The Japan Trademarks Dataset: A First Analysis [*]

Invited Researcher: Jeremy Nathan SHEFF [**]

This paper discloses and describes a new research dataset on the trademark registration system of Japan. The dataset was built using data provided by the National Center for Industrial Property Information and Training (INPIT), and builds on earlier work by Motohashi et al. for the National Institute of Science and Technology Policy (NISTEP). The current dataset includes application-level information including prosecution history events; registration and maintenance events; mark subject matter and text; and other information. Ancillary files to be made available in future releases include applicant/owner, attorney, and examiner information; Paris Convention and Madrid Protocol claims; goods and services classifications. This paper catalogs the dataset contents and provides some overall descriptive information. In addition, this paper will propose and offer illustrations of possibilities for both domestically-focused and comparative empirical research on industrial property made possible by the new dataset, for example by collating and comparing the data described herein with comparable datasets from other national intellectual property offices (particularly the USPTO) and related sources such as economic data.

I. Background

Scholars and policymakers analyzing the workings of industrial property systems—and lawyers, businesspeople, and firms who interact with those systems—are increasingly turning to empirical data to guide their analyses and judgments. While patent systems have attracted particular attention, trademarks are another form of valuable industrial property—by some measures the *most* valuable[1]—whose owners face risks and tradeoffs when they interact with the legal system. This can be a particularly complex problem when individuals and firms seek trademark protection in multiple countries. Lawyers often rely on judgment based on personal experience when advising clients on appropriate courses of action in registering trademarks. Their judgment would likely be better informed by empirical evidence from beyond their own personal experience. Similarly, empirical analysis of fine-grained national and international trademark data, and comparison to similar data sources from other countries, may assist Japanese policymakers and their counterparts in other nations' intellectual property offices, as well as scholars trying to understand the workings of the trademark system and its relationship to the legal profession and the economy.

---

[1] *See, e.g.*, Danny M. Awdeh & Brian R. Westley, *More Valuable than Patents? Recognizing—and Leveraging—Trademark Assets While Avoiding Potential Pitfalls Along the Way*, LANDSLIDE (Sept./Oct. 2015), available at http://www.americanbar.org/publications/landslide/2015-16/september-october/more_valuable_patents_recognizingand_leveragingtrademark_assets_while_avoiding_potential_pitfalls_along_way.html (last visited July 15, 2016).

This project contributes to the store of public knowledge regarding the workings of trademark registration systems, both internationally and within Japan. It does so by collecting and reporting empirical data on Japan's trademark registration system—including data on international registrations—which may be used in conjunction with similar datasets from other national and supra-national offices to comparatively assess applicant and examiner behaviors, application characteristics and trends, registration events and outcomes, and interactions between trademark systems and broader economic and social indicators. As the culmination of that effort, this Report discloses and describes a new data resource: the Japan Trademarks Dataset, a research-quality dataset containing comprehensive, application-level information on the trademark registration system administered by the Japan Patent Office (JPO). The Report further offers examples of the kinds of internal and comparative analyses that the Japan Trademarks Dataset makes possible, particularly in conjunction with the Trademark Case Files Dataset released by the United States Patent and Trademark Office (USPTO).

1. Trademark Registration in Japan, with Comparison to the United States

Japan's trademark system is registration-based, as contrasted with the use-based system in the United States. Thus, where a federal US trademark registration may only be issued upon a showing of actual use of a mark in commerce,[2] a trademark registration in Japan may be issued upon filing of an application for registered rights even without establishing actual use.[3] However, failure to use the registered mark can result in rescission of registered rights under the Japan Trademark Act ("JTA").[4]

In order to establish registered trademark rights in Japan, the claimant must file an appropriate application and pay a fee. Trademark applications are subject to a multi-stage review by the JPO. If the JPO finds no grounds for refusal, it will issue a final decision of registration.[5] Upon payment of the appropriate fee, JPO then issues the registration certificate to the applicant and publishes the registration in the JPO's Trademark Gazette.[6] Once a registration issues, its holder enjoys "an exclusive right to use the registered trademark in connection with the designated goods or designated services."[7] Registered trademark rights last for ten years, but can be renewed

---

[2] *See* 15 U.S.C. § 1051 (limiting registration to marks that have been actually used in commerce, though allowing applications for registration before use commences based on a verified declaration of bona-fide intent to use the mark in commerce).

[3] 商標法[Trademark Act] Law No. 127 of April 13, 1959, as last amended by Act No. 55 of July 10, 2015 [hereinafter "JTA"], art. 3 (allowing registration of marks intended to be used in commerce); *id.* art. 18(1) ("A trademark right shall arise upon registration of establishment of such right").

[4] *See id.* art. 50 (establishing a right for any party to bring an action at JPO for extinguishment of registered trademark rights on grounds of three consecutive years of non-use).

