**U.S. Patent and Trademark Office**
**OFFICE OF THE CHIEF ECONOMIST**
Economic Working Paper Series

# Creative Destruction and Strategic Protection: Evidence from Pharmaceutical Patenting

Alan C. Marco
U.S. Patent and Trademark Office

Charles A. deGrazia
Royal Holloway, University of London
U.S. Patent and Trademark Office

**Original Version: September 2011**
**This Version: June 2017**

USPTO Economic Working Paper No. 2016-05

# Creative Destruction and Strategic Protection: Evidence from Pharmaceutical Patenting

Alan C. Marco
U.S. Patent and Trademark Office

Charles A. deGrazia
Royal Holloway, University of London
U.S. Patent and Trademark Office

**Original Version: September 2011**
**This Version: June 2017**

**Abstract**

This paper investigates the strategic use of follow-on patenting, or patent "fencing," by investigating patenting behavior in the pharmaceutical industry. Using a large sample of drug patents, we develop an empirical test for strategic substitution versus strategic complementarity at the product level in pharmaceuticals. We compare patenting by pioneers to the patenting behavior of those firms who cite the pioneers. Non-parametric hazard estimates show a marked difference between the timing of follow-on patents by competitors (creative destruction) and follow-on patenting by pioneers (fencing). We find that fencing tends to decrease the degree of citation by other firms, indicating that follow-on patenting is characterized by strategic substitution. Follow-on patenting may correspond to increased or extended market power in the original substance.

*Keywords: Patents, Pharmaceuticals, Cumulative Innovation, Strategic Substitutes, Strategic Complements*

*JEL Classification Numbers: K19, L1, L65, O31, O32*

## 1 Introduction

One of the primary contributions of the new theory of industrial organization was in the subfield of oligopoly pricing. A convenient structure within which to examine oligopoly pricing are the concepts of strategic substitution and strategic complementarity, coined by Bulow, Geanakoplos, and Klemperer (1985). Fudenberg and Tirole (1984) complement the analysis by investigating investment under strategic substitution and strategic complementarity. However, to date empirical tests of strategic substitution and strategic complementarity are rare.

At the same time, the innovation literature has increasingly focused on R&D, including strategic uses of R&D and patenting (e.g., Lerner 1995, Hall and Ziedonis 2001, Ziedonis 2004, Noel and Schankerman 2013). Reinganum (1989) provides an early survey of some theoretical results, but economists have yet to propose a global framework for modeling R&D. Much of the empirical work relies on reduced form estimations, and for the most part has not framed the analysis in the context of strategic substitutes/substitution (SS) and strategic complements/complementarity (SC). In fact, until Bloom, et al. (2013), there had been little evidence as to whether R&D exhibits SS or SC.

In this paper, we develop an empirical test for SS/SC in the context of pharmaceutical patenting. Our results show that follow-on patenting in pharmaceuticals is characterized by strategic substitution, or "mutual avoidance," for both incumbents and entrants in the product space. One of the benefits of using pharmaceutical patenting is that – more so than any other industry –patenting can be seen as a proxy for R&D. In addition, a focus on patenting enables delineation of R&D intensity by product line. That is, the results do not reflect whether pharmaceutical R&D is a SS or SC. Rather, we find that, at the product market level, patenting is a SS. This result is similar to the results of Lerner (1995) who examines patenting behavior of new biotechnology firms. He finds that firms with high litigation costs avoid "crowded" patent subclasses. In our results as well as Lerner's, this finding does not imply that the firms patent less in aggregate. Thus, aggregate R&D could still be characterized by SC, which we do not address in this paper.

Follow-on patenting in pharmaceuticals by pioneer drug firms is one way to "fence" a drug. Follow-on patents create a barrier to entry for a narrowly defined drug market, and also allow firms to lengthen the patent protection of a chemical entity. For instance, the extended release form of a drug may embody a separate patent, and indeed a separate New Drug Application (NDA) at the

Food and Drug Administration (FDA). Thus, while the patent on the new chemical entity (NCE) expires, the extended release form of the drug is still protected. Specifically related to "patent fencing", Sternitzke (2013) examines the use and timing of strategic pharmaceutical patent types (offensive versus defensive) in one product line, PDE5 inhibitors. Sternitzke (2013) finds that defensive patents are filed early in the patent's lifecycle while offensive patents, with "low potential to substitute prior filings economically", are filed in the later stages of the NCE patent's lifecycle. Ziedonis (2004) examines a similar fencing phenomenon in the context of defensive patenting in the semiconductor industry. She finds that firms patent more aggressively in technological areas where the ownership of patent rights is highly fragmented. She does not explicitly test for SS/SC.

In addition to "fencing", there is a vast literature relating strategic behavior and the role of intellectual property in the pharmaceutical industry. Gilchrist (2016), advancing the optimal patent term length literature of Gilbert and Shapiro (1990) and Gallini (1992), found that an additional year of an incumbent's market exclusivity increases subsequent drug entry by 0.2 drugs, where the first entrant's market exclusivity acts as an "implicit subsidy towards non-infringing substitutes." Another branch of the economics of intellectual property literature in the pharmaceutical industry relates strategic behavior and reaction to generic market penetration. Recent studies of generic market penetration, including a 2014 working paper by Branstetter, et al., find that generic penetration market decreases early-stage innovation within the same therapeutic class while simultaneously noting the stability of drug development across all therapeutic areas. Hemphill and Sampat (2012) examine the strategic behavior of potential entrants into pharmaceutical markets through patent challenges, finding that firms use patent challenges to target high-sales drugs. These challenges mostly target "lower quality and later expiring" patents and aim to limit an incumbent's ability to "evergreen" branded drugs.

Our paper builds upon the few empirical studies SS/SC in the industrial organization literature discussing SS/SC. Most recently, Bloom et al. (2013) investigates strategic interaction between product market rivals in R&D investment. The authors utilize the Mahalanobis extension to the Jaffe (1986) measure of spillovers to compute a measure of product market rivalry at the firm level. Bloom, et al. (2015) found evidence of strategic complementarities in own and product market rivals' R&D but the result was not robust when R&D is considered endogenous. Cockburn and Henderson (1994) present one of the first tests of SS/SC in R&D. They gather project level

data from ten pharmaceutical manufacturers, using R&D expenditures as a measure of inputs; and patents, new drugs, and drug sales as measures of output. They find that R&D is characterized by SS at the firm level. However, no robust relationship was found at the project level.

In a working paper, Dewo, Gans, and Hirschberg (2005) develop competing theoretical bases for SS and SC, where the complementarity can arise in a patent race model. Using Compustat data on 31 incumbent firms and Venture Economics data on startups, they predict SS between incumbent and entrant firms. They find mixed results, so that R&D can be characterized as either SS or SC. For pharmaceutical firms in their sample, R&D for incumbents appears to be either SS, or non-responsive to entrant R&D. Note that the firm-level result differs from that of Grabowski and Baxter (1973), who find that R&D is characterized by SC in pharmaceuticals.

