explore other proxies of patent quality as well. Our first proxy is the inventor per patent. The idea is that if a patent involves more inventors, it is likely that the patent requires more research input and expertise from multiple areas. Therefore, the patent is technologically more sophisticated and thus is of higher quality. For this reason, the inventor per patent can be used as a proxy for the quality of a patent.

By taking the inventor per patent as the dependent variable, we have conducted a similar patent level DiD regression test as in equation (3), using the same treatment and control group firms constructed by PSM. The results have been shown in Table 4 Panel A. As can be seen, since the coefficient estimate of $list_i * before_{it}^{-2\&-3}$ is statistically insignificant, the parallel trend assumption for the treatment and control group before the IPO is hold. The coefficient estimates for $list_i * current_{it}$, $list_i * after_{it}^{1}$ and $list_i * after_{it}^{2}$ are all statistically significant, indicating that after the IPO, the IPO firms increase their human capital deployment for each patent more than the non-IPO firms do. This result lends support to our hypothesis that IPO leads to higher quality firm innovation.

In addition to use the inventor per patent as the proxy for patent quality, we also explore the ratio of invention patent over total patent as another proxy for the quality of innovation. And we carry out the same univariate DiD test as in Table 3 Panel C. Column 1 of

Table 4 Panel B reports the change of average invention patent ratio for the IPO firms. We compute the change of the ratio by first subtracting the invention patent ratio over three years immediately preceding the IPO from the invention patent ratio counted over the three years immediately after the IPO for each treatment firm. The differences are then averaged across the treatment group. Similarly, we compute the average change of the invention patent ratio for the control group in column 2. Column 3 reports the DiD estimates that are the difference between columns 1 and 2. Column 4 reports the p-value of testing the null hypothesis that the DiD estimates are zero. We find that the DiD estimates of the increase in invention patent ratio is 0.043 after IPO with the p-value equals to 6.8%. The DiD test result is economically significant, given the fact that the average invention patent ratio for the firms before IPO is only 0.09.

Both the inventor per patent test and the invention ratio test indicate that firms not only increase the quantity, but also the quality of innovation output after the IPO. These results are in sharp contrast to Bernstein (2015)'s findings that the quality of firm innovation declines after IPO in the US context. Bernstein (2015) suggests that the decline is due to the agency problem between managers and shareholders. The failure of innovation may cost the manager's job if shareholders attribute the failure to the manager, thereby pushing managers to cut back those more innovative yet more risky R&D projects. This type of agency problem, while might be important in developed countries, is less relevant in developing countries. Developing country firms are usually not at the forefront of technology development as their counterparties in developed countries. The innovation activity in developing countries is not aiming at major technological breakthrough but rather incremental product improvement, which is much less risky thereby leading to less agency problem. Moreover, the financial resource gained through IPO is more beneficiary for firms in developing countries given firms in developing countries often face much more severe financial constraints. The benefit of financing through IPO could then dominate the negative impact brought by IPO in developing country context. We will formally test the relevance of financial constraints and corporate governance later in Section 4.4.

4.3 The scope of innovation

The scope is another important characteristic of innovation. On the one hand, focusing innovation in its core businesses allows a firm to strengthen its competitive advantage over competitors; on the other hand, innovation extended to other industries, especially industries unrelated to firm's current core businesses, can help firms acquire new growth opportunities. Therefore, the choice of a firm's innovation scope after the IPO is an interesting empirical question to explore.

We study the scope of innovation by first classifying a firm's patents into two categories: patents that are related to a firm's main business (defined as related patents); and patents that are unrelated to a firm's main business (defined as unrelated patents). Specifically, we define a patent with technology class which can be linked to a firm's Chinese Industry Classification (CIC) as related patent, and unrelated patent, otherwise. Practically, we first map each firm's CIC code to the international standard industrial classification (ISIC) code following Dean and Lovely (2010). Then, we use a concordance table provided by Ulrish Schmoch et al. (2003) to link the ISIC code to the International Patent Classification (IPC), which classifies a patent's technology classes. The link is developed based on whether a patent's technology area is closely related to a certain industry. Our procedure eventually links a firm's industries to the directly related patent technology classes(IPCs). If a patent's IPC belongs to one of these IPCs, it is defined to be a related patent, otherwise unrelated.

Based on the above patent classification, we exam how firms' will allocate their innovation resource in core and non-core business after IPO. We do so by investigating the changes in total patent applications and invention patent applications for both the related and unrelated patents surrounding IPO in a DiD framework, using the same propensity score matched sample of treatment and control group firms. Once again, we used a seven-year window centered around the IPO year and estimate the following regression model separately for the related and unrelated patent applications,

$$\mathrm{y_{it}} = \alpha + \beta_1 list_i * before_{it}^{-2\&-3} + \beta_2 list_i * current_{it} + \beta_3 list_i * after_{it}^1 + \beta_4 list_i * after_{it}^2 + \beta_5 list_i * after_{it}^3 + \beta_6 before_{it}^{-2} + before_{it}^{-1} + current_{it} + after_{it}^1 + after_{it}^2 + after_{it}^3 + year_t + firm_i + \varepsilon_{it} \quad (4)$$

where we define the independent variables the same way as in equation (3). The dependent variables we use in our tests are dummy variable $\mathrm{Dunrelated_{it}}$, which equals one if firm i has unrelated patents and zero otherwise; the number of technology classes that firm i's unrelated patents belong to, $\mathrm{UnrelatedClass_{it}}$; the natural logarithm of one plus related total patents, $\ln RelatedPat_{it}$; the natural logarithm of one plus unrelated total patents, $\mathrm{I}n\ UnrelatedPat_{it}$; the natural logarithm of one plus related invention patents, $\ln RelatedInv_{it}$; and the natural logarithm of one plus unrelated invention patents, $\mathrm{I}nUnrelatedInv_{it}$.

We report the results in Table 5. The coefficient estimation of β_1 for the six regressions are all statistically insignificant, suggesting that the parallel trend assumption behind the DiD test is likely to hold. The coefficient estimates of β_3, β_4 and β_5, on the other hand, are positive and statistically significant for all tests except for β_3 of Unrelated Class. The results in column (1) suggest that IPO firms are more likely to innovate beyond their core business than those matched unlisted firms after the IPO. In addition, IPO firms innovate in more unrelated technology classes according to the test in column (2). The results in columns (3)-(6) further indicate that IPO boosts the innovation quantity and quality in both the core and non-core business of the firms. In developing countries, an IPO firm, with sufficient financial resources, would have plenty of opportunities to access into new business. Our results confirm that IPO firms not only consolidate their advantage in their existing core businesses by increasing innovation (both quantity and quality), but also expand their innovation effort by entering into non-core businesses.

4.4 The cross-section analysis

The impact of IPO on firm innovation activity varies according to different firm characteristics. To explore how our results in the current section varies according to these characteristics could help us understand the underlying mechanism behind the positive impact of the IPO on innovation.

In this section, we explore the cross-sectional differences of IPO using the same propensity score matched sample in a seven-year window centered around the IPO year. we apply a difference in difference in difference (DiDiD) framework by estimating the following regression model:

$$\ln Pat_{it}\,(\ln Inv_{it}) = \alpha + \beta_1 list_i * \mathrm{After}_{it}^{1\&2\&3} * X_{it} + \beta_2 list_i * After_{it}^{1\&2\&3} + \mathrm{After}_{it}^{1\&2\&3} * X_{it} + \mathrm{After}_{it}^{1\&2\&3} + firm_i + year_t + \varepsilon_{it} \quad (5)$$

where $\mathrm{After}_{it}^{1\&2\&3}$ is equal to one if the observation is either one year, two years or three years after IPO, and zero otherwise. X_{it} are relevant firm variables. In our case, X_{it} could be a corporate ownership indicator, such as a dummy variable *SOE*, which equals to one if the treatment group firm is a state-owned enterprise and zero otherwise; Or a dummy variable *EXEHLD*, indicating whether the management team holds firm stocks (one if management holds, zero if not) after the IPO; Or a dummy variable $Duality_i$, indicating if the CEO and the president is the same person (one if it is, zero otherwise) after the IPO; Or a dummy variable $Fixed$, which equals one if a firm's fixed asset ratio (FixAT_AT) is above the median before the IPO measured in 1 year pre-IPO.

