A Study of Factors Influencing Invention Patent Grant in Biotechnology (Postprint)

Authors: Feng Rentao, Yu Xiang

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] After an invention patent application is published and before a substantive examination decision is rendered, it remains in a state of legal uncertainty. Evaluating the prospects for patent grant during this stage is of great theoretical and practical significance. [Method/Process] This study examines 88,304 invention patent applications filed by domestic applicants in the biotechnology field with the China National Intellectual Property Administration (CNIPA) from 1985 to 2014. Indicators based on patent documents were constructed, and Logistic regression models were employed to explore factors influencing patent grant. Grouped regression and coefficient difference tests were used to analyze variations in patent grant determinants across different types of innovation entities. [Results/Conclusion] Logistic regression results indicate that in the biotechnology field, indicators such as the number of claims, number of inventors, whether early publication was requested, word count of the first claim, and number of pages in the patent document are all positively correlated with the probability of patent grant. The patent grant rate for enterprises is lower than that for universities and research institutions but higher than that for individual applicants. Grouped regression coefficient difference tests reveal significant differences in how variables affect patent grant among different types of innovation entities. Therefore, it is essential to categorize different innovation entities separately in patent grant analysis.

Full Text

A Study on Influencing Factors of Patent Grant in the Field of Biotechnology

Feng Rentao[1], Yu Xiang[1,2]

[1]Fujian Province Institute of Intellectual Property, Fujian University of Technology, Fuzhou 350118

[2]School of Management, Huazhong University of Science and Technology, Wuhan 430074

Abstract: [Purpose/Significance] After an invention patent application is published, it remains in a pending legal status until a substantive examination decision is made. Assessing the prospects for patent grant during this stage is both theoretically and practically meaningful. [Method/Process] This study examines 88,304 invention patents applied for by domestic applicants in the biotechnology field at the China National Intellectual Property Administration (CNIPA) from 1985 to 2014. Based on patent documents, we construct relevant indicators and use logistic regression models to explore factors affecting patent grant. Through grouped regression and coefficient difference tests, we analyze variations in influencing factors across different types of innovation entities. [Result/Conclusion] Logistic regression results show that in the biotechnology field, the number of claims, number of inventors, whether early publication is requested, word count of the first claim, and number of patent document pages are all positively correlated with the probability of patent grant. Enterprise patents have lower grant rates than university and research institution patents but higher rates than individual patents. Grouped regression coefficient difference tests indicate significant differences in how variables affect patent grant across different innovation entities. Therefore, classification by innovation entity is necessary in patent grant analysis.

Keywords: patent grant; logistic regression; patent quality; patent maintenance time; biotechnology

1. Introduction

Under China's Patent Law, patents are categorized into three types: invention, utility model, and design. Among these, invention patents require substantive examination and are only granted after meeting requirements for novelty, inventiveness, and practicality, along with other conditions. Consequently, invention patents hold higher value than utility models and designs, which are directly granted after preliminary examination, and are considered more appropriate indicators of innovation capability [1].

A key distinction of invention patents is their publication before substantive examination. Whether a published application will be granted remains uncertain until examined by a patent examiner. The average period from publication to a grant decision exceeds two years [2], with some cases taking over five years. During this time, patents remain in a state of legal limbo, which explains why many studies on invention patent grants are forced to exclude the most recent data.

The grant outcome directly affects applicants' interests and concerns competitors as well. For applicants, an ungranted invention patent that has disclosed

technical information means anyone can freely use the technology, resulting in significant losses in capital investment and technical information disclosure. From a competitive perspective, timely assessment of grant probability after a competitor's patent publication helps in strategic planning and maintaining a favorable competitive position. Therefore, evaluating grant prospects before patent authorization is urgently needed in practice.

Obtaining grant is a prerequisite for patent value creation, and assessing grant probability forms the basis for patent quality evaluation during the application stage. Several indicators can be extracted from published patent documents to evaluate patent quality. Currently, patent maintenance duration is a widely recognized quality indicator [3], but it suffers from substantial time lag. Scholars have thus used maintenance duration as a quality metric to explore how indicators extracted from patent documents affect patent quality. Existing research has found that patent document length, number of claims, number of inventors, and number of IPC classification codes all have significant effects [4-5]. Granted invention patents typically exhibit higher quality than ungranted applications [6], suggesting that factors affecting maintenance duration likely influence grant probability as well. In empirical studies, W.M. Schuster et al. demonstrated using U.S. patent data that the number of inventors, patent family size, and technical field significantly affect patent grant [7]. However, whether these relationships exist in Chinese patents remains unexamined.