[5] JTA art. 16.

[6] *Id.* art. 18.

[7] *Id.* art. 25; *see also id.* art. 30.

indefinitely for additional ten-year terms upon filing of a renewal application and payment of appropriate fees within six months of each scheduled expiration; failure to renew results in extinguishment of the trademark right.[8] Other features of Japan's trademark registration system of relevance to the Japan Trademarks Dataset—with comparison to the US trademark system—are reviewed at greater length in the full Report.

2. Previous Studies

[Omitted]

II. Data Sources and Dataset Construction

The Japan Trademarks Dataset was constructed using archived Japan Patent Office (JPO) records made available in digital format by the National Center for Industrial Property Information and Training (INPIT) at the request of the Institute of Intellectual Property (IIP), which obtained the data on my behalf in connection with my participation in IIP's Invited Overseas Researcher Program. I am grateful to IIP and to INPIT for their support and assistance.

1. Sources

The INPIT data consists of hundreds of periodic archives covering a period from late 2001 to mid-2015, plus some additional serial archives pre-dating the late 2001 archives. INPIT's archives are encoded in SGML format. A detailed data specification (in Japanese only) is available at http://www.inpit.go.jp/info/standard/download/standard_dl/sgml5.3.html.

2. Construction

I extracted, cleaned, and organized the data from the SGML documents provided by INPIT using a combination of bash scripts, ruby programs, and STATA do-files. With respect to the bash scripts and ruby programs, I was greatly assisted by Kenta Ikeuchi, a Research Fellow at the Research Institutes of Economy, Trade, and Industry (RIETI) and one of the authors of a NISTEP Research Paper undertaking a similar project.[9] Mr. Ikeuchi, on behalf of his co-authors and with the consent of NISTEP, provided me with several of the scripts and programs commissioned by his

---

[8] *Id.* arts. 19-20, 40(2).
[9] Kazuyuki Motohashi, Kenta Ikeuchi and Jianwei Dang, *Constructing a Database of Design and Trademark Rights*, NISTEP Research Material No.249 (Apr. 28, 2016), available at http://www.nistep.go.jp/archives/27214 (last visited July 15, 2016).

research team to construct their datasets. I used several of these scripts and programs as a starting-point for my own code, and they were invaluable in helping me quickly understand the structure of the INPIT data. This paper and the accompanying dataset thus build on the earlier work done by Motohashi, Ikeuchi, and Dang by extracting, analyzing and reporting a broader selection of data from the INPIT trademark archives and providing additional descriptive and comparative analysis. I am very grateful to Mr. Ikeuchi and his colleagues, without whose generous assistance construction of the Japan Trademarks Dataset would have been a much slower and more difficult process.

III. Dataset Components

1. Overview

The Japan Trademarks Dataset contains rich bibliographic and historical data on 4,518,187 unique applications for registration of trademark rights in Japan.

2. Coverage
[Omitted]

3. Component Data Files
[Omitted]

(1) Primary Data File
[Omitted]

(2) Ancillary Data Files
[Omitted]

IV. Research Applications

The Japan Trademarks Dataset has a broad array of potential research applications. This Section will provide some illustrative examples.

1. Historical Research

The deep historical data preserved by JPO allows for interesting historical research into Japan's trademark registration system, but actually doing such research requires an accessible archive. The Japan Trademarks Dataset provides such an accessible archive. The value of the dataset for historical researchers is illustrated by using it to investigate the identity of Japan's oldest registered trademark. Several candidates are considered—the Yanagiya cosmetics company's graphical mark for hair tonic, the Hyakumangoku Shuzo [Million Stone Brewing Company] "Jukai" label for nihonshu, and the Kokonoe Mirin Company's sakura label. The Japan Trademarks Dataset allows evaluation of each mark (and any other candidate) by registration number, registration date, application filing date, and other characteristics, each of which supports a different trademark's claim to the status of oldest Japanese trademark. This analysis demonstrates how comprehensive, accessible data can allow researchers to investigate and tell illuminating and compelling stories about the history of the trademark system in Japan, at the level of individual brands.

2. Internal Analysis of Japan's Trademark Registration System

The Japan Trademarks Database enables researchers to engage in analysis of Japan's trademark registration system in ways that are not possible with the aggregate statistical data that has heretofore been available.

(1) Identification of Key Actors in the Trademark Registration System

The Japan Trademarks Dataset allows for comprehensive analyses of various aspects of the trademark system. For example, we can use it to identify those attorneys who prosecute the largest number of applications for trademark registrations. The top 15 trademark attorneys identified in the dataset have each been associated with between 14,000 and 24000 applications over their lifetimes. Leading the list is 石川 義雄 (Ishikawa Yoshio), former adviser to the Japan Trademark Association

and former section manager of the Trademark Division of JPO, who passed away in 2015.[10]

(2) Check on JPO Statistical Reports

The NISTEP Research Paper by Motohashi et al. revealed that the INPIT data on registrations seems to have some discrepancy with the aggregate registration numbers reported by JPO and WIPO.[11] The discrepancies are substantial—as much as 10% of the annual registration totals. The Japan Trademarks Dataset allows us to perform a similar analysis for trademark applications, revealing a similar discrepancy.