To save the space, in all the panels in Table 6, we only report the coefficient estimates for key variables $list_{it} * After_{it}^{1\&2\&3}$ and $list_{it} * After_{it}^{1\&2\&3} * X_i$. All other explanatory variables are suppressed. The dependent variables in our regression include both the total patent applications and invention patent applications.

In columns (1)-(2) of Table 6 where X_i is $EXEHLD$ dummy, the coefficient estimate of $list_{it} * \mathrm{After}_{it}^{1\&2\&3} * EXEHLD_i$ is positive and statistically significant for both total patents and invention patents, indicating that firms which grant stocks to their management team tend to witness a larger increase in innovation activity (both quantity and quality) after the IPO. Since the management team stock holding aligns the interest of managers and shareholders, it provides incentive for managers to engage in more innovation activities to boost a firm's long term value. Our result thus demonstrates the importance of management incentive structure in innovation activity after the IPO.

In addition, according to columns (3)-(4) of Table 6, we find that firms which have CEO and board president being the same person tend to have a larger increase in total patents and invention patents after the IPO, as indicated by the positive and statistically significant coefficient estimate of $list_{it} * After_{it}^{1\&2\&3} * Duality_i$. Since the CEO and president being the same person is a strong indicator of management entrenchment (which makes he/she more tolerant to the risk associated with innovation), our result shows that career concern could affect the innovation output after the IPO as well.

Overall, our results demonstrate that corporate governance could play an important role in innovation output after the IPO.

The impact of the IPO on innovation also varies according to the tightness of a firm's financial constraint before IPO. Following the literature (e.g., Kiyotaki and Moore 1997), we construct the variable *Fixed* as a proxy for the tightness of firms' financial constraint. *Fixed* equals to one if a firm's fixed asset/total asset is larger than the median level among all firms before its IPO, and zero otherwise. The idea is that a firm with more fixed asset has more collateral and therefore it is easier for the firm to borrow from banks or issue bonds in capital markets. A lower fixed asset/total asset ratio represents a tighter financial constraint. We report our results in columns (5) and (6) of Table 6. The coefficient estimates of the triple interaction term are negative and statistically significant. This shows that firms facing a tighter financial constraint (lower fixed asset ratio) before IPO have a larger increase in innovation activity after the IPO, both in quantity and quality.

We further explore the ownership structure and its impact on innovation output. The results are reported in columns (7)-(8) of Table 6. We find that POEs experience a much larger increase in innovation activities after IPO than SOEs.[***] On the one hand, managers in SOEs often have weaker incentive compared to their counterparts in POEs. For example, POEs often grant more stocks to the management team. On the other hand, SOE top managers are more likely subject to career concerns as indicated by the fact that SOEs are much more likely having president and CEO being separate persons. Both factors illustrate the importance of corporate governance in determining the impact of IPO. In addition, the SOE dummy could also be interpreted as a proxy for the tightness of financial constraint, as SOEs tend to face much less

[***] In an unreported exercise, We also separately conduct the DiD test as in equation (3) for SOEs and POEs subsamples, respectively. We find a significant positive impact of the IPO on innovation for POEs but insignificant IPO effect for SOEs.

obstacles to raise capital (see Song, Storesletten, and Zilibotti 2011). The negative results of SOE dummy is, therefore, also consistent with the financial constraint channel mentioned above.

5. Human capital and innovation

Our previous results demonstrate that IPO leads to an increase in innovation quantity, quality, and scope. It is natural to raise the question: what directly drives the enhanced innovation activity? The annual research investment, such as R&D cost, could be an important factor behind the innovation increase. Since the R&D information is limited in our data, we use inventor number as a proxy for innovation input.[†††] By leveraging our patent dataset, which contains the detailed inventor information, we calculate the number of inventors who produce patents for each firm-year observation.

Since IPO brings a firm more resource and market reputation, it could be easier for an IPO firm to hire external inventors and retain the internal ones. To explore the hypothesis that firms hire more inventors after IPO, we use the same DiD test framework in equation (3) and the same propensity score matched sample of treatment and control group firms. The sample is restricted to the three-year window before and after the IPO for both the treatment and matched control group firms as before. Table 7 panel A reports the results. The first column documents the average change in the number of inventors for the treatment group. We compute the changes by first subtracting the total number of inventors over the three-year window before the IPO from the total number of inventors over the three-year window after the IPO for each treatment group firm. The differences are then averaged across the treatment group. By the same token, we compute the average changes in the amount of inventors for the

[†††] We only have two year (2005 and 2006) R&D data available in CIS dataset w. Publicly-listed firms in CSMAR only reports R&D expenditures after 2006. Therefore in our matched data sample R&D data are limited.

control group and report the results in the second column. In the third and fourth columns, we report the DiD estimates that are the differences between column (1) and column (2) and the corresponding p-value of the test. We find that IPO helps a firm increase the number of inventors by 3.016 over three years, which is statistically significant at the 1% level. Comparing to the average number of inventors in a treatment group firm before the IPO, which is 3.75, IPO helps a firm increase its inventor count by 80% on average.

In addition to hiring more inventors, IPO could also help a firm to retain its best talents. To investigate this mechanism, we have carried out an analyst level analysis. Following the existing studies (e.g. Bernstein 2015)[‡‡‡], we identify two groups of inventors. The first group is "stayers": the inventors who produce at least one patent in a firm over the three years before and three years after the firm's IPO respectively. The second group of inventors is "leavers": the inventors who produce at least one patent in a firm over the three years before the firm's IPO and at least one patent in a different firm over the three years after the IPO.

If retaining inventors is a way through which IPO firms can better manage their human capital, we would expect to observe that inventors are less likely to leave the firm after IPO. By combining both "stayers" and "leavers" in our sample, we intend to focus on inventors that produce patents before the IPO, and examine the likelihood that these inventors leave the firm. We run the following model:

$$Leaver_l = \alpha + \beta Treat_l + \gamma X_{i,-1} + year_t + industry_j + \varepsilon_{it} \quad (6)$$

[‡‡‡] In China's patent application data, there is no identification number granted to each individual inventor as in the US patent data. Therefore, we use the name of the inventor and her patent's technology field (IPC) to identify each individual inventor. For instance, if an inventor named A filed a patent in the wireless communications technology before, and if a new patent is filed also in wireless communications technology with inventor named A, we would deem that it is the same inventor, who has filed the new patent. However, if the patent is in chemical, we would take the new patent filed by a different inventor who also named A.

where *l* indexes inventor, *i* indexes firm, *j* indexes industry, and *t* indexes time. $Leaver_l$ is a dummy variable, which equals to one if an inventor belongs to "leavers" and zero, otherwise. $Treat_l$ is a dummy variable indicating if the inventor *l* works for a treatment group firm before IPO (one if she does and zero, otherwise). $X_{i,-1}$ represents firm level control variables one year before the IPO. $year_t$ and $industry_j$ control for year and industry fixed effects. Results in Table 7 Panel B show that inventors in IPO firms are 69.2% less likely to leave the firm during the three years after the IPO, with the coefficient statistically significant at the 1% level. This result confirms our hypothesis that a firm's capacity to retain their talents has substantially improved after the IPO.

In addition to testing the inventor hypothesis, we also conduct similar DiD test on the total patent (invention patent) applications per inventor, and find statistically insignificant results, indicating that the productivity of inventors does not change much after the IPO. All these evidences indicate that the increase in patents, both invention and non-invention, are driven mainly by the increase in the number of inventors, rather than the increase of inventors' productivity.

6. The economic implication of the innovation activity

Our results so far suggest that IPO can trigger more active innovation in firms. It is time to ask the "bottom line" question: whether the enhanced innovation activity is value-creating or value-destroying? On the one hand, empirical studies indicate that innovation is positively associated with firm value (Hall, Jaffe, and Trajtenberg, 2005). On the other hand, there is a literature arguing that overinvestment in innovation could be damaging to firm value. For instance, Hirshleifer, Low and Teoh (2012) and Gompers (1996) argue that overconfident managers, who chase short term interest, could overinvest in innovation that may not serve the best interest of shareholders. In our setting, it could be the case that managers simply use patent applications as a "window dressing" tool to cajole investors and boost firm valuation in the short run. It therefore does not have a long-run impact on the firm value. Which story is true? Our uniquely matched data can provide an answer.