Sundaram, John, and John (1996) perform one of the most direct examinations of SS/SC for R&D. To determine SS/SC, it is sufficient to determine the sign in the change in marginal profitability of R&D spending, with respect to changes in competitors' R&D spending. Their approach is to examine the change in stock prices as a reaction to announcements of new R&D spending by competitors. They hypothesize that in the aggregate – across industries – it is not unexpected to get an insignificant effect. If some industries/firms are characterized by SS (so that the sign is negative), and some are characterized by SC (positive sign), then the average effect could easily be zero. Thus, it is not surprising if other studies find very weak evidence for SS/SC. In a sample of 125 announcements by 65 firms, the authors distinguish SS firms from SC firms by determining the direction of change in stock price as a reaction to R&D announcements by competitors. They subsequently classify SS firms and SC firms on the basis of this sign. Their analysis determines the factors that influence the magnitude of the reaction to R&D announcements.

Finally, the strategic patenting and SS/SC literature is related also to models of entry in industrial organization, specifically strategic deterrence models or models with initial investment (fixed cost) requirements. For example, Ellison and Ellison (2011) develop and test a model of entry deterrence to determine if firms act strategically as patents near expiration by altering advertising, product, and pricing decisions. They find that entry deterrence depends nonmonotonically on market size, where only incumbents in medium-sized pharmaceutical markets exhibit entry deterrence behavior.

In this paper, we examine the patenting behavior of pharmaceutical firms based on data contained in the FDA's Orange Book, as well as patent citation and patent assignee data from PatentsView. We examine the data at the product level, using the patent on a new chemical entity (NCE) as the basis for defining the product. Patents that cite the original patent represent research intensity by both the pioneer firm and competitors. We estimate the dynamics of patent citations using a hazard rate model. One advantage of this method is that we do not depend on annual or quarterly observations, but can observe patenting in continuous time. Relative to more aggregated data, continuous data make the timing of the "reactions" in the reaction functions easier to identify.

In the following section, we describe some of the relevant institutional details specific to patents in the pharmaceutical industry. Section 3 presents some of the theoretical considerations of strategic substitution and strategic complementarity with R&D. Section 4 describes the empirical specification used for the hazard model, followed by a description of the data in section 5. The results are presented in section 6, and section 7 concludes.

## 2 Pharmaceutical Patenting

In 1984, the U.S. Congress approved amendments to the federal Food, Drug and Cosmetic Act (The Drug Price Competition and Patent Term Restoration Act, commonly referred to as the Hatch-Waxman Act). A primary goal of Hatch-Waxman was to reduce the delay in the approval of generic drugs by way of Abbreviated New Drug Applications (ANDAs). The ANDAs allow generic manufacturers to essentially eliminate clinical trials by utilizing the submitted clinical trials for the pioneer drug without themselves conducting independent trials on the generic drug. Thus, the generic manufacturers do not incur the same fixed costs for drug discovery and development as the pioneering firm, and can bring the generic drug to market at a much lower cost.[1]

The success of Hatch-Waxman in spurring generic entry was met by a strategic response by pioneer firms, in the form of follow-on patenting. Until the passage of the Medicare Prescription Drug Improvement and Modernization Act of 2003 (Medicare Act), incumbent pioneers could enjoy a 30 month stay on generic approval by threatening patent litigation. This stay is possible only if the pioneer firm has unexpired patents listed with the Food and Drug Administration (FDA)

[1] In order to obtain FDA approval for the generic, the generic manufacturer must demonstrate that its drug is bioequivalent to the FDA-approved pioneer drug.

in its Orange Book. Incumbent firms have an incentive to "pad" their list of Orange Book patents covering the drug in question. As long as some patents are in force, the threat of litigation (and the automatic 30-month stay) is credible. Additionally, incumbent firms may have an incentive to delay patent issuance for follow-on patents, so that the threat of litigation is extended. The Medicare Act limited incumbent firms to a single 30-month stay.

Because of strategic patenting behavior, the last decade has seen a great deal of litigation activity in pharmaceuticals. First, many lawsuits are filed in order to obtain the stay. Few of these cases actually go to judgement; however, the settlement process itself has the potential to allow incumbents and the first generic entrant to collude (Hovenkamp, Janis, and Lemley 2003). Further, incumbent firms have come under fire from class action lawsuits alleging anti-competitive patenting behavior. For example, AstraZeneca (Prilosec) and SmithKline Beecham (Paxil) were each alleged to have strategically delayed patents so as to extend their exclusivity. Each firm had initiated more than ten patent infringement suits.

At the same time there are opportunities for "entry" within a particular drug line. For instance, it is possible for a competitor to patent an isomer of the NCE that is the basis of a drug. This isomer would infringe the patent on the NCE, but would represent a different marketable form of the drug. Other follow-on patents based on the original drug are described below. Patent holders can either invest in these follow-ons, or allow competitors to patent. The question remains whether a patent holder is more likely to "fence" in its property in the face of increased competition, or when the competition is limited. Similarly, will competitors tend to attack a well-fenced product, or will they tend to attack the unfenced property? These questions are addressed more formally in section 3.

In this paper, we use a large sample of drug patents to investigate patent citation behavior both before and after the expiration of the initial patent on the chemical composition. Using both non-parametric and parametric methods, we find that follow-on patenting tends to delay patent citations by other firms. Citations by other firms tend to be on patents with fewer follow-on patents, but which are still under patent protection.

### 2.1 Follow-on Patenting

For some years, patent researchers have been exploiting the growing availability of patent statistics to study, among other things, patent value (Pakes 1986, Schankerman and Pakes 1986,

Schankerman 1998), incentives for R&D (Kortum and Lerner 1999, Sakakibara and Branstetter 2001, Hall and Ziedonis 2001), strategic behavior (Grindley and Teece 1997, Hall and Ziedonis 2001, Noel and Schankerman 2013, Sternitzke 2013, Bloom, et al. 2013), and consolidation (Marco and Rausser 2002). Pharmaceutical papers have paid great attention to the protection afforded by patents. Pharmaceutical products are undoubtedly fertile markets in which to investigate the value of patenting. In comparison to markets like electronics, patented pharmaceutical products are well-defined, and generic entry and therapeutic substitutes are relatively easy to quantify. Nonetheless, attention has been paid to the observable characteristics of the patents themselves only recently.[2]

This paper investigates the strategic use of follow-on patenting to effectively extend the patent protection of the original drug patent. We define the original drug patent as the first patent listed by the incumbent firm for the brand name drug in the FDA's Orange Book. A follow-on patent, or self-citation, is defined as any patent in the same Orange Book New Drug Application (NDA) as the original NCE patent or a patent which cites the original NCE patent that is also assigned to the incumbent. Most self-citations in our sample are not contained within the original drug's NDA but are linked to the original drug patent through forward citation. All other citations in our sample represent patents assigned to other firms that cite the original drug patent.

