Benchmarking U.S. University Patent Value and Commercialization Efforts: A New Approach*

June 7, 2019

David H. Hsu Wharton School University of Pennsylvania dhsu@wharton.upenn.edu

Po-Hsuan Hsu Faculty of Business and Economics University of Hong Kong paulhsu@hku.hk

Tong Zhou Lingnan College Sun Yat-Sen University zhout59@mail.sysu.edu.cn

Arvids A. Ziedonis Faculty of Economics and Business KU Leuven ziedonis@kuleuven.be

^{*} We are grateful for helpful comments from Pierre Azoulay, Lauren Cohen, Stephen Dimmock, Chuck Eesley, Jeff Furman, Michelle Gittelman, Douglas Hanley, Thomas Hellmann, Yael Hochberg, William Kerr, Wonjoon Kim, Leonid Kogan, Josh Lerner, Kai Li, Willaim Mann, Kazuyuki Motohashi, Abhiroop Mukherjee, Ivan Png, Ariel Stern, Scott Stern, Neil Thompson, Kevin Tseng, Yanbo Wang, Yizhou Xiao, Yu Xu, and participants at the AIEA-NBER Conference, SFS Cavalcade Conference, and the Hong Kong Joint Finance Conference.

Benchmarking U.S. University Patent Value and Commercialization Efforts: A New Approach

June 7, 2019

Abstract: Despite the significance of patented university research, it is difficult to measure the economic value of their patented inventions and observe the extent to which universities are able to capture such value through patent licensing. Moving beyond assessing commercialization performance by simple statistics, we propose a new approach to benchmarking university patents and commercialization performance based on comparative corporate patent value. Our procedure involves matching university patents to patents granted to public corporations with similar patent characteristics to estimate the "potential value" of these university patents by the stock market reactions to matched corporate patent grants. We then calibrate an empirical patent valuation model for these estimated values of university patents by employing technology-level licensing data from a leading US research university. In aggregate, we compare the estimated potential values of a university's patent portfolio to its annual licensing income, and find that universities realize on average 5-9% of the estimated potential value through licensing income. Finally, we investigate the correlates of university-level potential patent value and suggest avenues for future research.

Keywords: university patents; patent value; patent licensing

JEL classification: G12; L30; O30

I. INTRODUCTION

According to the Association of University Technology Managers' (AUTM) 2015 survey of U.S. university technology transfer operations, its members filed 16,000 patent applications and received about 6,500 patent grants that year.¹ In addition, over 1,000 new ventures were formed. 879 new products based on university research were reported to have been introduced that year, and new and existing licensed products generated \$29B in product sales. Recent examples of influential scientific discoveries from university research include Emory University's HIV drug Emtricitabine, New York University's anti-inflammatory agent, Remicade, to treat rheumatoid arthritis, and the University of Pennsylvania's recent pioneering work in CAR-T immunotherapy. Scientific advances have occurred not just in the life sciences; university-based breakthroughs have been achieved in cryptography (such as the RSA encryption algorithm), computing, and many other fields.

Scientific discoveries based on university research have also generated significant income. For example, over the 1991-2010 time period, licensing revenues accrued to participants of the AUTM survey averaged 22.6% of endowment income based on an estimated endowment payout of 4% per year, or 11.3% based on an 8% payout (Figure I plots the distribution of technology licensing income to endowment payouts across all university-years), to make a comparison to this more widely-discussed source of university income.²

[INSERT FIGURE I AROUND HERE]

¹ Although patenting by U.S. universities occurred as early as the 1920s, the Bayh–Dole Act in 1980, which allows universities to own patent rights resulting from federally-funded research, is associated with a rise in university patenting in licensing since 1980 (Mowery et al. 2004). A series of studies have examined the patenting and commercialization performance of universities since the Act (Henderson et al., 1998; Mowery, et. al. 2004).

 2 Brown et al. (2014) have examined the recent investment performance of university endowments and show that endowment payouts have become an increasingly important component of universities' revenues in recent decades.

While these statistics suggest that income from commercializing university research is significant, it is difficult to assess whether the realized licensing revenues are "small" or "large." Most assessments are based on simple statistics such as counts or dollar amounts (e.g., Huggett (2017) or AUTM annual reports). In general, estimating the private economic value of patents or patent portfolios is difficult, as observable market transactions of patent sales or licenses are rare (the transfers do not occur regularly, and even when they do, the transactions are privately struck between parties). There have been some efforts to value corporate patents based on forward citations, corporate acquisition events, observed patent renewal fees at various stages of the patent lifecycle, or through patent disputes (e.g., litigation).³ Compounding the valuation problem is that these metrics are typically unrepresentative of the full distribution of patent values.

In this paper, we aim to evaluate the "potential" economic value of university patents (benchmarked against a similar patent portfolio granted to private firms) and to understand how much of that value has been captured through licensing revenue. We do so through the following steps: (1) we use patent-level licensing income data (including unlicensed patents) associated with 1,586 patents from a leading U.S. research university to identify patent characteristics which correlate with licensing income; (2) we use those patent characteristics to match university patents to publiclyheld corporate patents and to estimate the potential value of university patents; (3) we use the patent-level licensing income data to evaluate our estimated university patent value; (4) we use the university-level licensing income of 167 AUTM-member universities to analyze the proportion of estimated university patent value that has been commercialized; and (5) using the AUTM data, we analyze university characteristics and inputs that correlate with university patent value.

 3 Trajtenberg (1990), Harhoff et al. (1999), and Hall et al. (2005) have documented a positive relation between forward citations and market value. Lanjouw (1998) and Schankerman (1998) examine the relation between patent value and patent renewal. Bhagat et al. (1994), Lerner (1995), and Bessen and Meurer (2012) have examined the market reactions to firms' involvement in patent litigation.

With that overview, we wish to provide more detail about the estimation process. We first analyze detailed patent-level licensing income data (actual licensing revenues, including patents that remain unlicensed) from a leading U.S. research university from 1974 to 2018, and find that patent quality (i.e., forward citations) and generality are important characteristics in explaining patent-level licensing income. We then identify and match U.S. university-assigned patents granted between 1976 to 2010 to a standardized list of U.S. universities. We match each patent to a publicly-held corporate patent displaying similar characteristics (described in detail below). A university patent is then assigned the median of the values of matched corporate patents estimated by Kogan et al. (2017), which is our estimated potential value.⁴ To examine potential errors in our sampling and matching procedure, we conduct a simulation analysis (described below) based on randomly selected university patents and corporate patents, and find modest sampling error. It is also worth noting that while we use Kogan et al.'s (2017) estimated patent value due to its public access, our matching procedure can be based on any estimation method for corporate patent value.

We then employ the detailed patent-level licensing income data to calibrate an empirical patent valuation model. We extrapolate from this patent-level data to a sample of 167 AUTM-member universities reporting aggregate commercialization results in AUTM annual reports from 1991 to

⁴ The corporate patent value estimated by Kogan et al. (2017) is based on stock market reaction to the announcement of corporate patents, which is defined as the increase in market value in the three-day period around patent approval announcements, after adjusting for benchmark returns, idiosyncratic stock return volatility, and various fixed effects (more details are provided in the Supplemental Appendix). Such a market reaction-based valuation approach follows Austin (1993) and has been widely used in the economics literature; see Bhagat et al. (1994), Lerner (1995), and Bessen and Meurer (2012) for patent litigation, and Chen et al. (2005) for new product announcements. An alternative way to evaluate the value of corporate patents is to manually collect or purchase the disclosed licensing contracts by public firms (see Kankanhalli et al., 2019); however, even those disclosed contracts are subject to selection issues and redactions. We also acknowledge the following biases in benchmarking university patents against corporate patents. On the one hand, the market reaction to patents assigned to publicly-listed companies reflects not only technological merits but also marketing and production synergies that are not available to universities (Sampat and Ziedonis, 2004). Thus, matching university patents to corporate patents may *overestimate* the value of university patents. On the other hand, it is well known that the total economic value of an invention consists of private rent and public benefits. Since market reaction to corporate patents only reflects private rents, the proposed approximation may *underestimate* the total economic value of university patents. Moreover, we evaluate each patent separately and thus unavoidably neglect the potential complementarity of patents in a patent portfolio.