To tackle the "bottom-line" question regarding the economic consequences of innovation, we focus only on the IPO firms, and aggregate all subsidiary firms' patent applications at the parent firm level. Following the literature, we use Tobin's Q to measure firm value. Tobin's Q is measured as the summation of market value of equity and book value of debt, divided by book value of total assets. Because it usually takes time for the real impact of innovation to be reflected into firm valuation, to avoid the short term market noise we use $\Delta Q_{i,t+3\rightarrow t+6}$, the change in Tobin's Q from the end of the third year after the IPO to the end of the sixth year after the IPO, to measure the increase in firm value in the long run. We then estimate the following model:

$$\Delta LnQ_{i,t+3\rightarrow t+6} = \alpha + \beta Pat_{i[1,3]}(Inv_{i[1,3]}) + \gamma X_{i,3} + year_t + \varepsilon_{it} \quad (7)$$

where *i* index firm, and t indexes year. $Pat_{i[1,3]}$ $(Inv_{i[1,3]})$ is firm i's total patent (invention patent) filing over the three years after IPO. $X_{i,3}$ represents a group of firm-level control variables measured at the end of the third year after the IPO, including the natural logarithm of Tobin's Q, ROA (net income divided by total assets), leverage (book value of total debt divided by total assets, asset (book value), and the ratio of fixed asset. In addition, we also control for the firm's stock return over the three years after the IPO. Finally, $year_t$ captures the year fixed effect.

We report the results in Table 8. As can be seen that the coefficient estimate for $pat_{i[1,3]}$ in column (1) is positive and statistically significant, indicating that firms which file for more patents over the three years after the IPO, tend to experience a higher increase in their Tobin's Q in the long run. This result shows that the innovation activity after the IPO indeed creates value for the firm in the long run. In addition, we find that the coefficient estimate for $Inv_{i[1,3]}$ in column (2) is also positive and statistically significant, indicating that the invention patent applications after the IPO is associated with firm long-term value creation as well. Moreover, by comparing the coefficient estimates of $pat_{i[1,3]}$ and $Inv_{i[1,3]}$, it can be seen that the invention patent filing (higher quality innovation activity) has a much stronger impact on firm value creation. Combining our previous results that IPO increases both the quantity and quality of firm innovation, it can be argued that IPO creates real value for the firm through boosting firm innovation activity in China. And the impact mostly comes from the enhanced quality of innovation after the IPO. Our results therefore show evidence against the "window dressing" hypothesis.

7. Conclusions

In this paper, we empirically investigate the real impact of IPO on firm innovation activity using a uniquely matched firm-level data in China linking patent filing and stock market information. To mitigate endogeneity and establish causality, we use a difference-in-difference approach, together with the propensity score matching algorithm to construct the treatment and control group. We find a significant positive causal effect of the IPO on firms' patent applications, both in quantity and quality. We also find that, after IPO, firms strengthen their innovation in core business and expand innovation into none core business as well.

We show that the impact of IPO on innovation varies across firm corporate governance structure, financial constraint and ownership. In particular, firms with governance structure better aligning the interest of managers and stockholders, facing tighter financial constraint, and owned by non-state see a larger increase in innovation after the IPO. Moreover, by using the inventor information in the Chinese patent data, we find that IPO allows a firm to not only better retaining internal inventors, but also hiring external inventors, which could be an important factor driving the enhanced innovation activity. Finally, we find that the increase in innovation activities creates value for a firm in the long run, which confirms that innovation has a *real* impact on Chinese firms.

Reference

Aghion, Philippe, and Peter Howitt, 1992, A model of growth through creative destruction, Econometrica LX, 323-351.

Aghion, Philippe, and Jean Tirole, 1994, The management of innovation, Quarterly Journal of Economics 109, 1185–1209.

Aghion, Philippe, John VanReenen, and Luigi Zingales, 2013, Innovation and institutional ownership, American Economic Review 103, 277–304.

Agrawal, Ajay K., Iain M. Cockburn, Alberto Galasso, and Alexander Oettl, 2012, Why are some regions more innovative than others? The role of firm size diversity, NBER Working paper.

Akcigit, Ufuk, and William R. Kerr, 2013, Growth through heterogeneous innovations, NBER Working paper.

Asker, John, Joan Farre-Mensa, and Alexander Ljungqvist, 2014, Corporate investment and stock market listing: A puzzle?, Working paper, New York University.

Atanassov, Julian, Vikram Nanda, and Amit Seru 2007, Finance and innovation: The case of publicly traded firms. Working Paper, University of Oregon.

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

Berle, Adolf A., and Gardiner C. Means, 1932, The Modern Corporation and Private Property (Macmillan, New York).

Bertrand, Marianne, and Sendhil Mullainathan, 2003, Enjoying the quiet life? Corporate governance and managerial preferences, Journal of Political Economy 111, 1043–1075.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing, Journal of Development Economics 97.2, 339-351.

Brau, James C., and Stanley E. Fawcett, 2006, Initial public offerings: An analysis of theory and practice, Journal of Finance 61, 399–436.

Brav, Alon, and Paul A. Gompers, 1997, Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies, Journal of Finance 52, 1791–1821.

Brav, Omer, 2009, Access to capital, capital structure, and the funding of the firm, Journal of Finance 64, 263–308.

Cai, Hongbin and Qiao Liu (2009),"Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms," Economic Journal 119, pp.764-795.

Chemmanur, Thomas J., Shan He, and Debarshi K. Nandy, 2010, The going-public decision and the product market, Review of Financial Studies 23, 1855–1908.

Dean, Judith M., and Mary E. Lovely, 2010, Trade growth, production fragmentation, and China's environment, China's Growing Role in World Trade, University of Chicago Press, 429-469.

Gao, Huasheng, Jarrad Harford, and Kai Li, 2014, Investor horizon and CEO turnover-performance sensitivity, Working paper, University of British Columbia.

Gao, Huasheng, Po-Hsuan Hsu, and Kai Li, 2014, Public equity markets and corporate innovation strategies: Evidence from private firms, Working paper, University of British Columbia.

Gu, Yuqi, Connie X. Mao, and Xuan Tian, 2013, Creditor interventions and firm innovation: Evidence from debt covenant violations. working paper.

Hall, Bronwyn H., Zvi Griliches, and Jerry A. Hausman, 1986, Patents and R&D: Is there a lag? International Economic Review 27, 265–283.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, 2005, Market value and patent citations,

RAND Journal of Economics 36, 16–38.

Hall, Bronwyn H., and Josh Lerner, 2010, The financing of R&D and innovation, in Bronwyn H.

Hall and Nathan Rosenberg, eds.: Handbook of the Economics of Innovation, (Elsevier-North, Oxford, UK).

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

Hirshleifer, David, Angie Low, and Siew Hong Teoh, 2012, Are overconfident CEOs better innovators?, The Journal of Finance 67.4, 1457-1498.

Holmstrom, Bengt, 1982, Managerial incentive problems—A dynamic perspective, in Essays in Economics and Management in Honor of Lars Wahlbeck (Swedish School of Economics, Helsinki).

Holmstrom Bengt, 1989, Agency costs and innovation, Journal of Economic Behavior & Organization, 12(3): 305-327.

Hsieh, Chang-Tai, and Peter J. Klenow, 2009, Misallocation and Manufacturing TFP in China and India, The Quarterly Journal of Economics 124.4, 1403-1448.

Ibbotson, Roger, and Jay R. Ritter, 1995, Initial public offerings, in Robert A. Jarrow, Vojislav Maksimovic, and William T. Ziemba, eds.: Handbooks of Operations Research and Management

Science (North-Holland, Amsterdam).

Jain, Bharat A., and Omesh Kini, 1994, The post-issue operating performance of IPO firms, Journal of Finance 49, 1699–1726.

Jensen, Michael, and William H. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, Journal of Financial Economics 3, 305–360.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2012, Technological innovation, resource allocation, and growth, NBER Working paper.

Loughran, Tim, and Jay R. Ritter, 2004, Why has IPO underpricing increased over time?, Financial Management 33, 5–37.

Michaely, Roni, and Michael R. Roberts, 2012, Corporate dividend policies: Lessons from private firms, Review of Financial Studies 25, 711–746.