Follow-on patents are generally process or product patents that relate to the original chemical composition of the pioneer drug. Reference to an example is helpful in illustrating the ways in which incumbents (or other firms) may cite the original patent. For example, the initial patent for Zantac is patent number 4,128,658. The patent refers to a chemical composition as described in figure 1. Any reference to the original patent may be a substitute, derivative, form, process, or use. A *substitute* chemical is one that accomplishes the same therapeutic purpose through an entirely different chemical structure. For example, patent number 4,239,908 has the chemical structure shown in figure 2. A *derivative* is a chemical that belongs to the same class of compounds as the pioneer chemical and treats the same ailment. Many of the functional groups

[2] Hemphill and Sampat (2012) differentiate types of pharmaceutical patents between those that cover an active ingredient (AI) and those that do not. Non-AI patents tend to be of "lower quality" and are more likely to be declared invalid by federal courts. Sternitzke (2013) uses defensive and offensive blocking definitions from Blind, et al. (2006), where defensive blocking patents prevent competing firms from reducing a firm's "own technological room to maneuver," by patenting the technological area around the original invention; offensive block patents prevent competing firms from utilizing or extending their inventions by patenting in the same technological area as the original invention.

are the same, and the mechanism of action is usually the same. The derivative patent and the pioneer patent have a common origin from which they were both derived. In many cases, the pioneer chemical was the first of its kind, so all subsequent derivatives derive from this chemical. For instance, patent number 4,128,658 represents a derivative of the original molecule for Zantac.

Figure 1: Chemical Composition for Zantac

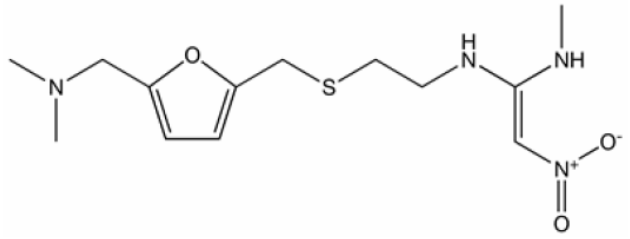



Figure 2: Substitute

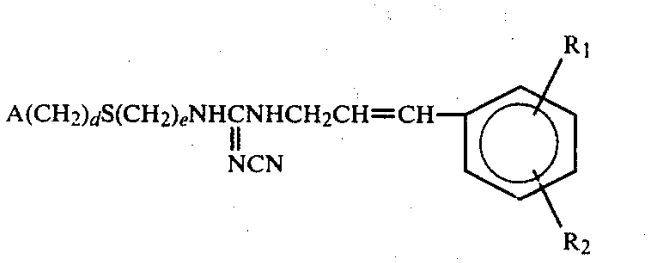
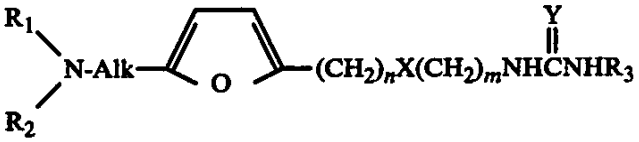



Figure 3: Derivative




A *form* is a patent for a different physical form of the pioneer chemical. This could include, for example, a more soluble form of the drug so that it could be injected into the blood, or the drug incorporated into a stable pill form. A form is the same chemical, but in a different environment so that it functions more effectively or more conveniently. Specifically, patent number 4,880,636

"relates to an improved polymeric film coating for a ranitidine Hydrocholride (HCl) [Zantac] tablet in which the plasticizer triacetin has been added to the polymeric film coating medium."

A *process* patent is a novel way of producing the pioneer chemical. Many process patents are for a more cost effective way of making the intermediates in the production of the pioneer chemical, rather than the pioneer chemical itself. Patent number 5,621,120 is "a process for the manufacture of Form 1 ranitidine hydrochloride (N-[2-[[[5(Dimethylamino)methyl]-2-furanyl]methyl]thio]ethyl-N'-methyl-2- nitro-1,1-ethenediamine) hydrochloride."

A *use* patent is one that describes the use of the pioneer chemical to treat an ailment that is substantially different from the ailment described in the pioneer chemical's patent. For example, patent 5,364,616 "relates to methods for prevention or treatment of gingivitis or periodontitis comprising topical administration, to gingival tissues of the oral cavity, of a composition comprising a safe and effective amount of a selective histamine-2 receptor antagonist compound." The original compound (Zantac) treats gastric problems of the stomach and esophagus, while the '616 patent describes a treatment for oral diseases.

Unrelated patents are ones that do not fit into the other categories. Many are truly unrelated to the pioneer patents, but some are related indirectly. An unrelated patent may be, e.g., a process for preparing a competing product. Or, it may be a drug that treats side effects of the pioneer drug. For instance, patent 6,780,880 is a description of a process that "determines the isomer ratio of chemical compositions." It does not mention Zantac in any way.

The pharmaceutical industry provides a useful case study for investigating the value of patent citations. Many empirical patent papers presume that more highly cited patents are more valuable. While this presumption is in all likelihood true, it is important to recognize that some citation activity can be "negative." In the context of pharmaceuticals, citations by the pioneer may correspond to the degree a pioneer is able to "fence" its original substance with follow-on patenting. However, citations by other firms can be either positive or negative: they may signal the degree of inter-molecule substitution, indicating a potential loss to patent holders. Alternatively, they may indicate cumulative innovation by other firms, which may entitle the incumbent firm to some rent.

Equally important as the quality of citations, is the timing of citations. A claim in antitrust litigation is that incumbent's strategically time their follow-on patents in order to extend the period of exclusive marketing and manufacture. Table XX shows an example of patent citations for

Ceftin, manufactured by Glaxo. One can observe that early in the patent's life, Glaxo produces many patents that cite the pioneer patent. These patents generally cover new forms of the drug. Late in the patent's life – after expiration – there is a flurry of patenting activity by the generic manufacturer Ranbaxy. Midway through the patent's life, there is one patent by Sumitomo on a compound that competes directly with Ceftin.