2010. We find that a university's potential patent value is positively associated with its license income and startups founded. We estimate through this approach that an average university in the sample realized 5-9% of their potential patent value in licensing income.

Finally, we discuss the university characteristics and inputs that correlate with patent value creation. Among university-level variables we collect from the AUTM survey data or other sources, R&D investment, the number of faculty members, and the number of full-time employees in the technology transfer office (TTO) have explanatory power for a university's patent value, while Carnegie research ranking and the presence of an affiliated medical school have no or only weak explanatory power.

Overall, we propose a new benchmarking approach for universities and their stakeholders to evaluate the economic value of patent portfolios held by universities and assess university technology commercialization efforts. This approach may be informative to universities and policymakers for resource and asset allocation decisions relying on evidence-based indicators.⁵ We also add new evidence to the literature on universities' patent value and licensing income, as well as the factors influencing their performance in commercialization.⁶ A novelty of our approach is that we introduce matched corporate patent values into the evaluation. Our empirical evidence thus offers new insights to the assessment and realization of the value of university patents. Nevertheless, we also acknowledge the limitations of our estimation approach and research design

⁵ For example, UMETRICS is a recent initiative in organizing all input and output indicators related to science activities in universities (see Weinberg et al., 2014; Lane et al., 2015).

⁶ We find a university patent's forward citations and generality to be significantly and positively related to its licensing income. Using the licensing income data of University of California and Columbia University in the 1980s and 1990s, Sampat and Ziedonis (2004) find that the number for forward citations predicts if a patent is licensed but not the amount of revenue. Lach and Schankerman (2008) find that U.S. universities that give higher royalty shares to faculty members are associated with higher license income. Azoulay et al. (2007) and Audretsch et al. (2009) examine the determinants of the commercialization of research done by university scientists. In addition, Thursby and Kemp (2002), Thursby and Thursby (2002), Di Gregorio and Shane (2003), Siegel et al. (2003), Belenzon and Schankerman (2009), and Sampat (2006) have examined why some universities exploit their intellectual property more effectively than do others in terms of licensing patents and startup creation.

and discuss them in detail in the concluding section.

Our study also speaks to the analysis of rent-sharing of innovation output. Our estimate for university patent value is based on corporate patent value that comprises not only the direct technical value of the invention, but also complementary marketing and production (Teece, 1986; Arora et al., 2001). Because universities are unable to realize the economic value of their patented inventions by reaching the market themselves, we interpret the "conversion" rate of 5 to 9% in our AUTM dataset as a rent share to, or economic value created by academic researchers through upstream research activity with the remaining value accruing to the downstream licensee.⁷

II. DATA

We first describe the process of collecting university patents and associated information in Section A. In Section B, we describe the patent licensing dataset of a prominent U.S. research university, which allows us to associate patent characteristics with actual patent licensing revenue. We then explain our matching process for university and corporate patents and the estimation of the potential value of university patents in Section C.

A. University Patent Data

We first collect data on patents granted to U.S. universities from 1976 to 2010. Specifically, we manually construct a list of assignees and corresponding identifiers (PDPASS) that are U.S. universities, institutes, and foundations. We first examine the National Bureau of Economic Research (NBER) patent assignee file (1976-2006) and identify all assignees in the category of

⁷ Relatedly, a recent study by Kline et al. (2017) also uses Kogan et al.'s (2017) patent value and shows that 29% of patent-induced operating surplus is transferred to workers (including inventors and non-inventors).

"U.S. University."⁸ We then use other sources to identify research institutes and other entities affiliated with universities.⁹ We then manually search possible names (universities, research institutes, and foundations) in other non-university categories in the NBER patent assignee file and extract related unique identifiers (known as "PDPASS" in the dataset). For example, the hospital of the School of Medicine of Tufts University is listed under the "U.S. Hospital" category. Also, the Purdue Research Foundation affiliated with Purdue University is listed in the category of "U.S. Institute." This process results in a list of 362 U.S. universities which received at least one patent in the sample period. The complete list of the university-PDPASS pairs is reported in Table OA.III of the Supplemental Appendix.

Based on the university-PDPASS pairs, we construct a dataset of U.S. university patents. We then combine the patent and citation data from NBER (Hall et al., 2001), Patent Network Dataverse of Harvard University (Li et al., 2014), and Patentsview to construct a dataset that includes detailed information on each patent granted to U.S. universities from 1976 to 2010.¹⁰ The resulting sample consists of 77,880 university-linked patents.

We then assemble the following patent characteristics variables commonly used in the prior literature on university patenting: (i) *Quality* is defined as the number of forward citations received by a patent within five years after its grant year (Trajtenberg, 1990; Sampat and Ziedonis, 2004; Hall et al., 2005);¹¹ (ii) *Generality* is defined as one minus the Herfindahl-Hirschman Index (HHI)

⁸ For example, Harvard University has several different names in this category, including "Harvard College," "Harvard President & Fellows of Harvard College," "Harvard Univ. Office of Tech Transfer," etc.

⁹ Some university patents are assigned to categories other than universities, such as institutes (e.g., university hospitals) or research corporations affiliated to universities. We use the *U.S. News National University Rankings* and *Top 100 Worldwide Universities Granted U.S. Utility Patents* published by National Academy of Inventors to help identify universities and their affiliates in our sample.

¹⁰ The NBER database is downloadable via: http://www.nber.org/patents/; Patent Network Dataverse of Harvard University is downloadable via: https://dataverse.harvard.edu/dataverse/patent; and the Patentsview database is downloadable via: http://www.patentsview.org/web/.

¹¹ Lanjouw and Schankerman (2004) find that forward citations explain 48% of the variation of their patent quality index. Harhoff et al. (1999) and Hall et al. (2005) show that forward citations are associated with higher patent valuation from survey and stock price data, respectively.

of patent subcategory citations received from forward citing patents (Trajtenberg et al., 1997; Hall et al., 2001); (iii) *Originality* is defined as one minus the HHI of patent subcategory citations of the focal patent (Trajtenberg et al., 1997; Hall et al., 2001); (iv) *Basicness* is defined as the ratio of the number of references to prior non-patent documents divided by the total references in the focal patent, which reflects the direct dependence on scientific and academic knowledge (Trajtenberg et al., 1997; Fleming and Sorenson, 2004; Ali and Gittelman, 2016); and (v) *Claims* denotes the number of claims of each granted patent, which defines the coverage and scope of a patent (Lerner, 1994).

In Table I, we report the averages for all five measures for the university patents in our sample and patents assigned to U.S. public firms.¹² We find that university patents receive significantly more forward patent citations (5.55 vs. 4.97) on average, are more general (0.44 vs. 0.38), are more original (0.42 vs. 0.36), are more "basic" (0.47 vs. 0.11), and contain more claims (20.39 vs. 16.31) compared to corporate patents.¹³ These differences are largely consistent with the literature (e.g., Trajtenberg et al., 1997; Henderson et al., 1998). We take these five characteristics into account in our matching procedure.

[INSERT TABLE I AROUND HERE]

We also observe that university patents are concentrated in certain technology fields such as Drugs, Chemicals, and Surgery and Medical Instruments, as shown in Panel C of Table OA.I in the Supplemental Appendix. We define technology fields by the two-digit subcategory codes of Hall et al. (2001). Our matching procedure takes technology field differences into account by matching

¹² Our corporate patents include 1,361,771 patents granted to assignees in U.S. public firms (i.e., assignees with GVKEY identifiers) in the NBER assignee file from 1976-2010.

¹³ Consistent patterns are observed in different sample periods (Panel A in Table OA.I in the Supplemental Appendix), in distribution (Panel B in Table OA.I in the Supplemental Appendix), and in different technology subcategories (Panel D to Panel G in Table OA.I in the Supplemental Appendix).

each university patent to corporate patents in the same technology field (to be discussed in more detail later).