Mikkelson, Wayne H., Megan M. Partch, and Kshitij Shaw, 1997, Ownership and operating decisions of companies that go public, Journal of Financial Economics 44, 281–307. Moser, Petra, 2007, Why don't inventors patent? NBER Working paper.

Nanda, Ramana, and Matthew Rhodes-Kropf, 2013, Investment cycles and startup innovation, Journal of Financial Economics 110, 403-418.

Pagano, Marco, Fabio Panetta, and Luigi Zingales, 1996, The stock market as a source of capital: Some lessons from initial public offerings in Italy, European Economic Review 40, 1057–1069.

Pagano, Marco, Fabio Panetta, and Luigi Zingales, 1998, Why do firms go public? An empirical analysis, Journal of Finance 53, 27–64.

Pástor, Ľuboš, Lucian A. Taylor, and Pietro Veronesi, 2009, Entrepreneurial learning, the IPO decision, and the post-IPO drop in firm profitability, Review of Financial Studies 22, 3005–3046.

Porter, Michael, 1992, Capital disadvantage: America's failing capital investment system. Harvard Business Review 70, 65-82.

Robinson, David, 2008, Strategic alliances and the boundaries of the firm, Review of Financial Studies 21, 649–681.

Saunders, Anthony, and Sascha Steffen, 2011, The costs of being private: Evidence from the loan market, *Review of Financial Studies* 24, 4091–4122.

Seru, Amit, 2014, Firm boundaries matter: Evidence from conglomerates and R&D activity, *Journal of Financial Economics* 111, 381–405.

Sheen, Albert, 2009, Do public and private firms behave differently? An examination of investment in the chemical industry, Working paper, UCLA.

Schumpeter, Joseph, 1943, *Capitalism, Socialism and Democracy* (London: Allen Unwin)

Schmoch, Ulrich, et al, 2003, Linking technology areas to industrial sectors, *Final Report to the European Commission,* DG Research 1.

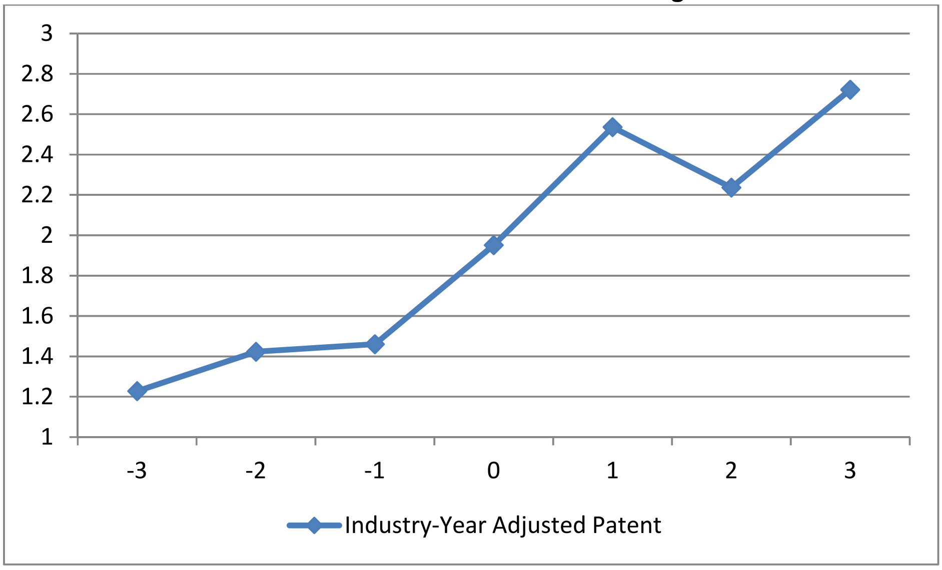
Solow, Robert M., 1957, Technical change and the aggregate production function, *Review of Economics and Statistics* 39, 312–320.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, 2011, Growing Like China, *American Economic Review* 101(1), 196.233.

Figure 1： Industry and year-adjusted innovation output around IPO

This figure presents firms' average number of industry and year-adjusted patent applications around the IPO year. We adjust each firm's patent application by subtracting the average number of patent applications of all firms in the same 2-digit industry and year. Panel A shows that the time trend for total patent counts (including invention, design and utility patents) and Panel B shows the time trend for invention patent counts only. The x-axis indicates the number of years before or after IPO.

Panel A: Number of total Patent filing



Panel B: Number of invention Patent filing

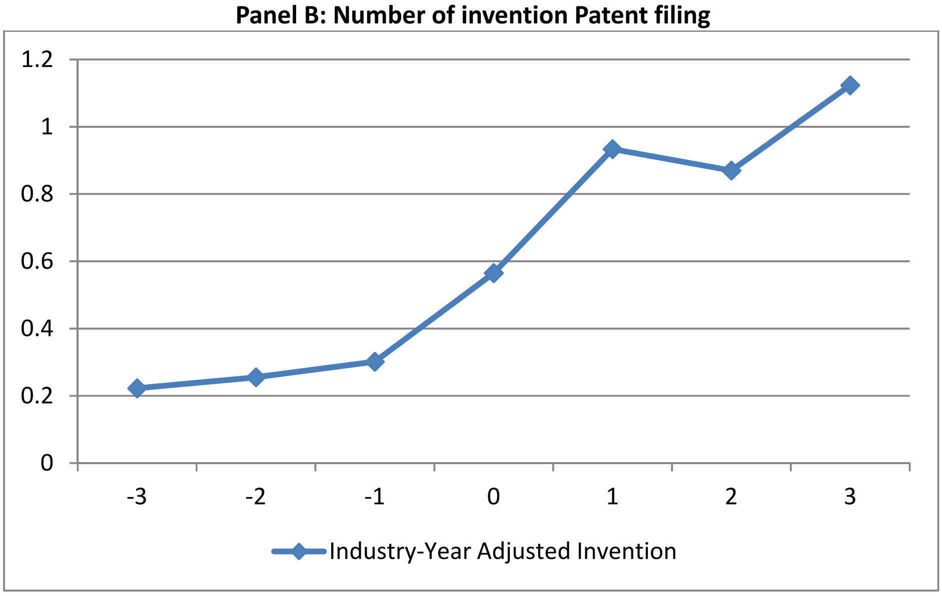
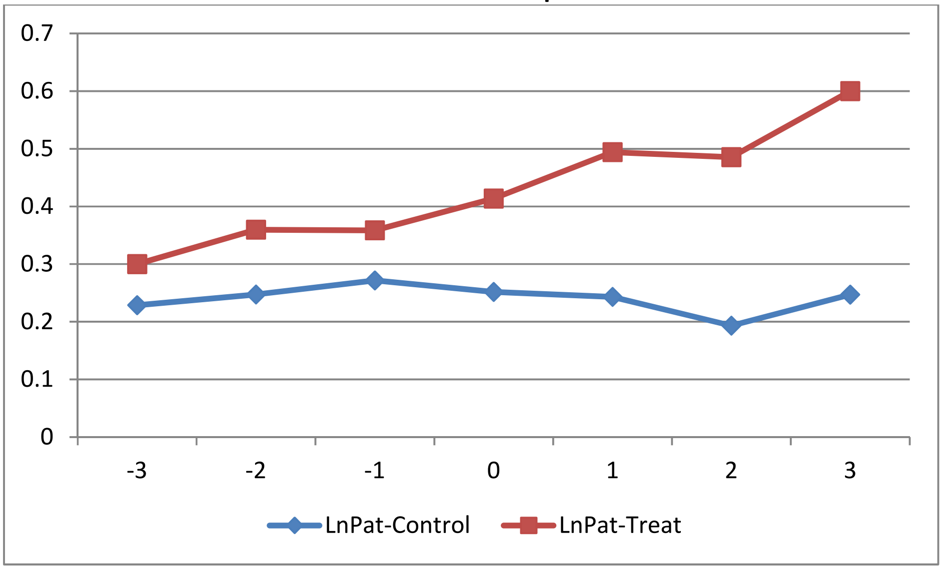


Figure 2: Innovation outputs around IPO for treatment and control firms

This figure presents mean values of innovation outputs around IPO. Panel A shows the time trend for the natural logarithm of one plus patent counts and Panel B shows the time trend for the natural logarithm of one plus invention counts. The square line is for treatment firms that incurred a IPO in a sample year. The diamond line is for propensity-score-matched non-list firms (control group). The x-axis indicates the years before or after IPO.