Since different technology fields are at different development stages and patents play varying roles in protecting technological innovation, patent grant situations and influencing factors may differ substantially across fields. This study focuses on the biotechnology field, primarily because biotechnological innovation heavily depends on patent protection. Biotechnology is also a typical knowledge-based field with a high proportion of university and research institution patent applications, and abundant existing research facilitates comparative analysis [9-10].

Whether an individual invention is granted is determined by CNIPA examiners after examination. Most indicators highly correlated with patent value can be obtained from published patent documents before examiners make their decisions. If these indicators are indeed highly correlated with patent value, they may enable prediction of grant prospects. This study constructs models from published invention patent document data to predict grant probability, providing a foundation for determining patent value during the pending period and testing whether indicators that significantly affect maintenance duration also influence patent grant.

Different innovation entities—enterprises, universities, and individuals—exhibit different operational models, resource bases, and patent application motivations, leading to variations in the technical foundation, objectives, and strategies of their patent applications, which may result in different grant outcomes. Existing research has identified significant differences in patent value across innovation entities, with enterprise invention patents showing significantly longer maintenance durations than individual and university patents [8]. In foreign studies,

E. Sapsalis et al. found minimal differences in quality and influencing factors between university and enterprise biotechnology patents in Belgium, though they did not test for significance [9]. In China, do significant differences exist in grant rates among enterprises, universities, individuals, and other entities? Do influencing factors differ across these entities? These questions have received limited research attention. Quantitative analysis of such differences and their causes can help understand how patent applications from different entities are formed and provide targeted recommendations for improving grant rates.

2. Data and Model

2.1 Data Invention patent data were obtained through the Incopat online patent database platform. Biotechnology patents were identified by their primary IPC classification codes. Following OECD's IPC classifications for biotechnology [11], we retrieved 88,304 invention patents applied for by domestic applicants at CNIPA from 1985 to 2014 (excluding patents still under examination), which constitute our study sample.

According to the OECD classification, biotechnology IPC codes include: A01H1/00, A01H4/00, A61K38/00, A61K39/00, A61K48/00, C02F3/34, C07G (11/00, 13/00, 15/00), C07K (4/00, 14/00, 16/00, 17/00, 19/00), C12M, C12N, C12P, C12Q, C12S, G01N27/327, G01N33/ (53*, 54*, 55*, 57*, 68, 74, 76, 78, 88, 92).

After publication, invention patent applications enter substantive examination upon applicant request. The process from examination entry to grant typically takes about two years (only after publication), and even longer from the application date. As of July 16, 2020 (our data collection date), 2.26% of patents applied for in 2014 remained under examination. From 2015 onward, the examination rate for all years exceeded 7%. To analyze a longer time series while minimizing the impact of pending patents, this study only considers invention patents applied for at CNIPA by domestic applicants through 2014. Given that patent funding and policy environments differ in Taiwan, Hong Kong, and Macau, patents from these regions are excluded.

2.2 Variables The dependent variable is whether an invention patent was granted (1 = granted, 0 = not granted). Drawing from factors considered in patent maintenance duration studies, we use data obtainable from published patent documents as independent variables, including: number of claims, patent document pages, whether early publication was requested, number of IPC classifications, number of inventors, whether foreign applications were filed, and whether a patent agency was used.

Patent system modifications may affect grant conditions and policies. Since China's patent system was established in 1985, the Patent Law has been amended three times (1992, 2000, 2008). We use the implementation years of these amendments as turning points, dividing the study period into four

sub-periods: 1985-1992, 1993-2000, 2001-2008, and 2009-2014, to test the impact of legal changes on grant rates.

Significant regional differences exist in China's technological innovation capacity and economic development, which may affect invention patent application quality. We therefore divide the country into four regions—eastern, central, western, and northeastern—following National Bureau of Statistics classifications. As discussed, innovation entity type may also significantly affect patent grant. We categorize innovation entities into four types: enterprises, individuals, universities, and others, using the first applicant as the reference when multiple applicants exist.