These discrepancies do not necessarily reveal any error on the part of either the JPO or the Japan Trademarks Dataset—though they may. Or, they may simply illustrate the point that there are many ways to tabulate aggregated data, and different tabulation methods may be more or less relevant for particular analytical purposes. Instead of relying on the single aggregate measures made available by JPO, the Japan Trademarks Dataset allows researchers to generate their own aggregate measures to suit whatever research purpose they are pursuing.

(3) Analysis of JPO Performance

Comprehensive application-level data allows for other types of evaluation of JPO's performance of its trademark functions. This includes evaluations that are impossible using aggregate measures such as those disclosed in JPO's annual reports. Two examples are registration rates and individual examiner performance.

(i) Registration Rates

The proportion of applications that proceed to registration cannot be calculated using the aggregate data in JPO's annual reports, because the cohorts reported in annual application aggregates are different from those reported in annual registration aggregates—there is an average lag of several months between application and registration. Application-level data is therefore required. Using the application-level data in the Japan Trademarks Dataset, we find that registration rates hit a low of approximately 68% in 2000, and since 2002 have been decreasing from a high of approximately 80% to a low of approximately 75%.

---

10 【訃報】石川義雄先生ご逝去 |2015年 | お知らせ | 日本商標協会(JTA), 日本商標協会(JTA), http://jta.tokyo/information/detail/153/.
11 Motohashi et al., *supra* note 9, at 17-18 & tbls. 13-14.

(ii) Examiner Performance

The Japan Trademarks Dataset allows analysis of individual examiner performance along metrics such as pendency times and registration rates. Such analysis reveals substantial variation in the examiner corps. Pendency times for applications that proceed to registration show a standard deviation of approximately 75 days around the mean pendency time (for the entire dataset) of approximately 260 days. The distribution of registration rates is somewhat more compact around its central tendency—a standard deviation of 11.2% around a mean of approximately 73.5%. Interestingly, however, the registration rate distribution is also significantly skewed—indicating that while more examiners have above-average registration rates, those with below-average registration rates must be responsible for a larger number of applications.

(4) Correlation of Application Characteristics with Registration Outcomes

The Japan Trademarks Dataset enables researchers to evaluate the correlation between various application characteristics and registration outcomes. For example, we can analyze the difference in registration rates between applications with which an attorney is associated and those for which there is no attorney data. Such an analysis shows that attorney attorney representation appears to make a huge difference in the chances of an application proceeding to registration. Attorney success rates are between 90 and 100 percent, while unrepresented applicants have success rates of between 40 and 75 percent. This analysis also suggests that the severe drop in registration rates around 2000 was largely attributable to *pro se* applicants who may have been less likely to comply with legal requirements and more likely to abandon their applications if they ran into difficulty in examination. Similar correlations may be examined to investigate the relative success rates of individuals versus firms, of domestic versus foreign applicants, of applications in particular similarity groups or goods and services classes, and so on. This type of correlation analysis is likely to be of interest not only to researchers, but also to attorneys attempting to predict their clients' needs, and to policymakers attempting to allocate resources or implement reforms.

3. Comparative Analysis

An important goal of the Japan Trademarks Dataset is to allow for comparison of trademark registration data across countries—particularly advanced economies with heavily used trademark registries. Any of the analyses reviewed above, and many more, can be performed comparatively, demonstrating important differences between national systems, or important international dynamics

affecting the Japanese system. Below are two illustrations of the types of comparative analysis that the Japan Trademarks Dataset can be used for: comparison of use-based and registration-based trademark systems (using USPTO data for comparison), and comparison of domestic registration with international registration under the Madrid Protocol.

(1) Use-Based vs. Registration-Based Systems

One common criticism of registration-based systems is that they can generate "cluttered" registers: registrations may issue that ultimately never get used, misleading the public into thinking that certain marks are occupied by bona fide users when in fact they are not.[12] Such clutter can be a problem if it prevents or discourages those who would make bona fide use of a promising trademark from actually engaging in that use, without providing any offsetting benefit to the registrant or the public.

Comparative analysis using the Japan Trademarks Dataset and the USPTO's Trademark Case Files dataset shows that trademark applications in the US are *published for opposition* at roughly the same rate as applications are *registered* in Japan. However, because even published US applications that draw no opposition must still be supported by proof of actual use before the USPTO will issue a registration, many otherwise registrable US applications are ultimately abandoned because the applicant fails to actually use the mark for which they applied. The result is a roughly 25-percentage-point reduction in the number of applications that proceed to registration in the US. If the rate of failure to use is roughly constant between the US and Japan, we can estimate that approximately 25% of new trademark registrations in Japan are pure clutter—registrations for marks that will never be used in commerce. This clutter may be a serious issue, as it may linger on the register for many, many years.