Panel A: Number of patents



Panel B: Number of inventions patent application

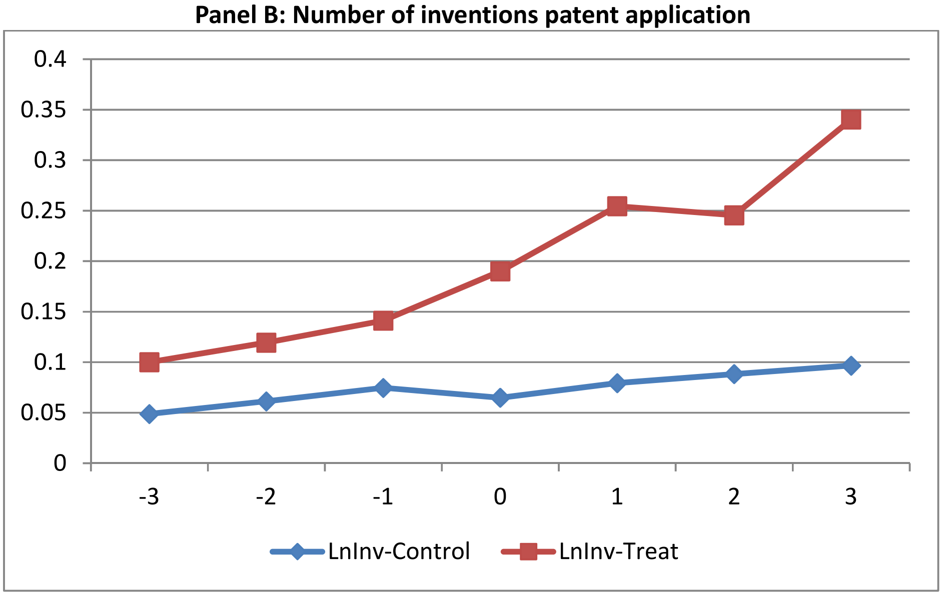


Table 1: Summary statistics

This table presents summary statistics for variables constructed on our merged data sample of listed and non-listed firms from 1998 to 2007. All variables are winsorized at the 1^{st} and 99^{th} percentile of their distribution.

	N	Mean	S.D.	P25	Median	P75
Pat	999114	0.18	1.91	0.00	0.00	0.00
Invention	999114	0.03	0.62	0.00	0.00	0.00
Ln(AT)	999114	10.04	1.37	9.06	9.87	10.87
FixAT_AT	999114	0.33	0.20	0.17	0.30	0.45
Interest_AT	999114	0.01	0.02	0.00	0.01	0.02
Admin_AT	999114	0.08	0.08	0.03	0.06	0.10
Ln(Age)	999114	2.22	0.80	1.79	2.20	2.71
Leverage	999114	0.59	0.27	0.40	0.60	0.78
Liquidity	999114	0.06	0.29	-0.11	0.06	0.24
ROA	999114	0.08	0.15	0.00	0.03	0.10
HI	999114	0.05	0.07	0.01	0.02	0.05
HI^2	999114	0.01	0.05	0.00	0.00	0.00
SOE	999114	0.12	0.33	0.00	0.00	0.00
EX	999114	0.34	0.47	0.00	0.00	1.00
Ln(Stock)	999114	0.18	0.58	0.00	0.00	0.00

Table 2 Panel A: Raw patterns on innovation dynamics surrounding IPO

This table reports the OLS regression results that estimate the innovation dynamics surrounding IPO. Using a sample of all listed firms, we retain firm-year observations for a seven-year window centered in the IPO year and we estimate the pooled OLS regression of the following model:

$$\mathrm{Pat}_{it}(\mathrm{Inv}_{it}) = \alpha + \beta_1 before_{it}^{-3} + \beta_2 before_{it}^{-2} + \beta_3 current_{it} + \beta_4 after_{it}^{1} + \beta_5 after_{it}^{2} + \beta_6 after_{it}^{3} + \varepsilon_{it}$$

The dependent variable is eitherPat_{it}, firm i's industry and year adjusted total patents filed in year t, or Inv_{it}, firm i's industry and year adjusted invention patents filing in year t. We adjust innovation output variables by subtracting the average values of innovation outputs of all firms (excluding the listed firms) in the same 2-digit industry and year. $before_{it}^{-3}$ ($before_{it}^{-2}$) is a dummy that equals one if a firm-year observation is from 3 (2) year before the IPO year, and zero otherwise. $current_{it}$ is a dummy that equals one if a firm-year observation is in its IPO year and, zero otherwise. $after_{it}^{1}$ ($after_{it}^{2}$, $after_{it}^{3}$) is a dummy that equals one if a firm-year observation is 1 (2,3) year after the IPO year, and zero otherwise. Therefore, the omitted group (benchmark) is the observations 1 year before the IPO year. Standard errors clustered by industry are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

VARIABLES	(1) Pat	(2) Inv
before^3	-0.262	-0.086
	(0.313)	(0.102)
before^2	-0.054	-0.045
	(0.218)	(0.069)
current	0.433*	0.246*
	(0.220)	(0.123)
after^1	1.012**	0.610***
	(0.423)	(0.212)
after^2	0.757**	0.561***
	(0.292)	(0.197)
after^3	1.255**	0.820**
	(0.532)	(0.315)
Constant	1.465***	0.302***
	(0.369)	(0.061)
Observations	2,118	2,118
R-squared	0.005	0.012

Table 2 Panel B: OLS regression of innovation outcomes on IPO

This table reports the OLS results estimating the effect of IPO on innovation output variables. We estimate the pooled OLS regression of the following model:

$$\text{LnPat}_{it+\text{n}}(\text{LnInv}_{it+\text{n}}) = \alpha + \beta list_{it} + \gamma X_{it} + year_t + industry_j + \varepsilon_{it}$$

using a sample of all listed and non-listed firms from 1998 to 2007. The dependent variable LnPat_{it+n} is the natural logarithm of one plus total number of patents applied in one (t+1), two (t+2), and three (t+3) years after year t and results are reported in columns (1) – (3) respectively. The dependent variable LnInv_{it+n} is the natural logarithm of one plus the number of inventions patents applied in one (t+1), two (t+2), and three (t+3) years after year t and results are reported in columns (4) – (6), respectively. list_{it} is a dummy variable that equals one if IPO occurs in year *t* for firm *i*, and zero otherwise. Year fixed effects year_{it} and industry fixed effects industry_{it} are included in all regressions. All other variables are as defined in the Appendix. Standard errors clustered by industry are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

VARIABLES	(1) LnPat_+1	(2) LnPat_+2	(3) LnPat_+3	(4) LnInv_+1	(5) LnInv_+2	(6) LnInv_+3
List	**0.336*** **	**0.312*** **	**0.406*** **	**0.205*** **	**0.183*** **	**0.263*** **
	(0.071)	**(0.053)**	**(0.093)**	**(0.041)**	**(0.038)**	**(0.076)**
Ln(AT)	0.048***	0.050***	0.052***	0.016***	0.017***	0.019***
	(0.007)	(0.007)	(0.008)	(0.003)	(0.003)	(0.003)
FixAT_AT	-0.066***	-0.071***	-0.079***	-0.017***	-0.022***	-0.025***
	(0.013)	(0.014)	(0.015)	(0.004)	(0.005)	(0.006)
Admin_a	0.214***	0.223***	0.235***	0.062***	0.066***	0.074***
	(0.035)	(0.037)	(0.039)	(0.014)	(0.015)	(0.016)
Ln(Age)	-0.004**	-0.005***	-0.005***	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
SOE	0.008	0.009	0.010	0.006**	0.006**	0.006**
	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)
EX	0.025***	0.026***	0.026***	0.006***	0.007***	0.007**
	(0.006)	(0.007)	(0.007)	(0.002)	(0.002)	(0.003)
Leverage	-0.024***	-0.023***	-0.024***	-0.007***	-0.009***	-0.009***
	(0.004)	(0.004)	(0.005)	(0.002)	(0.002)	(0.002)
Liquidity	-0.007*	-0.007*	-0.008*	0.001	-0.000	-0.001
	(0.004)	(0.004)	(0.004)	(0.001)	(0.002)	(0.002)
Interest_At	-0.033	-0.056*	-0.047	0.036**	0.030	0.029
	(0.032)	(0.033)	(0.037)	(0.017)	(0.019)	(0.019)
ROA	0.053***	0.068***	0.088***	0.020***	0.028***	0.038***
	(0.009)	(0.011)	(0.015)	(0.004)	(0.006)	(0.008)
HI	-0.034	-0.037	-0.050*	-0.003	-0.006	-0.009
	(0.026)	(0.027)	(0.029)	(0.011)	(0.013)	(0.013)
HI^2	0.066*	0.076**	0.081**	0.023	0.029	0.028
	(0.034)	(0.036)	(0.038)	(0.020)	(0.023)	(0.025)
Constant	-0.440***	-0.459***	-0.482***	-0.159***	-0.171***	-0.186***
	(0.065)	(0.069)	(0.072)	(0.028)	(0.031)	(0.032)
Observations	839,761	709,997	580,425	839,761	709,997	580,425
R-squared	0.064	0.065	0.067	0.039	0.040	0.043
Year fixed effect	Y	Y	Y	Y	Y	Y
Ind fixed effect	Y	Y	Y	Y	Y	Y
Firm fixed effect	N	N	N	N	N	N