## 3 Theoretical Considerations

Suppose that a patent holder (A) owning the exclusive production rights to a particular chemical may extend that patent protection through investing in follow-on patenting, $x_A$: The follow-on patenting extends the market power in that therapeutic area, so we assume that profit $\Pi_A$ is increasing and concave in $x_A$: At the same time, competitors may also invest in follow-on patenting, $x_B$. Because $x_B$ gives competitors a foothold in the product line of A, we assume that $\Pi_A$ is decreasing and convex in $x_B$: Additionally, we assume that $\Pi_A$ is a function of overall market size $\theta$ (therapeutic class) and the number of therapeutically equivalent substitutes y: Thus,

$$\Pi_A = \Pi_A(\theta, x_A, x_B, y)$$

With derivatives:

$$\frac{\partial \Pi_A}{\partial x_A} > 0\,, \frac{\partial \Pi_A}{\partial x_B} < 0, \frac{\partial \Pi_A}{\partial y_B} < 0, \frac{\partial \Pi_A}{\partial \theta} < 0$$

$$\frac{\partial^2 \Pi_A}{\partial x_A^2} < 0, \frac{\partial^2 \Pi_A}{\partial x_B^2} > 0$$

where θ is market size (therapeutic class).

In the context of the data, $x_B$ represents follow-on patenting of the type described above, except that "substitutes" are considered to be counted in $y_B$, because these represent NCEs that compete with the pioneer, but that do not infringe.

To determine the strategic response of competitors means to look at marginal profitability $\partial \Pi_A/\partial x_A$ with respect to changes in $x_B$; $y_B$ and vice versa. That is, we want to consider the sign of $\partial^2 \Pi_A/(\partial x_A \partial x_B)$. Equivalently, we are interested in the sign of $\partial x^*_A/\partial x_B$, the sign of the reaction

function. If this sign is negative, then we say that $x_A$ and $x_B$ are strategic substitutes. If this sign is positive, then we say that $x_A$ and $x_B$ are strategic compliments.

If follow-on patenting is characterized by mutual avoidance, so that firms tend to invest where competitors do not, then one would expect to see a negative sign at the product level, indicating SS. On the other hand, it could be that incumbent firms will aggressively defend their property, indicating SC. Finally, it could be that aggressive follow-on patenting by the incumbent is met with an avoidance strategy by competitors. This last scenario would lead to SC for the incumbents, and SS for the competitors.

## 4 Econometric Specification

An appropriate estimation strategy to investigate the timing of follow-on patenting and citations is duration, or hazard, estimation with occurrence dependence. If follow-on patenting decreases the willingness or ability of other firms to exploit the technology, then an impact on the rate of citations by other firms should be observed. Hazard estimation considers several dependent variables as a function of the current citation counts and other control variables. For follow-on patenting, the incumbent firm will choose to issue another patent fencepost in the next small interval of time when the value of doing so exceeds the reservation value (the status quo). Of course, the value of a fencepost for any particular drug is dependent upon the current level of exploitation by competitors, as well as market conditions like available substitutes and market size.

The hazard function, $\lambda(t)$, gives the likelihood that the incumbent firm will build a fencepost with a follow-on patent (self-citation), given that it has not built a fencepost for t periods. The hazard function is defined as $\lambda(t) = f(t)/(1-F(t))$, where $f(t)$ and $F(t)$ are the usual density and cumulative distribution functions. For simplicity, we estimate our baseline results using standard semi-parametric (Cox) and parametric (Weibull) survival models. Using the Weibull model, we then introduce additional covariates and specifications, including frailty models to account for unobserved heterogeneity. The Weibull distribution employed in this paper implies a hazard function of the form $\lambda(t) = \gamma\rho(t)^{\rho-1}$. This hazard function includes the exponential as a special case where $\rho = 1$. However, for values of $\rho < 1$; the hazard function will exhibit negative duration dependence (the spell duration decreases the probability of a follow-on patent in the next interval of time). For $\rho > 1$, the hazard function will exhibit positive duration

dependence, so that the length of a spell increases the likelihood of a follow-on patent. For both the exponential and Weibull models, the parameter $\gamma$ is modeled as $\kappa(x)$:

$$\kappa(x) = e^{X\beta + \varepsilon} \qquad ()$$

where $X$ is a matrix of observable characteristics (given in table XX). To control for additional covariates, we utilize proportional hazard models, where the respective hazard function can be decomposed into $\lambda(t|x) = \kappa(x)\lambda_0(t)$. Using the Weibull distribution, the baseline hazard then becomes $\lambda_0(t) = \rho(t)^{\rho-1}$.

Estimation involves maximum likelihood estimation where the censored observations are incorporated (Greene 1993 and Wooldridge 2010), viz.:

$$lnL = \sum_{uncensored} \lambda(t|\, x; \beta, \rho) + \sum_{all} ln(1 - F(t|x; \beta, \rho)) \qquad (3)$$

Estimating the equation for the hazard of citation or self-citation is similar to equation (2) above.

Additionally, we test for unobserved heterogeneity by estimating a frailty model, which enters the hazard function multiplicatively. The hazard function for observation $j$ for patent $n$ is specified as

$$\lambda\left(t_{nj}\middle|X_{nj}, \nu_n\right) = \nu_n \lambda\left(t_{nj}\middle|X_{nj}\right) \qquad (4)$$

where $\alpha_n$ (the "frailty") follows a gamma distribution with mean one and variance $\eta$ (the degree of heterogeneity). $\eta = 0$ implies no unobserved heterogeneity – the standard Weibull hazard model. The null hypothesis of no unobserved heterogeneity ($\eta = 0$) can be tested using a Likelihood Ratio test. The frailty model above is essentially a random effects model for hazard estimation. It should be noted that if unobserved heterogeneity is present but not modeled – making some more prone to citation than others based on unobservable characteristics – then the duration dependence parameter ($\rho$) will be asymptotically underestimated (Lancaster 1990, Wooldridge 2010). Therefore, one can observe a decreasing duration dependence in the population, even when individual-level duration dependence is *increasing*.[3]

[3] In mortality studies, some subjects may be more "frail" than others. Frail patients die early, leaving a more robust population alive. If the frailty is unobservable *ex ante*, then there will be an apparent decrease in mortality over analysis time. It is a type of

Estimation of Equation 3 proceeds via maximum likelihood, with censored observations incorporated much like the Tobit model (Greene 1993 and Wooldridge 2010). The log-likelihood function is

$$lnL = \sum_{uncensored} \lambda(t|\theta) + \sum_{all} \ln(1 - F(t|\theta)) \quad (4)$$

Where $\theta = (\beta,\rho,\eta)$.

## 5 Data

### 5.1 Drug Sample

The full drug sample is obtained from historical Orange Book data published by the Food and Drug Administration (FDA). The data describe all patents registered with the FDA that cover granted new drug applications[4] (NDAs) from 1983 to 2016, including patents issued during the same period.