Of course, in the meantime, registrations and renewals can generate significant fee income for the operator of the registry—in this case, JPO—and fees from those additional registrants can help keep fee levels lower for everyone. However, the additional fees come at a cost in the form of increased clutter in the registry—a cost that the Japan Trademarks Dataset allows us to begin to quantify using empirical analysis.

(2) Madrid Protocol Applications vs. Domestic Applications from Foreign Applicants

A common issue facing trademark owners who plan to market goods or services under the

---

[12] Georg von Graevenitz, *Trade Mark Cluttering–evidence from EU Enlargement*, 65 OXF. ECON. PAP. 721 (2013).

same mark in multiple countries is how best to secure rights in those countries where they expect they will want protection. An international registration under the Madrid Protocol offers cost savings, but also carries certain risks not present for domestic registrations. The Japan Trademarks Dataset allows for investigation of the extent to which foreign trademark holders seeking rights in Japan are opting for international as opposed to national registrations to secure those rights. A preliminary analysis along these lines demonstrates broad variation from country to country in the preference of those countries' domiciliaries for domestic versus international registrations—though some of this variation could be attributable to the absence or timing of accession to the Madrid Protocol.

4. Economic Analysis

Scholars often attempt to investigate the relationships between legal phenomena such as patenting and macroeconomic phenomena such as innovation or economic growth.[13] Or they attempt to analyze the relationship between legal events such as litigation or patent issuance and microeconomic indicia such as securities prices.[14] The Japan Trademarks Dataset provides fodder for both types of analyses. This section provides examples of each.

(1) Macroeconomic Correlations—GDP

Proponents of expansive intellectual property protection often suggest correlations between such protection and economic activity.[15] If such a correlation were present, we might still ask whether economic activity is a cause or an effect of increased intellectual property rights. But such a correlation cannot be assumed; it must be established empirically. Plotting quarterly aggregates of application data from the Japan Trademarks Dataset against quarterly real GDP data from the Economic and Social Research Institute of the Cabinet Office of the Government of Japan shows that trademark applications and GDP appear to be trending in roughly opposite directions—flying in the face of our most common assumptions about the relationship between IP and economic

---

[13] *See generally, e.g.*, JAMES BESSEN & MICHAEL JAMES MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK (2008); Knut Blind & Andre Jungmittag, *The Impact of Patents and Standards on Macroeconomic Growth: A Panel Approach Covering Four Countries and 12 Sectors*, 29 J. PRODUCTIVITY ANALYSIS 51 (2008).

[14] *See generally* Alan C. Marco, *The Value of Certainty in Intellectual Property Rights: Stock Market Reactions to Patent Litigation* (unpublished manuscript Nov. 15, 2005), available at http://ssrn.com/abstract=945009. As of this writing the author of this manuscript is Chief Economist of the USPTO.

[15] *See, e.g.*, World Intellectual Property Organization, Economic Development and Patents, http://www.wipo.int/patent-law/en /developments/economic.html ("The recent history seems to show that technology and knowledge are important factors for e conomic growth and development. Since the creation of the first mechanism to protect inventions in 15th century, the patent system has evolved with a view to promote innovation and encouraging economic development.").

growth. This kind of basic empirical testing of the assumptions underlying IP policies is crucially important to enlightened policymaking.

(2) Microeconomic Correlations—Securities Prices

Finally, the application-level data in the Japan Trademarks Dataset makes possible the analysis of correlations between individual firms' trademark activities and their financial performance. Such correlations have some theoretical basis. For example, if we think that trademark applications are evidence of new lines of business, we might expect that such events would be leading indicators of increased cash flows of their holders, and thus of increased market valuation. Abandonment (or non-renewal) of trademark registrations are theoretically more ambiguous. To the extent such events indicate closing lines of business, they could be lagging indicators of decreased revenues, but if the closing lines of business are unprofitable (either on an absolute basis or in comparison to a firm's continuing lines of business), they could instead be lagging—or even leading—indicators of improved profit margins. Proving any such relationship would take quite sophisticated econometric analysis, but the Japan Trademarks Dataset allows us to investigate whether the relationship is plausible. I illustrate such a plausibility check by comparing the trademark application and renewal activity of the Kao Corporation (the most prolific applicant in the Japan Trademarks Dataset) with that company's stock price. This comparison shows, perhaps unsurprisingly for a successful large company, that both data series are generally increasing as a function of time. More specific correlations, however, are difficult to identify.

V. Conclusion
[Omitted]

VI. Acknowledgments
[Omitted]