Standard errors in parentheses
*** p<0.01, ** p<0.05, *p<0.1

Table 3: Difference-in-differences (DiD) test results

This table reports diagnostic tests and the DiD results on how IPO affect firm innovation. Our sample contains firms that experienced IPO from 1999 to 2006 and non-listed firms. We match firms using a one-to-one nearest neighbor propensity score matching, with replacement, on a host of observable characteristics including all independent variables used in equation (2) for the year before the IPO, the growth in the number of patents g_pn, computed over the three-year period before IPO, 2-digit industry dummies, and year fixed effects. Definitions of all other variables are listed in the Appendix. Our treatment group contains those listed firms' observations at their IPO year. Our control group includes all non-listed firm-year observations. Panel A reports parameter estimates from the probit model used in estimating the propensity scores for the treatment and control groups. The dependent variable equals one for the firm-year belonging to the treatment group and zero for those belonging to the control group. The "Pre-Match" column contains the parameter estimates of the probit model estimated using the sample prior to matching. These estimates are then used to generate the propensity scores for matching. The "Post-Match" column contains the parameter estimates of the probit model estimated using the subsample of matched treatment-control pairs after matching. Robust standard errors are displayed in parentheses below each coefficient estimate. Panel B presents the univariate comparisons between the treatment and control firms' characteristics and their corresponding p-value testing the null hypothesis that the differences are zero. Panel C gives the DiD test results. *Patent* is the mean of firm i's number of patents in the three-year window before or after IPO. *Invention* is the average of firm i's number of inventions in the three-year window before or after IPO. Panel D reports the regression results that estimate the innovation dynamics of treatment and control firms surrounding IPO. The dependent variable is either $\ln \mathrm{Pat}_{\mathrm{it}}$, the natural logarithm of one plus firm i's number of patents in year t, or $\ln \mathrm{Inv}_{\mathrm{it}}$, the natural logarithm of one plus firm i's number of inventions in year t. $\mathrm{list}_{\mathrm{i}}$ is a dummy that equals one for treatment firms (listed firms) and zero for control firms (non-listed firms). $\mathrm{before}_{\mathrm{it}}^{2\&3}$ is a dummy that equals one if a firm-year observation is from 2 or 3 year before the IPO year (year 0) and zero otherwise. $\mathrm{current}_{\mathrm{it}}$ is a dummy that equals one if a firm-year observation is in the IPO year (year 0) and zero otherwise. $\mathrm{after}_{\mathrm{it}}^{1}$ ($\mathrm{after}_{\mathrm{it}}^{2}$, $\mathrm{after}_{\mathrm{it}}^{3}$) is a dummy that equals one if a firm-year observation is from 1 (2,3) year after the IPO year (year 0) and zero otherwise. Robust standard errors are displayed in parentheses under each coefficient estimate. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Table 3 Panel A: Pre-match propensity score regression and post-match diagnostic regression

VARIABLES	(1) Pre-match	(2) Post-match
Ln(AT)	0.256***	-0.046
	(0.013)	(0.043)
FixAT_AT	-0.373***	0.045
	(0.100)	(0.365)
Admin_AT	0.863***	0.141
	(0.178)	(0.992)
Ln(Age)	-0.185***	0.000
	(0.021)	(0.054)
SOE	0.243***	0.061
	(0.043)	(0.112)
EX	-0.081**	-0.007
	(0.036)	(0.113)
Leverage	-0.284***	0.048
	(0.086)	(0.333)
Liquidity	-0.006	-0.074
	(0.096)	(0.298)
Interes_AT	0.915	1.645
	(0.980)	(3.962)
ROA	0.644***	0.499
	(0.086)	(0.510)
HI	0.151	0.125
	(0.420)	(1.265)
HI^2	-0.599	-1.046
	(0.778)	(2.248)
Stock	0.098***	0.077
	(0.019)	(0.050)
g_pn	0.050	0.045
	(0.030)	(0.077)
Constant	-5.441***	0.408
	(0.183)	(0.627)
Observations	773,090	710
P-value of Chi-square	0.0000	1.0000
Pseudo-R^2	0.1652	0.0066
Industry fixed effect	Y	Y
Year fixed effect	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3 Panel B: Post-Match differences

	Control	Treat	Treat-Control	P-value
Ln(AT)	11.94	11.86	-0.08	0.45
FixAT_AT	0.29	0.29	0.00	0.85
Admin_AT	0.06	0.07	0.00	0.34
Ln(Age)	1.99	2.01	0.01	0.86
SOE	0.37	0.37	0.01	0.88
EX	0.42	0.42	0.00	1.00
Leverage	0.53	0.53	0.00	0.87
Liquidity	0.11	0.11	0.00	0.81
Interest_AT	0.01	0.01	0.00	0.63
ROA	0.08	0.09	0.01	0.36
HI	0.06	0.06	0.00	0.67
HI^2	0.01	0.01	0.00	0.53
Stock	0.57	0.69	0.12	0.16
g_pn	0.12	0.14	0.04	0.40

Table 3 Panel C: DID estimates

	Mean Treatment Difference (after-before)	Mean Control Difference (after-before)	Mean DID estimate (treat-control)	P-value for DID Estiamte
Patent	1.577	0.014	1.563	0.001
(s.e.)	0.405	0.262	0.482	
Invention	0.858	0.086	0.772	0.000
(s.e.)	0.198	0.072	0.210	

Table 3 Panel D: DiD analysis for innovation dynamics

	(1)	(2)
VARIABLES	LnPat	LnInv
$list_i * before_{it}^{2\&3}$	-0.043	-0.066
	(0.052)	(0.041)
$list_i * current$	0.075	0.043
	(0.050)	(0.031)
$list_i * after_{it}^{1}$	**0.164*** **	**0.093*** **
	(0.052)	**(0.033)**
$list_i * after_{it}^{2}$	**0.231*** **	**0.101*** **
	(0.054)	**(0.034)**
$list_i * after_{it}^{3}$	**0.291*** **	**0.192*** **

	(0.062)	**(0.042)**
$before_{it}^{2\&3}$	0.080**	0.043*
	(0.038)	(0.023)
$before_{it}^{1}$	0.116***	0.062***
	(0.044)	(0.024)
current	0.096**	0.047*
	(0.042)	(0.024)
$after_{it}^{1}$	**0.096****	**0.068*****
	(0.043)	**(0.024)**
$after_{it}^{2}$	**0.060**	**0.071****
	(0.047)	**(0.029)**
$after_{it}^{3}$	**0.127****	**0.078*****
	(0.050)	**(0.028)**
Constant	0.173***	0.064**
	(0.044)	(0.025)
Observations	4,253	4,253
R-squared	0.655	0.576
Year fixed effect	Y	Y
Firm fixed effect	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Difference in Difference (DiD) test results on innovation quality

This table reports the DiD test results on how IPO affects the innovation quality. We use the propensity-score-matched sample and retain firm-year observations for both treatment and control group for a 7-year window centered around IPO year. Panel A runs a regression with the dependent variable *inventor_per_patent* equals to average number of inventors per patent for firm *i* in year *t*. Panel B reports the univariate DiD test results for Inv_ratio, which is the ratio of invention patent over total patent number over the three-year window before and after IPO.