Drugs may be covered by more than one NDA, and NDAs may contain more than one patent. In the historical Orange Book data, the NDAs do not delineate which other NDAs are related to the same pioneer drug. For instance the original NDA for Prozac (capsule; oral) is not linked in any meaningful way with the NDA for Prozac Weekly (delayed release). Prozac Weekly contains follow-on patents. We account for this limitation by assuming that any new NDA for a drug will contain the original patent for the chemical composition. The patent on the molecule is generally the first listed (oldest) patent for the drug. Thus, we join NDAs in drug families based on the first listed patent. Spot checking reveals that the assumption is robust, but that there are exceptions. The 1,528 NDAs in the data comprise 1,120 unique first patents.

Patent expiration and NDA approval data were also taken from the historical Orange Book. In cases where two patent expiration dates are listed (e.g., with pediatric exclusivity), the later date was used. The NDA approval date determines when a drug is first granted marketing approval by the FDA. If there are multiple NDAs associated with a particular patent, the first approved NDA is used. In the historical Orange Book, therapeutic class data is sparse. Therefore,

fallacy of composition. In the patent context, highly cited patents drop out early and the clock on analysis time (duration) is reset. As the duration of a spell increases, weakly cited patents remain, leading to an apparently lower citation rate.

[4] The historical FDA Orange Book includes data for granted NDAs but not for applications that were subsequently rejected or abandoned.

to define therapeutic area in broad terms, we use the FDA Center for Drug Evaluation and Research (CDER) office to which an NDA was assigned as a therapeutic class alternative.[5] The CDER offices and divisions represent broad but help to define the general market for each drug. For example, Lodine (NDA 18,922), a nonsteroidal anti-inflammatory drug with active ingredient etodolac, was assigned to the Division of Anesthesia, Analgesia, and Addiction Products. Please see Table XX in the appendix for a comprehensive list of CDER offices and divisions.

**5.2 Patent Data**

Patent citation data, patent grant dates, and assignee data were obtained from PatentsView.[6] Patent claims data was obtained from the USPTO's Patent Claims Research Dataset.[7] So-called "forward citations" are citations received by the patent from subsequent patents. These are called simply "citations" in this paper. Self-citations are forward citations made by patents assigned to the owner of the pioneer patent or patents contained within the same NDA as the NCE patent.[8]

Forward citations are commonly associated with higher patent value. The rationale is that if a patent is frequently cited, then it may be the basis for cumulative innovation, and therefore technologically important. However, higher forward citations may also have a negative impact on value if citations reflect replacement by new technologies – Schumpeter's *creative destruction*. In the pharmaceutical context, it is likely that both effects occur. More citations will occur for blockbuster drugs, and they will also occur in crowded therapeutic classes. The question is whether behavior on the part of the patent holder can alter the degree to which competitors will pursue R&D in that technology area. For instance, Lerner (1994) finds that small inventors will avoid crowded technology areas.

In our empirical analysis, we do not disaggregate different types of forward citations by other firms. For a sample this large, coding individual patents (process, form, derivative, etc.) is infeasible. We simplify our analysis by treating each subsequent patent as an identical inventive step and leave analysis by patent type to future work. All citations in the empirical analysis are

[5] CDER office data provided by the FDA.
[6] Bulk data and data descriptions can be found at www.patentsview.org.
[7] Available at https://www.uspto.gov/learning-and-resources/ip-policy/economic-research/research-datasets.
[8] We define an assignee to be any original assignee listed on the face of the patent. If multiple assignees are listed, we do not differentiate between these assignees when determining a patent's "self-citation" designation.

for the first listed patent only. The first listed patent represents the patent on the compound, which is arguably the groundbreaking technology for the drug. For each original patent, we track the dates of both self-citations and citations by other firms and calculate cumulative totals. In all specifications, we normalize the citation counts by the number of citations *made* (backwards citations + 1, to avoid division by zero) by the original patent. This is an additional attempt to control for class-specific differences in citation rates. Finally, the percentage of NDA patents in force is calculated using the patent expiration dates from the historical Orange Book.

### 5.3 Summaries

The final sample includes 1,528 identified drugs (NDAs), and nearly 31,000 unique patents (either follow-on patents or citations). Subsequent patents may cite several of the "innovative" patents in the sample, so the final sample is over 51,000 observations. Citations made and CDER division are not available for every original patent in the sample, so there are some data losses in the estimations.

Table XX gives descriptions of the variables and summary statistics. The interquartile range is used because the distributions are skewed. Figures XX and XX show histograms for the number of Orange Book patents and the number of citations per drug in the sample. Note that each drug must have at least one Orange Book patent (the original patent) and that anything in excess of one is defined as a self-citation based on its date of issuance.

## 6 Results

### 6.1 Non-Parametric Estimation

The estimation consists of non-parametric Kaplan-Meier hazard estimation, semi-parametric Cox proportional hazards estimation, and parametric estimation using the Weibull distribution. Figures XX to XX show the non-parametric estimates of the hazard rate for various types of patenting. The hazard of self-citations is increasing for the first five years of a patent's life, and then decreases gradually throughout its life. In this case, we further separate self-citations into two categories: (1) those self-citations that are linked to the pioneer drug patent for a drug through the FDA Orange Book; and (2) those self-citations that cited the pioneer drug patent and are assigned to the same patent assignee. In (1), the hazard of the Orange Book self-citations (Figure XX) peaks at three years after the pioneer drug patent was granted and decreases

thereafter, implying that the likelihood of self-citation through the FDA Orange Book is decreasing but non-negligible three years after grant. Figure XX depicts the hazard of all self-citations owned by the same assignee but not included in the FDA Orange Book. The hazard of these self-citations follows a similar pattern to (1) but peaks at five to six years after grant and decreases thereafter.

The hazard of citation increases until thirteen years post-grant and remains steady until nearly thirty years after allowance, at which time the hazard rate spikes. Citations can be separated into two subgroups, inside and outside citations, where inside citations include citations within the same primary U.S. Patent Classification (USPC). The separation of inside from outside citations shows that the two types of forward citations appear to have different hazard rates. The hazard for outside citation increases for most of the patent's life, declining only at the very end, while the hazard of inside citation increases until thirteen years after allowance, after which its hazard rate decreases slowly. After 23 years post-grant, the inside citation hazard rate begin to drop substantially. These non-parametric estimates show a marked difference in patenting behavior by not only firm type but also citation type.

### 6.2 Semi-parametric and Parametric Estimation

In order to determine whether follow-on patenting has an effect on later innovation by others, it is necessary to turn to semi-parametric and parametric models to estimate the effects of covariates on the hazard function of each citation type. First, we estimate baseline Cox and Weibull duration models using aggregated self-citation and citation counts to determine whether the pharmaceutical industry exhibits strategic substitutability or complementarity at the product level. Based on our model, the semi-parametric and parametric models allow for a natural test of strategic complementarity or substitution in follow-on patenting by the original assignee or competitors. In the case of self-citations, a coefficient for the stock of citations by competitors, presented as hazard ratios, which is statistically greater than one implies strategic complementarity. In other words, the likelihood of an additional self-citation within the time interval (t,t+dt) increases with an additional competitor citation at time t. The owner of the original drug patent for a particular product line is therefore more likely to seek another patent on its product with a marginal increase in the stock of citations. If the hazard ratio is significantly less than one, follow-on patenting exhibits strategic substitution in the pharmaceutical industry.