B. Patent Licensing Dataset

We collect a complete patent licensing dataset from a leading U.S. research university to facilitate our estimation of university patent value as compared to corporate patent value. The dataset includes 7,797 unique technologies and 779 licensing contracts. Among unique technologies, 2,246 licensed and 5,551 unlicensed, from 1974 to 2018. Some technologies are patented and some are not. A licensing contract (i.e., agreement) includes one or more technologies with related patent numbers (if associated patent applications were filed and granted), licensing status, execution date, license fee, maximum royalty rate, exclusivity in licensing or not, lifetime revenue, technology fields, etc. There are on average 2.88 technologies included in a licensing contract. Among the 779 licensing contracts, 227 are exclusive, 12 are co-exclusive, and 540 are non-exclusive. The licensing revenue reflects the total amount of cash received based on licensing, royalties, or equity. According to our contact at the university, the majority of the startups do *not* include an equity component in the license; thus, lifetime revenue mainly comes from license fees and royalties.

We focus on 765 licensed patents and 821 unlicensed patents from 1976 to 2010 and calculate a patent's licensing revenue as the lifetime revenue of the contract divided by the number of patents involved. Because patents in our data have different "lifetimes" to be licensed, truncation bias is a potential concern. For example, the lifetime revenue from a patent that was recently granted could be zero if it has not been licensed or may be underestimated as we only know its income by 2017. We take a conservative approach and do not attempt to extrapolate or estimate the future income from those patents that are subject to such truncation bias. Note that all revenue income recorded at this level is inclusive of the amount which will be shared with the inventor and the inventor's

department (which together represent an average of about 30% of the gross licensing revenues at this institution).

Our dataset also includes unlicensed patents, and their licensing revenue is set as zero if the patent is never observed to result in positive revenue. This conservative setting unavoidably underestimates the value of university patents for several reasons: first, some unlicensed patents may be valuable to industries but were unlicensed for any number of reasons. Second, these patents may have been exploited by firms without receiving royalties because universities may be less aggressive in enforcing IP rights. Finally, to be conservative in the estimation, we do not impute positive patent licensing income to unlicensed patents, though some of these patents may result in licensing income outside of the observation window (particularly for patents granted in more recent years).

Table II presents the patent-level summary statistics of variables of 765 licensed patents and 821 unlicensed patents from 1976 to 2010. Among the licensed patents, 333 are exclusive, 135 are coexclusive, and 297 are non-exclusive. Among these 765 licensed patents, their mean, median, standard deviation, and maximum of lifetime revenue (in thousands) are 185, 17, 725, and 11,256, respectively. The mean and median of maximum royalty rate are 3.44% and 3%, respectively. The mean, median, standard deviation, and maximum of licensing fee (in thousands) are 11, 1, 52, and 625, respectively.

We also compare patent characteristics of licensed patents to those of unlicensed patents. We find that, on average, licensed patents are created by more inventors (2.73 vs. 2.44), are more highly cited $(7.28 \text{ vs. } 3.13)$, are more general $(0.46 \text{ vs. } 0.40)$, are more original $(0.37 \text{ vs. } 0.34)$, are more basic (0.55 vs. 0.49), and contain more claims (15.86 vs. 14.58) than unlicensed ones. Each of these differences is statistically significant.

[INSERT TABLE II AROUND HERE]

C. Estimating University Patent Value

To identify which patent characteristics are correlated with realized licensing revenues, we perform OLS regressions of patent lifetime revenue on quality, generality, originality, basicness, the number of claims, and fixed effects for years and subcategories using the 1,586 patents (765 licensed patents and 821 unlicensed patents) from a leading U.S. research university. The lifetime revenue of each patent is inflation-adjusted based on the consumer price index (CPI) of the grant year of the patent: we scale the lifetime revenue of a patent by the CPI from the Federal Reserve Bank of St. Louis (the index is normalized as 1 for years 1982-1984). Results are reported in Table III. Regardless of whether lifetime revenue is specified in a linear form (Panel A) or in a log-one-plus form (Panel B), both quality and generality are estimated with positive and statistically significant coefficients in most cases. Other patent characteristics do not appear to positively correlate with realized licensing revenues. We therefore conclude that both quality and generality are key determinants of university patent value, and use them in our subsequent matching and benchmarking efforts for all university patents. Interestingly, Henderson et al. (1998) also focus on these two characteristics in comparing university and corporate patents. We also argue that, as generality reflects externality and spillover effects, it captures public benefits of university patents to a certain extent.

[INSERT TABLE III AROUND HERE]

The next step is to estimate the potential value of all university patents by benchmarking them against similar corporate patents. First, we collect the patent value of corporate patents (i.e., patents

assigned to public firms) from Kogan et al. (2017) ,¹⁴ which calculates the value of a patent granted to a U.S. public firm using stock market reaction to the announcement of the patent – we proxy a public firm's patent value with the 3-day appreciation of market capitalization of this firm around the announcement, adjusted for measurement noise and various fixed effects. The details of the estimation is provided in the Supplemental Appendix. The estimated patent value is also inflationadjusted based on the CPI (the index is normalized as 1 for years 1982-1984). Kogan et al.'s (2017) market reaction-based valuation approach follows Austin (1993) and is consistent with the valuation of patent litigation of Bhagat et al. (1994), Lerner (1995), and Bessen and Meurer (2012) and the valuation of new products of Chen et al. (2005).

Second, we benchmark each university patent to similar corporate patents to estimate its potential value based on a matching method based on quality, generality, technology field, and grant year. A university patent's potential value is the stock market reaction to a similar patent that is granted to the private sector. Specifically, the potential value of a university patent is set to that of a corporate patent that is in the same patent sub-category, granted in the same year, and has the shortest sum of distances to the focal university patent in terms of quality and generality.¹⁵ If there are multiple corporate patents satisfying the above criteria, we set the university patent value to be the median value of those corporate patents. We find that the quality and generality of university patents are similar to those of matched corporate patents: 7.67 vs. 7.58 in average quality and 0.47 vs. 0.47 in average generality.

Table IV presents the distribution statistics of estimated patent value (*PatVal(Matched)*) of 77,880

$$
\frac{|4-6|}{4+6}+\frac{|0.3-0.1|}{0.3+0.1}=0.7.
$$

¹⁴ The patent value data is downloadable via: https://iu.app.box.com/v/patents.

¹⁵ As shown in Table III, these two characteristics (quality and generality) have the greatest explanatory power of university licensing revenues. The distance along a patent characteristic is measured by the absolute value of the difference between two values divided by their sum. For example, if a university patent has its forward citation as 4 and its generality as 0.3 and a corporate patent has it forward citation as 6 and its generality as 0.1, then their distance is equal to 0.7, calculated as

university patents and that of 1,361,177 corporate patents (*PatVal*). We find that the average (median) value is \$14.77 (\$5.33) million among university patents, compared to \$12.09 (\$3.62) million among corporate patents. These figures suggest that university patents may be at least similarly valuable as corporate patents when they exhibit similar patent characteristics. Figure II illustrates the distributions of the values of university patents and corporate patents. They are highly comparable except for the left tail.

[INSERT TABLE IV AROUND HERE] [INSERT FIGURE II AROUND HERE]

We also implement simulations to examine potential errors if our collection of university patents is incomplete or if our matching method is biased. In each simulation, we randomly draw half of the university patents and half of the corporate patents within each subcategory-grant year group. We then benchmark each university patent to similar corporate patents, all from the random draw, and assign it the value of a similar corporate patent that is in the same patent sub-category, granted in the same year, and has the shortest sum of distances in terms of quality and generality. Similar to the previous procedure, we set the university patent value to be the median value of all corporate patents that satisfy the above criteria. We calculate the median and mean of the estimated values of all randomly drawn university patents and then calculate the absolute percentage deviation between the median (mean) and the full-sample median of 5.33 (mean of 14.77). By repeating the aforementioned procedure 500 times, we collect 500 absolute percentage deviations of median and mean and plot their frequency in histogram in Figure III. We find that all absolute percentage deviations of median are within the 5% range, and that most of absolute percentage deviations of mean are within 7.5%. These findings confirm that our matching and valuation method for university patents is reasonably robust to sampling of university and corporate patents – even when the sample size randomly drops by half, the majority of sampling errors in median and mean are

below 5% and 7.5%, respectively.

III. VALUATION TESTS

In Subsections A and B, we cross-check our valuation method with licensing data at both the patent level and the university level, respectively. In Subsection C, we examine university characteristics which correlate with the cross-university variation in patent value.