Table 4 Panel A: The inventor per patent regression

VARIABLES	(1) inventor_per_patent
$list_i * before_{it}^{2\&3}$	-0.102
	(0.083)
$list_i * current$	0.075
	(0.091)
$list_i * after_{it}^{1}$	**0.200***
	(0.106)
$list_i * after_{it}^{2}$	**0.299****
	(0.139)
$list_i * after_{it}^{3}$	**0.433****
	(0.180)
$before_{it}^{2\&3}$	0.071
	(0.060)
$before_{it}^{1}$	0.064
	(0.077)

$current$	-0.035
	(0.077)
$after_{it}^{1}$	-0.107
	(0.087)
$after_{it}^{2}$	-0.077
	(0.110)
$after_{it}^{3}$	-0.282*
	(0.161)
Constant	2.251***
$list_i * before_{it}^{2\&3}$	(0.092)
Observations	9,088
R-squared	0.611
Year FE	Y
Firm FE	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 Panel B: Univerate DiD estimates

	Mean Treatment Difference (after-before)	Mean Control Difference (after-before)	Mean DID estimate (treat-control)	P-value for DID Estimate
Inv_ratio	0.077***	0.033**	**0.043***	**0.068**
(s.e.)	0.018	0.016	0.024	

Table 5: DiD analysis for innovation scope

This table reports the regression results that investigate the effect of IPO on the expanding of innovation scope, and then estimate the related (unrelated) innovation dynamics of treatment and control firms surrounding IPO. We use the propensity-score-matched sample and retain firm-year observations for both treatment and control group for a 7-year window centered at IPO year. We estimate the following model:

$$y_{it} = \alpha + \beta_1 list_i * before_{it}^{2\&3} + \beta_2 list_i * current_{it} + \beta_3 list_i * after_{it}^1 + \beta_4 list_i * after_{it}^2 + \beta_5 list_i * after_{it}^3 + \beta_6 before_{it}^2 + before_{it}^1 + current_{it} + after_{it}^1 + after_{it}^2 + after_{it}^3 + year_t + firm_i + \varepsilon_{it}$$

The dependent variable y_{it} can be $Dunrelated_{it}$, which is a dummy that equals one if firm i have unrelated patents in year t and zero otherwise, $UnrelatedClass_{it}$, the number of technology class that firm i's unrelated patents belong to, $lnRelatedPat_{it}$, the natural logarithm of one plus firm i's number of related patents in year t, $lnUnrelatedPat_{it}$, the natural logarithm of one plus firm i's number of unrelared patents in year t, $lnRelatedInv_{it}$, the natural logarithm of one plus firm i's number of related inventions in year t, and $lnUnrelatedInv_{it}$, the natural logarithm of one plus firm i's number of unrelated inventions in year t. Patent's technology class is defined based on patent's IPC. We describe this detailed procedure in Section 4.3. Related innovation are those related to a firm's core business and unrelated innovation are those unrelated to a firm's core business. $\mathrm{list_i}$ is a dummy that equals one for treatment firms (listed firms) and zero for control firms (non-listed firms). $\mathrm{before_{it}^{2,3}}$ is a dummy that equals one if a firm-year observation is from 2 or 3 year before the IPO year (year 0) and zero otherwise. $\mathrm{current_{it}}$ is a dummy that equals one if a firm-year observation is in the IPO year (year 0) and zero otherwise. $\mathrm{after_{it}^1}$ ($\mathrm{after_{it}^2}$, $\mathrm{after_{it}^3}$) is a dummy that equals one if a firm-year observation is from 1 (2,3) year after the IPO year (year 0) and zero otherwise. All specifications include year and firm fixed effects. Robust standard errors are displayed in parentheses under each coefficient estimate. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Dunrelated	UnrelatedClass	LnrRelatedPat	LnUnrelatedPat	LnRelatedInv	LnUnrelatedInv
$list_i * before^{23}$	-0.005	-0.069	-0.036	-0.035	-0.040	-0.034
	(0.023)	(0.062)	(0.033)	(0.034)	(0.026)	(0.023)
$list_i * current$	**0.031**	**0.076**	**0.016**	**0.047**	**0.028**	**0.031**
	(0.025)	**(0.060)**	**(0.032)**	**(0.034)**	**(0.025)**	**(0.024)**
$list_i * after_{it}^1$	**0.054****	**0.085**	**0.081****	**0.076****	**0.069*****	**0.049***
	(0.025)	**(0.063)**	**(0.032)**	**(0.034)**	**(0.027)**	**(0.025)**
$list_i * after^2$	**0.061****	**0.217*****	**0.110*****	**0.113*****	**0.066****	**0.086*****
	(0.026)	**(0.068)**	**(0.036)**	**(0.036)**	**(0.029)**	**(0.026)**
$list_i * after^3$	**0.115*****	**0.337*****	**0.143*****	**0.198*****	**0.091*****	**0.147*****
	(0.030)	**(0.074)**	**(0.041)**	**(0.039)**	**(0.033)**	**(0.031)**
before^2	0.045***	0.088*	0.042*	0.057**	0.033*	0.025
	(0.016)	(0.049)	(0.025)	(0.026)	(0.020)	(0.016)
before^1	0.060***	0.063	0.059**	0.060**	0.039**	0.034**
	(0.019)	(0.053)	(0.024)	(0.027)	(0.018)	(0.017)
current	0.063***	0.068	0.062**	0.064**	0.040**	0.024
	(0.018)	(0.052)	(0.025)	(0.027)	(0.019)	(0.017)
after^1	0.063***	0.105*	0.048**	0.068**	0.043**	0.052***
	(0.020)	(0.056)	(0.022)	(0.028)	(0.018)	(0.019)
after^2	0.044**	0.022	0.058**	0.039	0.053**	0.040*
	(0.021)	(0.057)	(0.028)	(0.032)	(0.024)	(0.021)

$after^3$	0.057***	0.038	0.078***	0.035	0.063***	0.035*
	(0.022)	(0.056)	(0.027)	(0.030)	(0.022)	(0.020)
Constant	0.052***	0.158***	0.044*	0.097***	0.029	0.046**
	(0.018)	(0.053)	(0.024)	(0.028)	(0.019)	(0.018)
Observations	4,253	4,253	4,253	4,253	4,253	4,253
R-squared	0.536	0.588	0.571	0.650	0.501	0.556
Firm fixed effect	Y	Y	Y	Y	Y	Y
Year fixed effect	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Cross-section analysis of IPO's impact on innovation output

This table reports the effects of IPO on firm innovation output across firms with different characteristics. We use the propensity-score-matched sample and retain firm-year observations for both treatment and control group for a 7-year window centered around IPO year. We focus on comparing the 3-year window after IPO with 3-year window before IPO, and we drop the IPO year. The estimation model is as follows.

$$\ln Pat_{it}\ (\ln Inv_{it}) = \alpha + \beta_1 list_i * \text{After}_{it}^{1\&2\&3} * X_i + \beta_2 list_i * After_{it}^{1\&2\&3} + \text{After}_{it}^{1\&2\&3} * X_i + \text{After}_{it}^{1\&2\&3} + firm_i + year_t + \varepsilon_{it}$$

The dependent variable is either $\ln \text{Pat}_{it}$, the natural logarithm of one plus firm i's number of patents in year t, or $\ln \text{Inv}_{it}$, the natural logarithm of one plus firm i's number of invention patents in year t. X_i could be dummy variable SOE_i, which equals one if the IPO firm *i* is a state owned enterprise before the IPO and zero otherwise; or a dummy variable $EXEHLD_i$, which equals one if the management holds stock after IPO and zero otherwise; or a dummy variable $Duality_i$ indicating if the CEO and president are the same person; or a dummy variable $Fixed$, which equals one if a firm with fixed asset ratio (FixAT_AT) above the median before IPO. $\text{After}_{it}^{1\&2\&3}$ is a dummy that equals one if a firm-year observation is after IPO. list_i is a dummy that equals one for treatment group For brevity, we only report the coefficient estimates of variables $list_i * \text{After}_{it}^{1\&2\&3} * X_i$. All specifications include year and firm fixed effects. Robust standard errors are displayed in parentheses under each coefficient estimate. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