A similar interpretation holds for the likelihood of citations. Secondly, we control for the unobserved heterogeneity across product lines by running Weibull frailty models. Finally, using both semi-parametric and parametric models, we perform robustness checks on our results.

### 6.2.1 Strategic Substitution versus Strategic Complementarity

Tables XX to XX provide Cox and Weibull hazard estimates for aggregated self-citations and citations, where each regression was run using the entire sample of pioneer drug patents and their citations. Coefficients are expressed as hazard ratios, so that a value above one indicates a positive impact on the hazard rate. In each table, column one is the Cox estimate, column two is the Weibull estimate (equation (2)) and columns three estimates a frailty model (equation (4)). In the baseline self-citation and citation frailty models, we perform a likelihood ratio test (null hypothesis: $\eta = 0$) and find evidence of unobserved product-level heterogeneity at one percent significance.

Several results from the baseline regressions merit emphasis. First, each type of patenting exhibits "positive incidence dependence:" the current count of all self-citations is associated with a higher hazard for self-citations. The same is true for citations. Second, and more importantly, additional citations by competing firms decrease the self-citation rate significantly across all regressions reported in Table XX. Therefore, the likelihood of an additional self-citation by the pioneer pharmaceutical firm decreases by one percent when a competitor receives an additional patent in the same technology space. This result indicates the existence of strategic substitution in pharmaceutical patenting at the product level. In other words, the pioneer firm is less likely to seek an additional patent in the same product line when faced with an additional competitor citation. A similar relationship holds true for the hazard rate of competitor citations (see Table XX). The likelihood of an additional citation by a non-pioneer firm in a given product line decreases by one percent when the pioneer firm receives an additional patent in the same technology space (or more specifically, the same product line) as the pioneer drug patent. Therefore, competing firms also exhibit strategic substation in patenting.

### 6.2.2 Robustness Checks

In this section, we propose a number of robustness checks for our main results and present our findings. First, pioneer and competing firms may alter their patenting behavior once a

pioneer drug has been successfully approved by the FDA. To account for this potential change in behavior, we estimate equations (2) and (4) where risk exposure is limited to both (1) the time period between patent grant and the first NDA approval for a given drug (Tables XX and XX) and (2) the date of the first NDA approval for a given drug and the last date of observance, July 16, 2016 (Tables XX and XX). This adjustment allows for a clean look at strategic behavior in a given product space before and after the success of the pioneer drug is known. In general, the results for self-citations are consistent with the baseline results when run on the pre- and post-NDA subsets of the full sample. However, we note some caveats: (1) for the citations frailty model where risk exposure is limited to the time period between first marketing approval and last observed date, the coefficient on the stock of self-citations is barely insignificant, but the Cox and standard Weibull results (columns one and two of Table XX) are consistent with baseline results; (2) for the self-citations frailty model, where risk exposure in limited to the date of pioneer patent grant and date of first NDA approval, the coefficient on the stock of competitor citations is insignificant (column three of Table XX) but, again, the Cox and standard Weibull results are consistent with the baseline results (columns one and two of Table XX).

Secondly, market and technology characteristics, including market size, number of competitors, and saturation of the product space, may influence a firm's strategic patenting behavior. Due to data limitations, we cannot directly control for market- or firm-level data, nor can we directly control for the saturation of technology space over time, which is unobservable. Therefore, in order to control for market characteristics, technology space saturation, and other time-varying unobservables, we introduce year and therapeutic class interaction terms to the standard Weibull regressions (Table XX). We find that controlling for market, firm, and technology space unobservables does not impact the direction nor significance of the variables of interest, strengthening our main result.

### 6.2.2 Regressions by Therapeutic Class

In this section, we investigate whether or not firms exhibit SS across all therapeutic areas or if there exists heterogeneity within the pharmaceutical industry in regards to strategic behavior. One could argue that firms operating in different therapeutic classes may react dissimilarly due to variation in the underlying characteristics of each therapeutic class. To test this theory, we run the baseline Weibull model for each therapeutic class from 1983-2016. The

results are presented in Table XX. We find that there exists significant heterogeneity in the direction and significance for both the self-citation and citation regressions across therapeutic areas. For example, whereas the overall effect across all therapeutic classes of an additional competitor citation on the likelihood of an additional patent received by the pioneer firm is negative, coefficients on the stock of competitor citations by therapeutic class are sometimes positive, negative, or insignificant. Therefore, response heterogeneity exists across therapeutic class but we do not investigate which specific underlying characteristics determine the direction of the response, which we leave to future work.

### 6.2.4 Limitations

There are a number of limitations to our analysis. First, we do not directly account for changes in patent assignment after a patent is granted by the USPTO. PatentsView links patent assignee(s) at grant – of which there can be more than one – to patent number but does not include data on changes to assignment over time. In the pharmaceutical patent space, a firm might acquire either an individual patent or the intellectual property portfolio of a competing firm. If an incumbent's (competitor's) patent is sold to a competitor (the incumbent), then both the citations and self-citation counts will not reflect the true patent landscape. To gauge the magnitude of this issue, we matched the pharmaceutical patents in our dataset to the USPTO's Patent Assignment dataset, which includes recorded patent (re)assignments from 1970-2016. These assignments, or transfers from one individual or entity to another, are voluntary and therefore not all transfers between patent holders are recorded. Of the XX,XXX patents in our dataset, over thirteen thousand were reassigned to another individual or entity, but many of these reassignments occur within the same corporate structure (for example, patent 9,511,066 was reassigned from Purdue Pharma to Purdue Pharma L.P.). The overall percentage of patents that were not reassigned to an entity within the same corporate structure has not yet been calculated. Future work in this area should take this discrepancy into account and modify citation and self-citation citations to incorporate dynamic patent ownership changes.

## 7 Conclusion

This paper investigates the citation behavior of firms in the pharmaceutical industry. In particular, we develop a large sample of pioneer patents on new chemical entities. Follow-on patenting indicates research interests by both the incumbent firm and competitors. Hazard modeling provides an empirical test of strategic complementarity versus strategic substitution at the product level. The primary result is that firms, including both pioneer drug firms and competitors, exhibit strategic substitution in pharmaceutical patenting at the product level. This result cannot be aggregated to the firm level, because firms may practice "mutual avoidance" by finding niches within the larger pharmaceutical market.

Extensions to the analysis include exploitation of a smaller sample of pioneer patents, where individual citations to the pioneers can be classified into categories, such as process patents, product patents, and substitute chemical entities. Further, better information on therapeutic classes would enable some control for the level of inter-chemical competition as opposed to intra-chemical competition.