A. University Patent Value and License Revenue at the Patent Level

We now revisit the 1,586 licensed and unlicensed patents from the leading U.S. research university analyzed above. We cross-check our valuation method by comparing the estimated patent values to patent licensing revenue. In particular, we regress patent lifetime revenue on estimated patent value, controlling for other patent characteristics, year fixed effects, and technology field fixed effects (by subcategory), in both linear and log-one-plus forms (Panels A and B, respectively). As shown in Table V, our estimated patent value is significantly and positively associated with the actual realized licensing revenue in all specifications. Such a pattern is robust to linear and log forms, the inclusion of other patent characteristics (i.e., originality, basicness, and claims), and the inclusion of year and patent subcategory fixed effects. The fact that our estimated patent value explains realized licensing revenue suggests that our method indeed captures the variation in university patent value to a certain extent. The coefficients on *PatVal(Match)* in Panel A provide an estimate of the extent to which a university patent value is successfully commercialized. Taking Column (1) as an example, the coefficient on *PatVal(Match)* is 0.1484 when we do not include the intercept term, which implies that a patent worth \$1 million is associated with \$0.15 million of licensing income on average. When we include the intercept term in Column (2), the coefficient on *PatVal(Match)* becomes 0.0819 and implies that an increase of \$1 million worth of patent value is

associated with an increase of \$0.08 million in licensing income on average. The coefficient drops to 0.0608 but remains statistically significant when we include more control variables and fixed effects in the regression (Column (5)). As a result, Table V suggests that the university realizes approximately 6.1-14.8% of the total private value of corporate patents with similar characteristics.

[INSERT TABLE V AROUND HERE]

Given the importance of medical innovation in universities' patent portfolios (as shown in Panel C of Table OA.I in the Supplemental Appendix), we also report the estimated results based on 663 drug-related patents of the U.S. research university in Table OA.II in the Supplemental Appendix. We find that the coefficients on *PatVal(Match)* range between 7.1% and 12.6% in Panel A, which are largely consistent with the estimates in Table V.

B. University Patent Value, License Revenue, and Startups at the University Level

We proceed to examine the relation between patent value and license income at the university level by collecting the annual statistics of university-level license income and number of startups formed from the annual reports of AUTM.¹⁶ These two statistics allow us to measure the commercialization performance of a university's IP in two different dimensions. We include a total of 167 U.S. universities that have reported outcomes to the AUTM survey at least once between 1991 and 2010. We use the CPI to adjust all annual licensing incomes to the level of 1982-1984. In this sample, the average, median, and standard deviation of annual license income (in millions) are 4.50, 0.69, and 16.75, respectively. Moreover, the average, median, and standard deviation of the number of startups formed are 2.84, 1.00, and 4.51, respectively. We sum up the estimated values of all patents

¹⁶ For example, Northwestern University earned \$824 million in license income in 2008, which tops all university-year observations. In terms of total license income in 1991-2010, University of California, New York University, and Columbia University are the top three universities (with totals of \$1,805 million, \$1,790 million, and \$1,392 million, respectively).

granted to each of the 167 U.S. universities in each year. The average, median, and standard deviation of estimated university patent value (in millions) are 386.88, 109.00, and 957.07, respectively.

To understand to what extent a university's patent value explains its total license income spanning multiple years, we estimate a *cross-sectional* regression as follows: first, we define a university's patent value in year *t* as the sum of estimated values of all patents granted to the university in year *t*. We then calculate the time-series average of each university's patent value to be the main explanatory variable, *Average PatVal(Match)*, in Table VI. Second, since patents are valid up to 20 years and thus generate licensing income for multiple years, we calculate a university's total license income in a year using a straight line depreciation plan to discount its annual inflation-adjusted license income in the following 20, 15, and 10 years.¹⁷ We then calculate the time-series average of each university's total license income to be the dependent variable in Table VI. Last, we conduct cross-sectional regressions to regress universities' average total license income on *Average PatVal(Match)*.

We report the OLS regression estimations in both a linear form and a log form in Table VI Panel A. We find that the coefficients on *Average PatVal(Match)* are significant in all specifications, suggesting that the estimated potential patent value explains license income. Taking Column (1) based on 20-year license income as an example, the coefficient on *Average PatVal(Match)* is 0.0904. This indicates that an increase of \$100 million worth in a university's new patents is correlated with an increase of \$9.0 million worth in a 20-year license income stream on average. When we use 15- and 10-year license income (Columns (3) and (5)), the coefficients on *Average PatVal(Match)* are 0.0689 and 0.0474, indicating that an increase of \$100 million worth in a

 17 Each university-year observation is included in our regression sample for Table VI when the university shows up on the AUTM report in that year. In the 2,109 observations, we impose the missing license income of 36 observations (or 1.71% of the sample) to be zero. When we calculate future license income based on future 20, 15, and 10 years, the straight line depreciation rates are 5%, 6.77%, and 10%, respectively.

university's new patents is correlated with an increase of \$7 and \$5 million worth in a 15- and 10 year license income stream, respectively. These statistics suggest that universities realize 5-9% of the estimated value based on publicly-held corporate patents with similar characteristics as those from our sample of universities.

In Table VI Panel B, we examine the relation between our estimated university patent value and the number of startups formed at the university level. Similar to the approach used in Panel A, we use the time-series average of the discounted number of startups created by the university in the following 20, 15, and 10 years as the dependent variable.¹⁸ Results reported in Panel B are also based on cross-sectional regressions, for which we regress the time-series average of discounted number of startups created by a university in the following 20, 15, and 10 years on the university's *Average PatVal(Match)*. Results suggest a positive and statistically significant relation and confirm the intuition that more technologically capable universities create more new businesses. In terms of economic magnitude, Column (1) suggests that 42 (=0.0440*957.07) more startups will be formed in the following 20 years if the value of a university's patent portfolio increases by one standard deviation. Such an estimate is substantial given that sample average and median are 2.84 and 1.00, respectively, per year.

[INSERT TABLE VI AROUND HERE]

C. University Patent Value and University Characteristics

After showing that university patent value is able to explain both patent licensing and startup formation, we analyze the production function of university patent value to understand what inputs

¹⁸ Each university-year observation is included in our regression sample for Table VI when the university shows up on the AUTM report in that year. In the 2,109 observations, we impose the missing number of startups of 709 (or 33.62% of the observations) to be zero.

are crucial to valuable university patents. We collect the following university variables as "inputs": the five-year cumulative inflation-adjusted R&D expenditure (*R&D*, with a 20% obsolescence rate per year), a dummy variable indicating whether the sample university is a Carnegie-ranked research university or not (*Carnegie*), the number of full-time faculty members (*Faculty*), the full-time equivalents (*FTE*) in technology transfer office (*TTO*) in that year, and a dummy variable indicating whether the sample university has a medical school or not (*Medical School*). The information on R&D expenditure and the existence of a technology transfer office is collected from AUTM annual surveys. The data on full-time faculty numbers are manually extracted from the Carnegie report (1994, 2000, 2005, and 2010).¹⁹ The information on the presence of a medical school is collected from internet searches. We require a university to have non-missing values in *R&D*, *Faculty*, and *FTE* to enter into the sample. Table VII Panel A reports the cross-sectional correlation matrix of these university characteristics. Perhaps not surprisingly, some variables are highly correlated. For example, the correlation coefficient between R&D expenditure and the number of full-time faculty (number of full-time equivalents) is 0.891 (0.928).

We then regress the estimated potential value of all patents applied by (and later granted to) a university in a year on several university characteristics in a log-log form assuming a Cobb-Douglas production function of patent value. Results reported in Panels B and C are based on pooled regressions for university-year observations and cross-sectional regressions based on university observations, respectively.²⁰ To avoid multi-collinearity, we first include these characteristics in regressions one by one in Columns (1) to (5) in Panel B. We find that all are positively and significantly correlated with the output of patent value. When we include all five variables together in one regression, however, we find that only R&D expenditure, the number of full-time faculty

¹⁹ We assign the number of faculty members in 1994 to all years before 1994, and apply this rule to fill in the number for each university in all other years.