All variable definitions, expected effects on patent grant, and relevant literature are summarized in Table 1 . Expected relationships are derived from quantitative analyses of factors affecting patent maintenance duration or other quality indicators in existing research.

2.3 Model Since the dependent variable (patent grant) is binary, we employ logistic regression, a standard method for examining relationships between binary outcomes and independent variables widely used in patent maintenance duration studies.

Innovation entity type is handled in two ways: First, dummy variables are created for the four entity types (enterprise, university, individual, other), with enterprise as the reference (0). When multiple applicants exist, the first applicant determines classification. Second, dummy variables for eastern, central, western, and northeastern regions are created, with eastern as the reference (0).

3. Results

3.1 Descriptive Statistics Table 2 presents descriptive statistics. The rightmost three columns show means for granted (grant=1) and ungranted (grant=0) patents and t-test p-values. Except for first claim word count, all other variables have higher means in the granted sample, consistent with maintenance duration studies and Table 1 expectations.

Figure 1 [Figure 1: see original paper] shows annual invention patent applications by innovation entity. Before 2000, enterprises had the lowest application volume. After 2001, enterprise applications gradually surpassed individual applications. In most years, however, universities and other entities exceeded both enterprises and individuals. Biotechnology heavily depends on basic research; before 2000, China's biotechnology industrialization scale was small and enterprise innovation capacity was weak, resulting in fewer applications. After 2009, biotechnology companies such as BGI experienced explosive growth in patent applications, causing total enterprise patents to exceed individual patents and approach university levels.

Grant rates differ significantly across innovation entities. From 1985-2014, the

average grant rate for Chinese invention patents was 56.3%. However, average grant rates for enterprises, individuals, universities, and others were 49.9%, 39.2%, 61.5%, and 63.05%, respectively. Figure 2 [Figure 2: see original paper] shows annual grant rates by entity type. Before the second Patent Law amendment in 2000, all entities exhibited large fluctuations, primarily due to relatively small application volumes, particularly evident before the first amendment in 1992. After 2001, all four entities had substantial application volumes (Figure 1), and grant rates stabilized. Overall, university patents had grant rates similar to other entities and significantly higher than enterprises and individuals. After 2009, other entities' grant rates consistently exceeded university rates; since research institutions comprise the majority of "other entities," their grant rates are higher. These findings contrast sharply with patent value assessments based on maintenance duration, as existing research consistently shows enterprise patents have significantly longer maintenance durations than university patents [8]. However, our conclusions align with Liu Lijun and Song Min's research on agricultural biotechnology patents [18].

3.2 Logistic Regression Results Table 3 presents logistic regression results. Columns 2-3 show results for all 88,304 patents. The coefficients for number of claims, number of inventors, number of IPC classifications, foreign applications, cooperative applications, early publication requests, and patent document pages are all positive and significant at the 0.1% level, indicating these variables significantly increase grant probability. Except for first claim word count, these results align with maintenance duration studies—variables that enhance post-grant patent value also improve grant probability. For example, the early publication coefficient is 0.666, indicating that, holding other variables constant, patents requesting early publication have 1.936 times higher odds of grant (OR = e^0.666). Similar interpretations apply to other variables.

Notably, first claim word count has a coefficient of 0.060—small but significant at the 0.01% level. First claim word count partially reflects patent scope; generally, more words indicate more technical features and narrower scope. During examination, patents with narrower scope more easily satisfy novelty and inventiveness requirements, undoubtedly increasing grant probability. However, post-grant, narrower scope typically corresponds to lower value and shorter maintenance duration [4]. Thus, first claim word count's effect on grant probability opposes its effect on maintenance duration.

Columns 4-11 in Table 3 show regression results by innovation entity. Except for cooperative applications, other variables have consistent signs across entity types, indicating that our patent document-based indicators for predicting grant probability are applicable across different innovation entities.

Cooperative applications warrant attention: coefficients are significantly positive in regressions for all patents and enterprise patents, but significantly negative for university patents. Enterprise cooperative patents (with enterprise as first applicant) have significantly higher grant rates than other enterprise

patents, while university cooperative patents (with university as first applicant) have significantly lower grant rates than other university patents. This suggests cooperative applications do not universally improve grant prospects but depend on specific partners.