The development of a patent thicket or patent fence is a sunk cost. As such, it would be interesting to compare these results to the theoretical results of endogenous sunk costs (Sutton). Additionally, follow-on patents by competitors and substitute chemicals constitute what Schumpeter termed *creative destruction*. On the other had incumbent firms can invest in strategic protection through fencing, that reduces the rate of creative destruction and increases the life of the pioneer drugs. Whether this is, in fact, a net welfare loss depends on the optimal life of patents in pharmaceuticals.

References

Bloom, Nicholas, Mark Schankerman, and John Van Reenan. 2013. Identifying Technology Spillovers and Product Market Rivalry. *Econometrica* 81 (July): 1347-1393.

Bulow, J. I., J. D. Geanakoplos, and P. D. Klemperer. 1985. Multimarket Oligopoly: Strategic Substitutes and Complements. *Journal of Political Economy* 93 (June): 488-511.

Branstetter, L., C. Chatterjee, and M. Higgins. 2014. Killing the Golden Goose or Just Chasing It Around the Farmyard?: Generic Entry and the Incentives for Early-Stage Pharmaceutical Innovation. NBER Working Paper.

Cockburn, I., and R. Henderson. 1994. Racing to Invest? The Dynamics of Competition in Ethical Drug Discovery. *Journal of Economics and Management Strategy* 3 (September): 481-519.

Dewo, S. A., J. S. Gans, and J. Hirschberg. 2005. Has Investment in Start-Up Firms Driven Incumbent Innovative Strategy? Evidence From Semiconductor and Biotechnology Venture Capital Funded Firms. *Intellectual Property Research Institute of Australia Working Paper Number* 16/05.

Ellison, Glenn and Sara Fisher Ellison. 2011. Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbents Prior to Patent Expiration. *American Economic Journal: Microeconomics* 3 (February): 1-36.

Fudenberg, D., and J. Tirole. 1984. The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. *American Economic Review - Papers and Proceedings* 74 (May): 361-366.

Gallini, Nancy. 1992. Patent Policy and Costly Imitation. *RAND Journal of Economics, The RAND Corporation* 23 (Spring): 52-63.

Gilbert, Richard, and Carl Shapiro. 1990. Optimal Patent Length and Breadth. *RAND Journal of Economics, The RAND Corporation* 21 (Spring): 106-112.

Gilchrist, Duncan. 2016. Patents as a Spur to Subsequent Innovation? Evidence from Pharmaceuticals. *American Economic Journal: Applied Economics* 8 (October): 189-221.

Grabowski, H. G., and N. D. Baxter. 1973. Rivalry in Industrial Research and Development: An Empirical Study. *The Journal of Industrial Economic*s 21 (July): 209-235.

Greene, W. H. 1993. *Econometric Analysis, 2nd edition*. New York: MacMillan.

Grindley, P. C., and D. J. Teece. 1997. Managing intellectual capital: Licensing and cross-licensing in semiconductors and electronics. *California Management Review* 39 (Winter): 1-34.

Hall, B., A. B. Jaffe, and M. Trajtenberg (2001): The NBER Patent Citation File: Lessons, Insights and Methodological Tools. NBER Working Paper Number 8498.

Hall, B. H., and R. H. Ziedonis. 2001. The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995. *RAND Journal of Economics* 32 (Spring): 101-128.

Hemphill, C. Scott, and Bhaven Sampat. 2012. Evergreening, Patent Challenges, and Effective Market Life in Pharmaceutical. *Journal of Health Economics* 31 (March): 327-339.

Hovenkamp, H., M. Janis, and M. Lemley. 2003. Anticompetitive Settlement of Intellectual Property Disputes. *87 Minn. L. Rev. 1719*.

Jaffe, Adam. 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value. *American Economic Review* 76 (December): 984-1001

Kortum, S., and J. Lerner. 1999. What is Behind the Recent Surge in Patenting? *Research Policy*, 28 (January): 1-22.

Lancaster, Tony. 1990. *The Econometric Analysis of Transition Data.* Cambridge University Press.

Lerner, J. 1994. The Importance of Patent Scope: An Empirical Analysis. *Rand Journal of Economics* 25 (Summer), 319-333.

———. 1995. Patenting in the shadow of competitors. *Journal of Law and Economics*, pp. 463-495.

Marco, A. C. 2007. The Dynamics of Patent Citations. *Economics Letters* 94 (February): 290-296.

Marco, A. C., and G. C. Rausser. 2002. Mergers and Intellectual Property in Agricultural Biotechnology. in *Economic and Social Issues in Agricultural Biotechnology*, ed. by R. Evenson, V. Santaniello, and D. Zilberman. CABI Publishing, New York.

Noel, Michael, and Mark Schankerman. 2013. Strategic Patenting and Software Innovation. *Journal of Industrial Economics* 61 (September): 481-520.

Pakes, A. 1986. Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica* 54 (July): 755-784.

Reinganum, J. F. 1989. The Timing of Innovation: Research, Development, and Diffusion," in *Handbook of Industrial Organization, Volume I*, ed. by R. Schmalansee, and R. D. Willig, chap. 14, pp. 850-908. Elsevier Science Publishers B.V.

Sakakibara, M., and L. Branstetter. 2001. Do Stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Law Reforms. *RAND Journal of Economics*, 32 (February): 77-100.

Schankerman, Mark. 1998. How valuable is patent protection? Estimates by technology field. *Rand Journal of Economics* 29 (Spring): 77-107.

Schankerman, Mark, and A. Pakes. 1986. Estimates of the value of patent rights in European countries during the post-1950 period. *Economics Journal* 96 (December): 1052-1076.

Sundaram, A. K., T. A. John, and K. John. 1996. An empirical analysis of strategic competition and firm values the case of R&D competition,. *Journal of Financial Economics* 40 (March): 459-486.

Sternitzke, Christian. 2013. “An Exploratory Analysis of Patent Fencing in Pharmaceutical: The case of PDE5 Inhibitors. Research Policy 42 (March): 542-551

Wooldridge, Jeffrey. 2010. Econometric Analysis of Cross Section and Panel Data. Cambridge, Mass.: MIT Press.

Ziedonis, R. H. 2004. Don’t Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms. *Management Science* 50 (June): 804-820.