VARIABLES	(1) EXEHLD Vs none EXEHLD LnPat	(2) EXEHLD Vs none EXEHLD LnInv	(3) Duality Vs None Duality LnPat	(4) Duality Vs None Duality LnInv	(5) High Fixed ratio Vs Low Fixed ratio LnPat	(6) High Fixed ratio Vs Low Fixed ratio LnInv	(7) SOE Vs POE LnPat	(8) SOE Vs POE LnInv
list$After^{1\&2\&3}$	0.232**	0.207***	0.343***	0.244***	0.293***	0.238***	0.341***	0.233***
	(0.095)	(0.063)	(0.075)	(0.052)	(0.045)	(0.039)	(0.047)	(0.031)
list$After^{1\&2\&3} * EXEHLD$	0.470***	0.280***						
	(0.144)	(0.103)						
list$After^{1\&2\&3} * Duality$			0.526**	0.477***				
			(0.222)	(0.157)				
list$After^{1\&2\&3} * SOE$							-0.257***	-0.198***
							(0.078)	(0.051)
list$After^{1\&2\&3} * Fixed$					-0.098**	-0.162***		
					(0.038)	(0.048)		
Observations	1,464	1,464	1,464	1,464	3,543	3,543	3,543	3,543
R-squared	0.672	0.574	0.672	0.578	0.655	0.587	0.655	0.586
Year fixed effect	Y	Y	Y	Y	Y	Y	Y	Y
Firm fixed effect	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Inventor mobility and innovative productivity

This table reports the effects of an IPO on inventors' mobility and innovative activity. We use the propensity-score-matched sample and retain firm-year observations for both treatment and control group for a 7-year window centered around IPO year. Inventors are classified into three categories: stayers, leavers, and newcomers, as defined in the text. Panel A gives the DID results for yearly average of inventor number across firms, and per inventor total patent and invention patent. Inventor is yearly average of firm i's inventor number in the 3-year window before or after IPO, which is calculated by total inventor number divided by the year of firm survival. Patent (Invention) is per inventor patents (inventions) of firm i in the 3-year window before or after IPO, which is measured by total patents (inventions) divided by total inventor of firm i. Panel B reports the effect of IPO on inventor's departure. We estimate the model below:

$$Leaver_l = \alpha + \beta Treat_{lt} + \gamma X_{it} + list\_year_t + industry_j + \varepsilon_{it}$$

We include stayer and leaver in our sample, and the dependent variable equals one if the inventor leaves firm, and zero otherwise. IPO is a dummy variable equal to one if a firm experiences the IPO, and zero otherwise. In all specifications we control firm characteristics before IPO and inventor's total patent during the 3 years before IPO as inventor's productivity. The model is estimated using Probit, and list year fixed and industry fixed effects are added. Robust standard errors are reported in parentheses. *,**, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level, respectively.

Table 7 Panel A: DiD estimates of inventor and innovative productivity

	Mean Treatment Difference (after-before)	Mean Control Difference (after-before)	Mean DID estimate (treat-control)	P-value for DID Estimate
Number of Inventors	3.461	0.445	3.016	0.001
(s.e.)	0.848	0.232	0.879	
per capita total patents	-0.145	-0.720	0.575	0.242
(s.e.)	0.247	0.424	0.491	
Per capita Invention patents	0.081	0.014	0.067	0.379
(s.e.)	0.072	0.024	0.076	

Table 7 Panel B: DiD analysis for inventor mobility

VARIABLES	leaver	leaver
Treat	**-0.692*****	**-0.719*****
	(0.164)	**(0.166)**
Pre_productivity		-0.017
		(0.013)
Ln(At)	0.223*	0.224*
	(0.125)	(0.125)
FixAT_AT	0.881	0.843
	(0.862)	(0.864)
Admin_AT	-1.389	-1.428
	(1.697)	(1.698)
Ln(Age)	-0.116	-0.121
	(0.108)	(0.109)
SOE	0.621***	0.593***
	(0.194)	(0.195)
EX	-0.060	-0.039
	(0.216)	(0.217)
Leverage	-0.749	-0.748
	(0.811)	(0.814)
Liquidity	-0.216	-0.215
	(0.763)	(0.765)
Interest_AT	-20.971*	-20.181*
	(11.992)	(12.049)
ROA	-1.214	-1.309
	(1.316)	(1.321)
HI	3.306	2.838
	(3.747)	(3.777)
HI^2	-12.445	-11.327
	(11.232)	(11.340)
Constant	-8.320	-8.303
	(138.972)	(139.139)
Observations	600	600
List Year FE	Y	Y
Industry FE	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: The real effect of innovation: innovation and firm value.

This table reports the effect of innovation on firm value. We focus only on the IPO firms, and aggregate all subsidiary firms' patent applications at the parent firm level. The accounting information is collected from CSMAR. We estimate the following model:

$$\Delta \mathrm{Ln}Q_{i,t+3\rightarrow t+6} = \alpha + \beta Pat_{i[1,3]} Inv_{i[1,3],} + \gamma X_{it+3} + year_t + \varepsilon_{it}$$

The dependent variable is $\Delta \mathrm{Ln}Q_{i,t+3\rightarrow t+6}$, the change in natural logarithm of Tobin's Q from the end of the third year after IPO to the end of the sixth year after IPO. $Pat_{i[1,3]}$ ($Inv_{i[1,3]}$) is firm i's total patent (invention patent) filing over the three years after IPO. The firm characteristic variables in the regression are measured by the end of the third year after IPO and the definition of these variables are in the Appendix. Robust standard errors are reported in parentheses. *,**, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level, respectively.

VARIABLES	(1) $\Delta \mathrm{Ln}Q_{i,t+3\rightarrow t+6}$	(2) $\Delta \mathrm{Ln}Q_{i,t+3\rightarrow t+6}$
$\boldsymbol{Pat_{i[1,3]}}$	**0.005***	
	(0.003)	
$\boldsymbol{Inv_{i[1,3]}}$		**0.014***
		(0.007)
$AccuRET_{i[1,3]}$	-0.073	-0.074
	(0.103)	(0.103)
Q_{i3}	-0.123	-0.112
	(0.159)	(0.159)
$Ln(AT)_{i3}$	-0.580***	-0.537***
	(0.181)	(0.178)
$FixAT\_AT_{i3}$	-0.326	-0.497
	(0.768)	(0.752)
$Leverage_{i3}$	1.864**	1.663**
	(0.756)	(0.758)
ROA_{i3}	-2.528	-3.344
	(3.167)	(3.164)
Constant	10.824***	10.103***
	(3.792)	(3.750)
Observations	148	148
R-squared	0.514	0.514
Year FE	Y	Y

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix

Innovation Measure	
Pat	Total number of patents applied in a given year
Invention	Total number of invention patents applied in a given year
Related patents	Number of patents that are related to firm's core business, i.e., the number of patents that are mapped to a firm's field which is defined mainly based on 2-digit CIC industry.
Unrelated patents	Number of patents that are unrelated to firm's core business, i.e., the number of patents that are not mapped to a firm's field which is defined mainly based on 2-digit CIC industry.
Firm Characteristics	
Ln(AT)	Natural logarithm of total assets
FixAT_AT	Net fixed assets scaled by book value of total assets
Admin_AT	Administration expenditure divided by book value of total assets
Ln(Age)	Natural logarithm of the number of years since the firm's establishment
SOE	A dummy equals 1 if a firm's registered capital held by the state exceeds 50% or the "controlling shareholder" identifies the state as its controlling holder.
EX	A dummy equals 1 if a firm has exports
Leverage	Book value of total debt divided by book value of total assets
Liquidity_AT	The difference of current assets and debt divided by total assets
Interest_AT	Interest expenditure divided by book value of total assets
ROA	Operating profit divided by book value of total assets
HI	Herfindahl index based on annual sales in the cell of 2-digit CIC industry and province
Ln(Stock)	Natural logarithm of firm's stock, stock is the sum of the patents a firm has applied before a given year (the given year included)
Growth of patent	Growth in the number of patent computed over the three-year period before IPO,LnPat_t-LnPat_it-2, t is the IPO year
Tobin's Q	the sum of the firm's market value of equity plus the book value of its debt divided by the firm's total assets