Regarding regional effects, using eastern region as baseline, central and western regions show significantly higher grant probabilities than eastern regions, while northeastern regions show significantly lower probabilities. Although results vary across grouped regressions, central regions consistently exhibit significantly higher grant probabilities than eastern regions. Controlling for patent document indicators, regional differences may primarily reflect provincial patent subsidies and technological innovation capacity distributions. While eastern regions lead in economic development and overall innovation capacity, their invention patent subsidies are typically higher than central and western regions, and existing research agrees that subsidies generally reduce patent quality [19]. Subsidies during the application stage encourage more low-quality applications, reducing grant rates. Additionally, eastern regions may involve more cutting-edge biotechnology, which, while more valuable post-grant, faces fiercer competition and greater examination difficulty in emerging technologies. Regional effects on grant thus differ completely from effects on maintenance duration, as existing research consistently shows eastern patents have significantly longer maintenance durations than central and western patents [20].

Temporal effects are only significant in the second period (1993-2000) for the full sample and enterprise sample, indicating minimal differences across periods otherwise. This aligns with Figure 2, which shows relatively stable grant rates except for 1999-2000. Patent law amendments appear to have limited impact on grant rates. The core grant requirements are novelty and inventiveness; the law only modified novelty provisions in the 2008 amendment, and all amendments avoided inventiveness. The novelty modifications merely added foreign public use as a novelty-defeating circumstance, which is extremely rare in practice, especially for domestic applicants. Since inventiveness presupposes novelty, legislative amendments understandably had limited impact on grant outcomes.

3.3 Grouped Regression Coefficient Difference Tests Although Table 3 shows most variables have consistent effects across innovation entities, quantitative analysis is needed to determine whether coefficient differences are significant. If differences are insignificant, treating all patents as one group would be appropriate; otherwise, separate analysis by entity type is necessary.

Table 4 presents coefficient difference test results from the grouped regressions in Table 3. Overall, at the 0.01 significance level, no variable shows no significant difference across all entity pairs. For example, the effect of claim count differs significantly between individual and university patents, though both coefficients are significantly positive. Similar patterns exist for other variables. Table 4 shows that except for claim count, all other variables exhibit significant differences in at least three of the six pairwise comparisons, indicating that grouped

regression coefficients differ significantly across innovation entities. Therefore, analyzing patent grant influencing factors using indicators extracted from patent documents is more appropriate when conducted separately by entity type rather than pooling all patents.

3.4 Robustness Checks We conduct robustness checks using two approaches: (1) Modifying temporal and spatial controls by replacing four regions with 31 provinces and four time periods with 30 individual years (1985-2014). Overall and grouped regression results show minor changes in coefficients and significance levels, but variable signs remain consistent with Table 3, and grouped coefficient differences remain largely unchanged except for slight significance variations. (2) Standardizing variables to overcome scale differences and conducting logistic regression on standardized variables. The above conclusions remain essentially unchanged.

4. Conclusion and Discussion

This study examines how variables constructed from biotechnology patent documents affect invention patent grant and whether these effects differ across innovation entities. Logistic regression yields the following conclusions:

(1) In the biotechnology field, the number of claims, number of inventors, foreign applications, early publication requests, number of IPC classifications, first claim word count, and patent document pages are all positively correlated with grant probability. Except for first claim word count, cooperative applications, and regional variables, other variables' effects on grant align with their effects on maintenance duration—variables that enhance post-grant value also improve grant probability.

(2) Regression coefficients for relevant variables differ significantly across innovation entity groups, indicating that different entity types experience varying effects from these factors on invention patent grant. Therefore, classifying by innovation entity may be more appropriate when analyzing grant prospects in biotechnology.

Scientifically predicting patent grant prospects has important practical significance. Our model provides a quantitative analytical method for applicants, competitors, and policymakers to predict grant probabilities for large patent portfolios. However, this study has limitations. First, it focuses only on biotechnology patents; whether similar patterns exist in other fields requires further investigation. Feng Rentao [4] found significant differences in patent quality across most technology fields, so our results may not generalize. Second, while most coefficients are significant, Pseudo R^2 values range from 0.028 to 0.129, indicating many unconsidered factors affect patent grant. This is a common issue in patent document-based analyses, as corresponding Pseudo R^2 values are consistently low in domestic and international studies on patent value indicators extracted from documents [21]. How to combine patent document factors with

other potential influences for more accurate grant prediction warrants further exploration.