²⁰ For pooled regressions in Panel B, we use standard errors clustered by university to correct for errors in autocorrelation. For cross-sectional regressions in Panel C, we average all dependent variables and independent variables, and then run OLS regression using a total of 158 observations (due to missing information for some universities).

members, and the number of full-time equivalents in the TTO are statistically significant in Column (6). The coefficients of *Carnegie* and *Medical School* become insignificantly negative in Column (6), likely due to multi-collinearity. In terms of economic magnitude, a doubling of R&D expenditure, the number of full-time faculty, and TTO employees, is associated with patent value increases of 55%, 33%, and 42%, respectively, holding other variables fixed.

Panel C reports cross-sectional regressions, as longitudinal variation of the right-hand side variables may be modest. In particular, we regress the time-series average of total estimated potential value of patents applied by (and later granted to) a university in all years on the time-series averages of the five input variables. We find that when only one input variable is included at a time, each is positively and significantly correlated with university patent value, except in Column (5) for the existence of medical schools. When we include these five variables in one regression, we find that only R&D expenditure and the number of full-time faculty have statistically significant explanatory power for university patent value. Their coefficients suggest that doubling a university's research expenditure or full-time faculty is associated with increased patent value output by 64% or 30%, respectively, holding the other variables fixed.

[INSERT TABLE VII AROUND HERE]

III. CONCLUSION AND DISCUSSION

The degree to which universities should be in the business of commercially translating their scientific discoveries through patenting, licensing and startup efforts has long been debated (e.g., Bok, 1982; Etzkowitz, 1994; Mowery et. al. 2004; Åstebro et al., 2012). To better inform that debate and to more generally assess university technology commercialization efforts, we develop a novel approach to estimating (a) the potential economic value of university patents, and (b) the proportion

of this value that is accrued by the university through licensing. Further analyses suggest that research expenditure, the number of researchers, and full-time equivalents in TTO are key factors correlated with university patent value.

Our analysis contains some limitations that are worth further discussion. As we have acknowledged earlier, one limitation of our "market-based" measure of patent value is that stock market reaction to comparable patents held by corporations may reflect value capture expectations of more fullyintegrated operations (Teece, 1986; Arora et al., 2001) that may not be appropriate for academic institutions which lack the complementary assets often required for commercialization. Nevertheless, we are proposing a new approach that can be applied to *any* pricing data or methods for corporate patents.

Second, licensing income is not the only way for universities and academic researchers to be rewarded for their innovation output. For example, the founders of many companies in Silicon Valley started their innovation at Stanford University and later made significant donations back to the university. In addition, universities and their faculty members may lack incentive to patent their inventions or license these patented inventions – their efforts may be rewarded in other forms, such as publication, grants, and peer recognition (Lacetera, 2009). Furthermore, innovation taking place on campus also enhances the human capital of faculty members, lab researchers, and students, which is positively associated with future economic payoff at both the individual and aggregate levels. Thus, estimating the economic value of university patents or commercialization performance based on licensing income underestimates universities' intellectual contribution to society (such as spillovers).

Third, we calibrate our estimate of the fraction of patent value captured by the university based on outcomes at a single university. While our baseline university is a leading research institution and prolific patenter and licensor, its practices and outcomes may not be representative of all academic institutions. More broadly, unlike that of private firms, universities' missions extend beyond profit maximization and include broad long-term goals such as the creation and dissemination of knowledge and human capital without regard to private profit. Moreover, universities may be more likely to develop technology applicable to a wider spectrum of applications as compared to corporations, such as orphan drugs. Thus based on our analysis we cannot make social welfare arguments regarding the appropriate level of patent economic value that "should" be captured by universities.

Nevertheless, the efforts of licensing university patents importantly contribute to the mission of universities, as most powerfully illustrated by the contribution of licensing revenues to university operating budgets and compared to the more frequently discussed university endowments. The main purpose of our new benchmarking approach is to provoke conversations among university policy makers as to where they believe their institution should "sit" with regard to commercializing their intellectual property assets. Conducting such an exercise, we believe, requires rigorous statistical analysis in addition to the summary metrics often currently employed in assessing university commercialization performance. Our hope is that our estimates provide some sense of the contribution of upstream scientific research versus the downstream commercialization efforts by licensees in generating economic value from patented academic discoveries. Clearly, a complete understanding and a fair assessment of the economic value generated and captured by universities of their scientific discoveries through patents requires further research and data beyond our initial foray. We look forward to actively engaging in this research in the future.

References

- Ali, A., & Gittelman, M. (2016). Research paradigms and useful inventions in medicine: Patents and licensing by teams of clinical and basic scientists in Academic Medical Centers. *Research Policy*, 45(8), 1499-1511.
- Arora, A., Fosfuri, A., & Gambardella, A. (2001). *Markets for Technology: The Economics of Innovation and Corporate Strategy*. MIT press.
- Åstebro, T, Bazzazian, N., & Braguinsky, S. (2012). Startups by recent university graduates and their faculty: Implications for university entrepreneurship policy. *Research Policy*, 41(4), 663- 677.
- Audretsch, D. B., Aldridge, T., & Oettl, A. (2009). Scientist commercialization and knowledge transfer. pp. 176-201 in *Entrepreneurship, Growth, and Public Policy*, edited by Acs, Zoltán J., David B. Audretsch, and Robert J. Strom. Cambridge University Press, 2009.
- Austin, D. H. (1993). An event-study approach to measuring innovative output: The case of biotechnology. *American Economic Review*, 83(2), 253-258.
- Azoulay, P., Ding, W., & Stuart, T. (2007). The determinants of faculty patenting behavior: Demographics or opportunities? *Journal of Economic Behavior & Organization*, 63(4), 599- 623.
- Belenzon, S., & Schankerman, M. (2009). University knowledge transfer: private ownership, incentives, and local development objectives. *Journal of Law and Economics*, 52(1), 111-144.
- Bessen, J. & Meurer, M.J. (2012). The private costs of patent litigation. *Journal of Law, Economics & Policy*, 9, 59-95.
- Bhagat, S., Brickley, J.A. & Coles, J.L. (1994). The costs of inefficient bargaining and financial distress. *Journal of Financial Economics*, 35, 221-247.
- Bok, D. C. (1982). *Beyond the Ivory Tower: Social Responsibilities of the Modern University*. Harvard University Press.
- Brown, J. R., Dimmock, S. G., Kang, J. K., & Weisbenner, S. J. (2014). How university endowments respond to financial market shocks: Evidence and implications. *American Economic Review*, 104(3), 931-62.
- Chan L. K. C., Lakonishok J., & Sougiannis T. (2001) "The stock market valuation of research and development expenditures," *Journal of Finance*, 56(6): 2431-2456.
- Chen, S. S., Ho, K. W., & Ik, K. H. (2005). The wealth effect of new product introductions on industry rivals. *Journal of Business*, 78(3), 969-996.
- Di Gregorio, D., & Shane, S. (2003). Why do some universities generate more start-ups than others? *Research Policy*, 32, 209-227.
- Etzkowitz, H. (1994). Knowledge as property: The Massachusetts Institute of Technology and the debate over academic patent policy, *Minerva*, 32, 383-421.
- Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools (No. w8498). National Bureau of Economic Research.
- Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2005). Market value and patent citations. *RAND*

Journal of Economics, 36, 16-38.

- Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation frequency and the value of patented inventions. *Review of Economics and Statistics*, 81(3), 511-515.
- Henderson, R., Jaffe, A., & Trajtenberg, M. (1998). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. *Review of Economics and Statistics*, 80(1), 119-127.
- Huggett, B. (2017). Top US universities, institutes for life sciences in 2015. *Nature Biotechnology*, *35*(3), 203.
- Kankanhalli, G. and Kwan, A., & Merkley, K. J. (2019). Speech is silver, but silence is golden: Information suppression and the promotion of innovation. Working paper.
- Kline, Patrick, Petkova, N., Williams, H., & Zidar, O. (2017). Who profits from patents? Rentsharing at innovative firms. Working paper.
- Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*, *132*(2), 665-712.
- Lacetera, N. (2009). Academic entrepreneurship. *Managerial and Decision Economics*, 30(7), 443- 464.
- Lach, S., & Schankerman, M. (2008). Incentives and invention in universities. *RAND Journal of Economics*, 39(2), 403-433.
- Lane, J.I., Owen-Smith, J., Rosen, R.F., & Weinberg, B.A. (2015). New linked data on research investments: Scientific workforce, productivity, and public value. *Research Policy*, 44(9), 1659-1671.
- Lanjouw, J. O. (1998). Patent protection in the shadow of infringement: Simulation estimations of patent value. *Review of Economic Studies*, 65(4), 671-710.
- Lanjouw, J. O. and Schankerman, M. (2004). Protecting intellectual property rights: are small firms handicapped? *Journal of Law and Economics* 47, 45-74.
- Fleming, L., Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal*, 25(8-9), 909-928.
- Lerner, J. (1994). The importance of patent scope: an empirical analysis. *RAND Journal of Economics,* 25(2), 319-333.
- Lerner, J. (1995). Patenting in the shadow of competitors. *Journal of Law and Economics*, 38, 463- 495.
- Li, G. C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., Yu, A. Z., & Fleming, L. (2014). Disambiguation and co-authorship networks of the US patent inventor database (1975–2010). *Research Policy*, *43*, 941-955.
- Mowery, D.C., Nelson, R.R., Sampat, B.N., & Ziedonis, A. (2004). Ivory Tower and Industrial Innovation, Stanford University Press, Stanford, CA.
- Sampat, B.N. (2006). Patenting and US academic research in the 20th century: The world before and after Bayh-Dole. *Research Policy*, 35(6), pp.772-789.
- Sampat, B.N., & Ziedonis, A. (2004). Patent citations and the economic value of patents: A

Preliminary assessment. pp. 277–98 in *Handbook of Quantitative Science and Technology Research*, edited by Henk F. Moed, Wolfgang Glanzel, and Ulrich Schmoch. Dordrecht: Kluwer.

- Schankerman, M. (1998). How valuable is patent protection? Estimates by technology field. *RAND Journal of Economics*, 29, 77-107.
- Siegel, D.S., Waldman, D., & Link, A. (2003). Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: an exploratory study. *Research Policy*, 32(1), 27-48.
- Teece, D.J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, 15(6), 285-305.
- Thursby, J.G., & Kemp, S. (2002). Growth and productive efficiency of university intellectual property licensing. *Research Policy*, 31(1), 109-124.
- Thursby, J.G. & Thursby, M.C. (2002). Who is selling the ivory tower? Sources of growth in university licensing. *Management Science*, 48(1), 90-104.
- Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *RAND Journal of Economics*, 21, 172-187.
- Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, 5(1), 19-50.
- Weinberg, B.A., Owen-Smith, J., Rosen, R.F., Schwarz, L., Allen, B.M., Weiss, R.E., & Lane, J. (2014). Science funding and short-term economic activity. *Science*, 344(6179), 41-43.

Figures

Figure I: Frequency of the Ratio of License Income to Endowment Payout.

We first collect the annual statistics of university-level annual license income from the annual reports of the Association of University Technology Managers (AUTM, available from 1991 to 2010). We then collect the annual statistics of university endowments of AUTM universities in 1991-2010 from National Association of College and University Business Officers (NACUBO). We then estimate each university's endowment payout in a year by multiplying each university's endowment amount by annual returns of 4% (conservative estimate) and 8% (historical mean). We calculate the ratio of licensing income to endowment payout for each university-year observation and plot the frequencies (vertical axis) of this ratio in this figure. The dotted blue line denotes the ratio based on investment return of 8%, and the red solid line denotes the ratio based on investment return of 4%. The pooled average ratio of license income to endowment payout across all university-year observations is 11.3% based on investment return of 8% (and 22.6% based on investment return of 4%).

Figure II: Distributions of Value of Patents Granted to Public Firms and Universities. We illustrate the distribution of the estimated university patent values using the reported matching method in the text, as compared to the distribution of corporate patent values. Specifically, we report the frequency of patent value in each interval for patents granted to universities and corporates, respectively.

Figure III: Possible Sampling Error in the Estimation Method of University Patent Value.

We examine to what extent our university patent value estimation is subject to the sampling of university patents and corporate patents. To do so, we simulate 500 sampling draws, in which for each draw we randomly select half of the university patents and half of the corporate patents by their grant years and subcategories. We estimate the value of each university patent using the matching method described in Table IV. Then we compute the median and mean of university patent value in each simulation draw, calculate its absolute percentage derivation from the full-sample median (i.e., 5.33 as shown in Table IV) and full-sample mean (i.e., 14.77 as shown in Table IV), respectively. Their distributions are shown in the blue and red histograms, respectively.

Tables

Table I: Summary Statistics of Characteristics of Patents Granted to Public Firms and Universities in the U.S.

We compare the distribution of patent quality (i.e., the citations received in five years after the patent is granted), patent generality (i.e., one minus the HHI of citations received from other patents over patent subcategories), patent originality (i.e., one minus the Herfindahl-Hirschman Index (HHI) of citations given to other patents over patent subcategories), patent basicness (i.e., the ratio of the number of references to prior "non-patent documents" divided by the total references), and number of claims of patents granted to public firms and universities. The definitions of generality, originality, and basicness follow Trajtenberg et al. (1997). ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively, when comparing the mean characteristics of universities' patents with those of public firms' patents. Sample period: 1976-2010.

Table II: Summary Statistics of Lifetime Revenue and Characteristics of Licensed and Unlicensed Patents in A Research-Oriented University.

Using a patent licensing dataset provided by a prominent research-oriented U.S. university, we report the distribution of lifetime revenue, maximum royalty rate (in percentage), license fee, exclusivity, number of technologies (patented and unpatented) in the licensing agreement/package, number of patents in the agreement/package, number of inventors, patent quality, patent generality, patent originality, patent basicness, and number of claims of licensed patents in Panel A, and the distribution of number of inventors, patent quality, patent generality, patent originality, patent basicness, and number of claims of unlicensed patents in Panel B. Sample period: 1976- 2010.

Table III: Patent Characteristics and Actual University Patent Revenue.

We execute OLS regressions to examine the explanatory powers of patent characteristics for patent revenue. Our sample of patent revenue, including 765 licensed and 821 unlicensed patents, is obtained from a prominent research-oriented U.S. university. The dependent variable is the actual patent lifetime revenue in Panel A or the natural logarithm of one plus the patent lifetime revenue in Panel B, and the independent variable is one of the four patent characteristics (i.e., quality, generality, originality, basicness, and number of claims in Panel A and the natural logarithm of one plus quality, generality, originality, basicness, and the natural logarithm of one plus number of claims in Panel B). Lifetime revenue is split evenly to each patent in the same licensed agreement/package. If a patent is not licensed, we set its lifetime revenue as zero. Lifetime revenue is in \$ millions and adjusted for inflation.

Table IV: Distributions of Value of Patents Granted to Public Firms and Universities.

We benchmark the patent value of U.S. universities based on the value of patents granted to U.S. public firms. First, we follow Kogan et al. (2017) and calculate the value of a patent granted to a U.S. listed firm using its stock market's reaction – we proxy a public firm's patent value (*PatVal*) with the 3-day appreciation of market capitalization of this firm around the announcement of the patent, adjusted for measurement noise and various fixed effects. Second, based on the finding in Table III, we construct a matching model to estimate university patent value using quality and generality. Specifically, the value of a university patent is equal to that of a corporate patent which has the shortest sum of distances to the focal university patent in terms of quality and generality in the same grant year and in the same patent sub-category. The distance along a patent characteristic is measured by the absolute value of the difference between two values divided by their sum. For example, if a university patent has its quality as 4 and its generality as 0.3, and a corporate patent has its quality as 6 and its generality as 0.1, then their distance is equal to

$$
\frac{|4-6|}{4+6}+\frac{|0.3-0.1|}{0.3+0.1}=0.7.
$$

The university patent value takes the median value if the focal patent is matched to multiple corporate patents. Patent value is in \$ millions. This table reports the summary statistics of the estimated university patent values using the matching method as compared to the distribution of corporate patent values.

Table V: Market-Based Patent Value and Actual University Patent Revenue.