References

[1] Zhang Yafeng, Liu Haibo, Chen Guanghua, et al. Is patent a good innovation measurement indicator?[J]. Foreign Economics & Management, 2018, 40(6): 3-16.

[2] Wen Jiachun, Lu Bingke. Factors influencing patent substantive examination cycle[J]. China Science and Technology Forum, 2016, (12): 90-97.

[3] Guellec D, de la Potterie B P. Applications, grants and the value of patent[J]. Economics Letters, 2000, 69(1): 109-114.

[4] Feng Rentao. Analysis of factors influencing patent maintenance duration based on patent documents[J]. Information and Documentation Services, 2020, 39(7): 202-207.

[5] Song Shuang. Research on influencing factors of patent maintenance time in China: Based on patent quality considerations[J]. Library and Information Service, 2013, 57(7): 96-100, 105.

[6] Grimaldi M, Cricelli L. Indexes of patent value: A systematic literature review and classification[J]. Knowledge Management Research & Practice, 2020, 18(2): 214-233.

[7] Schuster W M, Davis R E, Schley K, et al. An empirical study of patent grant rates as a function of race and gender[J]. American Business Law Journal, 2020, 57(2): 281-319.

[8] Qiao Yongzhong. Empirical research on maintenance information of invention patents by different types of innovators[J]. Studies in Science of Science, 2011, 29(3): 442-447.

[9] Sapsalis E, de la Potterie B P, Navon R. Academic versus industry patenting: An in-depth analysis of what determines patent value[J]. Research Policy, 2006, 35(10): 1631-1645.

[10] Mowery D C, Ziedonis A A. Academic patent quality and quantity before and after the Bayh-Dole Act in the United States[J]. Research Policy, 2002, 31(3): 399-418.

[11] OECD. OECD patent databases, identifying technology areas for patents[EB/OL]. [2020-07-17]. https://www.oecd.org/sti/inn/o/40807441.pdf.

[12] Wu Hong, Fu Xiuying, Dong Kun. Empirical analysis of factors influencing patent maintenance time: Taking fuel cell patents as an example[J]. Library and Information Service, 2013, 57(24): 112-116, 100.

[13] Lerner J. The importance of patent scope: An empirical analysis[J]. The RAND Journal of Economics, 1994, 25(2): 319-333.

[14] Liu Xuefeng, Gao Xing. Research on influencing factors of wind technology invention patent maintenance time in China[J]. Science Research Management, 2015, 36(10): 139-145.

[15] Ye Jingyi, Li Chenle, Lei Zhen, et al. Early publication system, patent quality and technical knowledge dissemination[J]. The Journal of World Economy, 2012(8): 115-133.

[16] Cai Zhonghua, Hou Aoyu, Ma Huan. Empirical research on factors influencing patent maintenance time[J]. Science and Technology Management Research, 2015, 35(21): 160-163.

[17] Liu Yun, Wang Xiaoli, Yan Zhe. Research on patent quality measurement and regional comparison: Taking China's graphene industry as an example[J]. Science of Science and Management of S.& T., 2019, 40(9): 18-34.

[18] Liu Lijun, Song Min. Quality of agricultural patents in China: Comparison across different application periods, applicants and technical fields[J]. Scientia Agricultura Sinica, 2012, 45(17): 3617-3623.

[19] Zhang Jie. The puzzle of China's patent growth: Micro-empirical evidence from the perspective of local government policy incentives[J]. Wuhan University Journal (Philosophy & Social Sciences), 2019, 72(1): 85-103.

[20] Ma Ru, Luo Hui, Wang Hongwei, et al. Research on evaluation index system and measurement of high-quality regional economic development in China[J]. China Soft Science, 2019(7): 60-67.

[21] Fischer T, Leidinger J. Testing patent value indicators on directly observed patent value: An empirical analysis of Ocean Tomo patent auctions[J]. Research Policy, 2014, 43(3): 519-529.

Author Contributions: Feng Rentao: literature review, data collection and analysis, manuscript writing; Yu Xiang: research design, discussion of findings, manuscript revision.

Note: Figure translations are in progress. See original paper for figures.