We execute OLS regressions to examine the explanatory power of estimated patent value for patent revenue. Our sample of patent revenue, including 765 licensed and 821 unlicensed patents, is obtained from a large patent office in a prominent U.S. university. The dependent variable is the actual patent lifetime revenue in Panel A or the natural logarithm of one plus the patent lifetime revenue in Panel B, and the independent variable of interest is the estimated patent value (*PatVal(Match)*, using the matching methodology discussed in Table IV) in Panel A and the natural logarithm of one plus the estimated patent value in Panel B. We also control for patent originality, patent basicness, number of claims, grant year fixed effects, and patent sub-category fixed effects. Lifetime revenue and estimated patent value are in \$ millions and adjusted for inflation. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively.

Table VI: Association between Patent Value and License Income/Startups Formed in U.S. Universities.

In this table, we examine the explanatory power of our estimated university patent value (*Average PatVal(Match)*) for future license income and number of startups formed at the university level. To do so, we run cross-sectional regressions of future license income (in Panel A) and future number of startups formed (in Panel B) on patent value. We first define a university's patent value in year *t* as the sum of estimated values of all patents granted to the university in year *t*. We then calculate the time-series average of each university's patent value to be the main explanatory variable, *Average PatVal(Match)*. Second, since patents are valid up to 20 years and thus generate licensing income for multiple years, we calculate a university's total license income in a year using a straight line depreciation plan to discount its annual inflation-adjusted license income in the following 20, 15, and 10 years. We then calculate the time-series average of each university's total license income to be the dependent variable. Last, we regress universities' average total license income on *Average PatVal(Match)* in Panel A. In Panel B, we use the similar approach to calculate the total number of startups. License income and patent value are in \$ millions and adjusted for inflation. The data of license income and number of startups formed are from the annual reports of the Association of University Technology Managers (AUTM) 1991-2010. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively.

Table VII: Production Function of Patent Value in U.S. Universities

After finding that university patent value is economically relevant to both patent licensing and startup formation, we analyze the production function of patent value to analyze what inputs are important correlates of valuable patents. Panel A reports the correlation matrix of five university characteristic variables, the five-year cumulative R&D expenditure (*R&D*, with a 20% obsolescence rate per year), a dummy variable indicating whether the sample university is a Carnegie-ranked research university or not (*Carnegie*), the number of full-time faculties (*Faculty*), the full-time equivalents (*FTE*) in a technology transfer office or not (*TTO*) in that year, and a dummy variable indicating whether the sample university has a medical school or not (*MedicalSchool*). In Panel B, we run pooled OLS regressions in a log-log form to estimate the Cobb-Douglas production function of patent value in universities:

 $ln(PatVal_{i,t}) = Constant + \beta_1 \cdot ln(Raul_{i,t}) + \beta_2 \cdot Carnegie_{i,t} + \beta_3 \cdot ln(Faculty_{i,t})$ $+ \beta_4 \cdot T T O_i + \beta_5 \cdot F T E_{i,t} + \beta_6 \cdot \text{Medical School}_i + \ln(\alpha_t) + \ln(\varepsilon_{i,t})$

where $PatVal_{i,t}$ is the total value of patents (measured by $PatVal(Match)$) applied by (and later granted to) university i in year t and α_t is the year fixed effect. Standard errors are clustered by university to correct for errors in autocorrelation. In Panel C, we run crosssectional regressions in which variables are averaged across sample years for each university. *PatVal* and *R&D* are in \$ millions and adjusted for inflation. *R&D* and *FTE* come from the annual reports of the Association of University Technology Managers (AUTM) 1991-2010. The number of full-time faculties and research doctorates are collected from the Carnegie reports (1994, 2000, 2005, and 2010). All other variables are collected from online searches. We follow Chan et al. (2001) and use an obsolescence rate of 20% to compute patent value capital, following. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively.

Supplemental Appendix

Benchmarking U.S. University Patent Value and Commercialization Efforts: A New Approach

Patent Value Estimation Technique of Kogan et al. (2017)

We use estimates of the value of patents granted to U.S. public firms to proxy for the value of patents granted to U.S. universities. First, we use the economic value of each patent (*PatVal*) assigned to a public firm estimated by Kogan et al. (2017), which is the 3-day change in market capitalization of this firm around the announcement of the patent, adjusted for measurement noise and various fixed effects. The technical details of the estimation technique from Kogan et al. (2017) are provided below:

A firm's three-day announcement return for patent *j* (denoted as r_j) is the sum of two underlying distributions: (i) the value of newly granted patent *j* as a fraction of the firm's market capitalization (denoted as p_j), which is assumed to follow a truncated normal distribution with a mean equal to zero and a variance equal to σ_p^2 ; and (ii) the noise component in the three-day stock return unrelated to the newly granted patent (denoted as ε_j), which follows a normal distribution with a mean zero and a variance σ_{ε}^2 . With both σ_{p}^2 and σ_{ε}^2 known, Kogan et al. compute the expected patent value following Bayes' rule:

$$
E[p_j|r_j] = \delta r_j + \sqrt{\delta} \frac{\phi\left(-\sqrt{\delta} \frac{r_j}{\sigma_{\varepsilon}}\right)}{1 - \phi\left(-\sqrt{\delta} \frac{r_j}{\sigma_{\varepsilon}}\right)} \sigma_{\varepsilon},\tag{S.1}
$$

where \emptyset and Φ denote the probability density function and cumulative distribution function of a standard normal distribution, respectively, and δ is the ratio of signal to noise as defined below:

$$
\delta = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_\varepsilon^2}.
$$

In calculating $E[p_j|\mathbf{r}_j]$ in Equation (S.1), the values of two variables are vital: δ and σ_{ε}^2 . Kogan et al. assume δ to be constant across firms and time but allow σ_{ε}^2 to vary across firms and time. To estimate δ , they execute the following panel regression to compute the increase in volatility of firm returns around announcement days of newly granted patents:

$$
\ln(r_{fd})^2 = \gamma I_{fd} + cZ_{fd} + u_{fd}
$$

1

where r_{fd} is the three-day market-adjusted return of firm *f*, starting from day *d*. I_{fd} is a dummy for the day when there is any newly granted patent(s). *Z* is a vector of controls including the day-of-week fixed effects and the firm-year joint fixed effects. The calculation based on the above equation produces the estimate $\hat{v} = 0.0146$. Using the value of $\hat{\gamma}$, δ is estimated by the following equation:

$$
\hat{\delta} = (e^{\hat{\gamma}} - \mathbf{1})(1 - 2C_0^2 + e^{\hat{\gamma}} C_0^2)^{-1},
$$

where $C_0 = \phi(0) / (1 - \phi(0))$. The resulting estimate of $\hat{\delta} = 0.0145$.

To estimate the firm- and year-specific σ_{ε}^2 (we use the notation σ_{ε}^2 instead), Kogan et al. first follow Anderson and Terasvirta (2009) to non-parametrically estimate the market-adjusted daily return variance, σ_{ft}^2 , for each firm and each year. With the estimate of σ_{ft}^2 , the fraction of trading days that are announcement days for a firm in a year, d_{ft} , and the estimate of $\hat{\gamma}$, they compute the variance of the measurement error in the following equation:

$$
\hat{\sigma}_{\varepsilon,ft}^2 = 3\hat{\sigma}_{ft}^2 \left(1 + 3d_{ft}\hat{\gamma}/(1 - \hat{\gamma})\right)^{-1}.
$$

Inserting the previously estimated $\hat{\delta}$ and $\hat{\sigma}^2_{\epsilon_0 f t}$, they calculate the value of $E[p_j|\eta]$ in Equation (S.7).¹ Finally, they employ the following equation to compute the market value of patent *j*, θ_j , as the product of the estimated stock return associated with the patent, $E[p_j|r_j]$, multiplied by the market capitalization, M_{j} , of the firm granted with patent *j* on the day prior to the patent issuance announcement:

$$
\theta_j = (1 - \pi)^{-1} \frac{1}{N_j} E[\widehat{p_j} | r_j] M_j,
$$

Where π is the unconditional probability of a successful patent application (estimated to be 0.56 in Carley et al. (2015)), and N_j is the number of patents granted to the same firm on the same day.

 $\mathbb{E}[\mathbf{p}_j|\mathbf{r}_j]$ is not always positive in Kogan et al. (2017). Negative estimates are excluded.

Simulations to Check the Sensitivity of Patent Value Matches for Corporate Patents

We design a simulation process to verify our matching method based on the estimated patent value of Kogan et al. (2017). In each simulation, we randomly choose 5% of corporate patents as the out-of-sample from the universe of corporate patents. The ratio of 5% approximates the relative size of university patents to corporate patents. For each patent of the out-of-sample (the focal patent), we estimate its value by benchmarking it to similar patents from the remaining 95% (in-sample). Specifically, the value of the focal patent is set to that of an in-sample patent that is in the same patent sub-category, granted in the same year, and has the shortest sum of distances to the focal patent in terms of quality and generality.² If there are multiple in-sample patents satisfying the above criteria, we set the focal patent value to be the median value of those in-sample patents. After we finish the matching for each patent, we collect the simulated value for the 5% out-of-sample patents, we calculate their median and mean and then calculate the absolute percentage deviation between the median (mean) and the full-sample median (mean). By repeating the aforementioned procedure 500 times, we collect 500 absolute percentage deviations of median and mean and plot their frequency in a histogram presented in Figure OA.I. Overall, we find that the deviation is modest as most of them are within a 2% range.

$$
\frac{|4-6|}{4+6}+\frac{|0.3-0.1|}{0.3+0.1}=0.7.
$$

² The distance along a patent characteristic is measured by the absolute value of the difference between two values divided by their sum. For example, if the focal patent has its forward citation as 4 and its generality as 0.3 and an in-sample patent has it forward citation as 6 and its generality as 0.1, then their distance is equal to 0.7, calculated as

Figure OA.I: Accuracy of the Estimation Method of Patent Value.

We examine the accuracy of our matching method in estimating patent value. To do so, we randomly draw 5% of the corporate patents by their grant years and subcategories as focal patents. We benchmark the value of each focal corporate patent with the full sample of corporate patents using the matching method described in Table IV. We choose 5% of the corporate patents as focal because the number of corporate patents is 20 times the number of university patents in the full sample, as shown in Table IV. We repeat this simulation procedure 500 times. After each simulation run, we calculate the absolute percentage deviation of the median and mean estimated value of focal corporate patents from their median and mean true value, respectively. This distribution is shown in the blue and red histograms, respectively.

Table OA.I: Distributions of Characteristics of Patents Granted to Listed Firms and Universities in the U.S.

We compare the distribution of patent quality/importance (i.e., the citations received in five years after the patent is granted), patent originality (i.e., one minus the Herfindahl-Hirschman Index (HHI) of citations given to other patents over patent subcategories), patent generality (i.e., one minus the HHI of citations received from other patents over patent subcategories), and patent basicness (i.e., the ratio of the number of references to prior "non-patent documents" divided by the total references) of patents granted to listed public firms and universities. The definitions of patent originality, generality, and basicness follow Trajtenberg, Henderson, and Jaffe (1997). We also compare their distributions in the following three periods: 1976-1985, 1986-1995, and 1996-2010. We split our whole sample period (1976-2010) into three almost equal sub-periods to examine the evolution of patent forward citation, patent originality, and patent generality over time. We split our sample at 1985-1986 due to the adoption of the Bayh–Dole Act at 1980 and the surge of personal computer industry at early 1980s. We also split our sample at 1995-1996 for the ".com bubble" started around 1996-1997. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively, when comparing the mean of universities' patents with the mean of listed firms' patents.

Panel A

Panel A reports summary statistics for patent quality/importance (number of forward citations within 5 years) (A1), patent generality (A2), patent originality (A3), patent basicness (A4), and number of claims (A5) in the following three periods: 1976-1985, 1986-1995, and 1996-2010.

Panel B

Panel B shows the distributions of patent quality/importance (number of forward citations within 5 years), patent originality, patent generality, and patent basicness of listed public firms and universities. For quality/importance, we compute the frequency of observations within each category of entities for twelve intervals (citation equal to 0 (5th-20th percentile), citation equal to 1 (25th-40th percentile), citation equal to 2 (45th-50th percentile), citation equal to 3 (55th-60th percentile), citation equal to 4 (65th percentile), citation equal to 5 (70th percentile), citation equal to 6 (75th percentile), citation equal to 7 (80th percentile), citation larger than 7 and smaller than or equal to 9 (85th percentile), citation larger than 9 and smaller than or equal to 12 (90th percentile), citation larger than 12 and smaller than or equal to 19 (95th percentile), and citation larger than 19). For the distributions of patent originality, generality, and basicness (all bounded from 0 to 1), we report their frequencies in each equal bin between 0 to 1. For number of claims, we compute the frequency of observations within each category of entities for ten intervals (smaller than 4 (10th percentile), larger than 4 and smaller than 7 (20th percentile), larger than 7 and smaller than 9 (30th percentile), larger than 9 and smaller than 11 (40th percentile), larger than 11 and smaller than 14 (50th percentile), larger than 14 and smaller than 17 (60th percentile), larger than 17 and smaller than 20 (70th percentile), larger than 20 and smaller than 23 (80th percentile), larger than 23 and smaller than 31 (90th percentile), and larger than 31). We also compare their distributions in the following four periods: 1976-1985, 1986-1995, 1996-2010, and 1976-2010 (all years).

13

14

Panel C

Panel C compares the number of patents granted in each subcategory of listed public firms and universities. Percentages are reported within each category of entities. We also compare their distributions in the following four periods: 1976-1985, 1986-1995, 1996-2010, and 1976-2010 (all years). A Kolmogorov-Smirnov test confirms that distributions of patent counts across subcategories are statistically different (p-value<0.01) between universities and public firms.

Panel D

Panel D reports the mean patent quality/importance (number of forward citations within 5 years) in each subcategory within each category of entities (i.e., public firm and university). We report the statistical significance of the difference between corporates and universities in each subcategory with two-sample t-test. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively. We also compare their distributions in the following four periods: 1976-1985, 1986-1995, 1996-2010, and 1976-2010 (all years).

Panel E

Panel E reports the mean patent originality (one minus the HHI of citations given to other patents over patent subcategories) in each subcategory within each category of entities (i.e., corporate and university). We test the statistical significance of the difference between corporates and universities in each subcategory with two-sample t-test. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively. We also compare their distributions in the following four periods: 1976-1985, 1986- 1995, 1996-2010, and 1976-2010 (all years).

Panel F

Panel F reports the mean patent generality (one minus the HHI of citations received from other patents over patent subcategories) in each subcategory within each category of entities (public firm and university). We report the statistical significance of the difference between public firms and universities in each subcategory with two-sample t-test. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively. We also report their distributions in the following four periods: 1976-1985, 1986-1995, 1996-2010, and 1976-2010 (all years).

Panel G

Panel G reports the mean patent basicness (the ratio of the number of references to prior "non-patent documents" divided by the total references) in each subcategory within each category of entities (public firm and university). We report the statistical significance of the difference between public firms and universities in each subcategory with two-sample t-test. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively. We also report their distributions in the following four periods: 1976-1985, 1986-1995, 1996-2010, and 1976-2010 (all years).

Panel H

Panel H reports the mean number of claims in each subcategory within each category of entities (public firm and university). We report the statistical significance of the difference between public firms and universities in each subcategory with two-sample t-test. ***, **, * indicate significance levels of 1%, 5%, and 10%, respectively. We also report their distributions in the following four periods: 1976-1985, 1986-1995, 1996-2010, and 1976-2010 (all years).

Table OA.II: Market-Based Patent Value and Actual University Patent Revenue in the Sub-category of Drug.

We run the same regressions as in Table V but in the sample of patents in the drug subcategory.

Table OA.III: U.S. Universities and Their PDPASSs.

We manually match the university names with their PDPASSs using the assignee file (1976-2006) of the NBER patent data. We first examine the NBER patent assignee name file, focus on the assignees in the category of "U.S. University," and manually harmonize the PDPASSs for each university. To ensure full coverage, we search the university names in other categories and extract their PDPASSs. For example, Purdue Research Foundation of the Purdue University is in the category of "U.S. Institute." Our resulting sample is 362 universities.