# 8 Appendix

Table XX: Ceftin (Glaxo), Patent 4267320

| Citing Patent | Patent Assignee | Form | Process | Substitute |
|---|---|---|---|---|
| 4385054 | Glaxo | X | | |
| 4446317 | Glaxo | X | | |
| 4562181 | Glaxo | X | | |
| 4602012 | Glaxo | X | | |
| 4705784 | Sumitomo | | | X |
| 4820833 | Glaxo | X | | |
| 4897270 | Glaxo | X | | |
| 4994567 | Glaxo | | X | |
| 6060599 | Ranbaxy | | X | |
| 6323193 | Ranbaxy | X | | |
| 6346530 | Ranbaxy | | X | |
| 6384213 | Ranbaxy | | X | |
| 6485744 | Individual | X | | |
| 6534494 | Ranbaxy | | X | |
| 6833452 | Ranbaxy | | X | |

Table XX: Count of Self-citations and Citations (1983-2016)

| Year | Self-Citation | Citation |
|---|---|---|
| 1983 | 2 | 0 |
| 1984 | 2 | 2 |
| 1985 | 16 | 19 |
| 1986 | 23 | 22 |
| 1987 | 26 | 43 |
| 1988 | 37 | 55 |
| 1989 | 87 | 127 |
| 1990 | 75 | 146 |
| 1991 | 110 | 225 |
| 1992 | 115 | 247 |
| 1993 | 115 | 293 |
| 1994 | 129 | 309 |
| 1995 | 202 | 428 |
| 1996 | 263 | 517 |
| 1997 | 304 | 681 |
| 1998 | 371 | 839 |
| 1999 | 420 | 877 |
| 2000 | 372 | 941 |
| 2001 | 382 | 954 |
| 2002 | 348 | 1,201 |
| 2003 | 333 | 1,314 |
| 2004 | 244 | 1,104 |
| 2005 | 228 | 944 |
| 2006 | 305 | 1,318 |
| 2007 | 231 | 1,121 |
| 2008 | 268 | 1,330 |
| 2009 | 262 | 1,480 |
| 2010 | 401 | 2,366 |
| 2011 | 400 | 2,553 |
| 2012 | 544 | 3,065 |
| 2013 | 657 | 3,689 |
| 2014 | 730 | 4,199 |
| 2015 | 596 | 3,888 |
| 2016 | 249 | 1,833 |
| Total | 8847 | 38130 |

Table XX: Summary Statistics

| VARIABLES | (1) N | (2) Mean | (3) Std. Dev. | (4) Min. | (5) Max. | (6) p25 | (7) p50 | (8) p75 |
|---|---|---|---|---|---|---|---|---|
| Backward Citations | 1,120 | 12.03 | 19.71 | 0 | 247 | 2 | 6 | 13 |
| Independent Claim Count | 1,114 | 3.007 | 3.134 | 0 | 44 | 1 | 2 | 4 |
| Citations (normalized)* | 1,120 | 7.428 | 15.75 | 0 | 233 | 0.551 | 2.571 | 8 |
| Self-citations (normalized)* | 1,120 | 1.602 | 3.181 | 0 | 36.62 | 0.111 | 0.500 | 1.714 |
| Time-at-risk (years) | 1,120 | 17.96 | 8.217 | 0.0411 | 33.48 | 12.16 | 18.31 | 24.57 |

* Normalized by the number of backward citations for each patent.

Figure XX: Histogram of Orange Book listed patents per drug

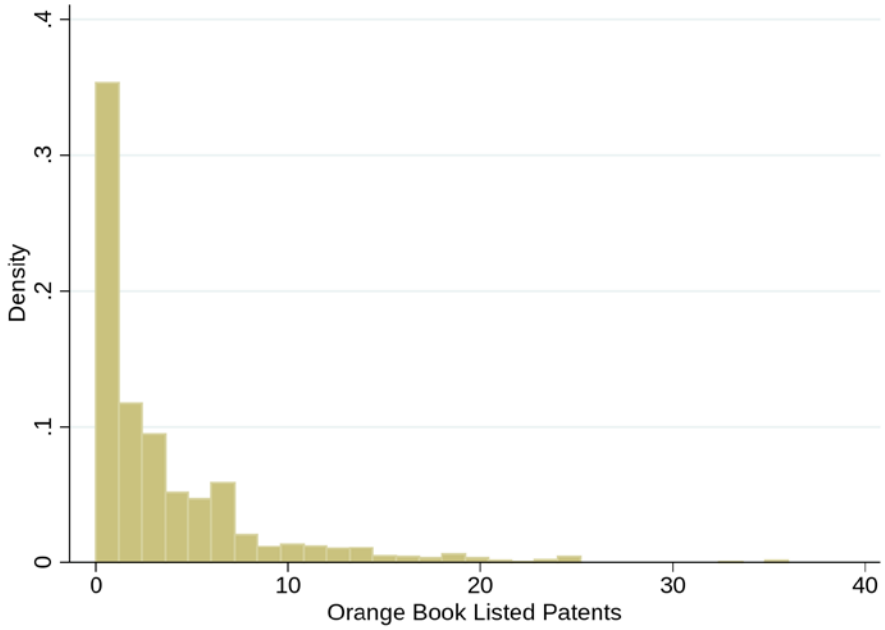


Figure XX: Histogram of self-citations per original drug patent

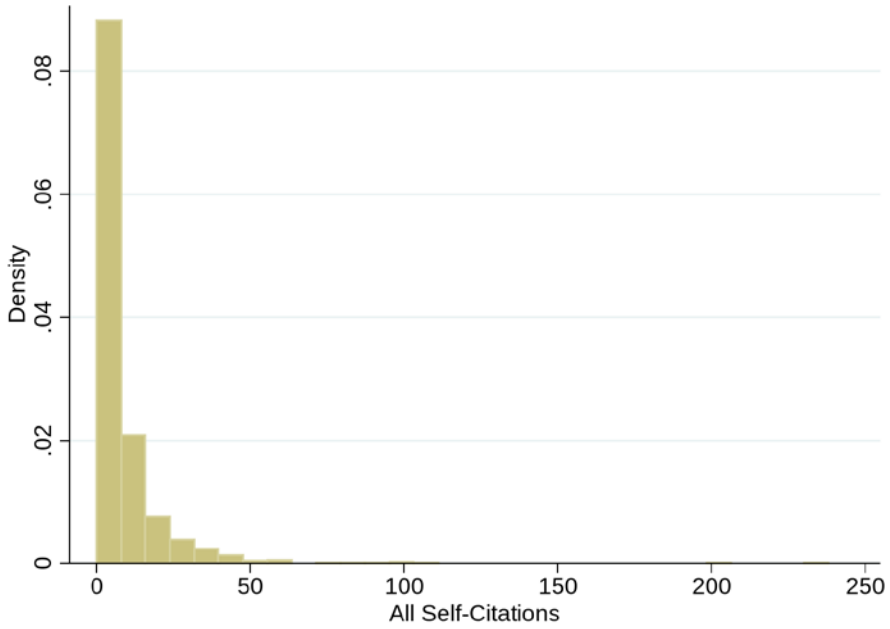


Figure XX: Histogram of citations per original drug patent

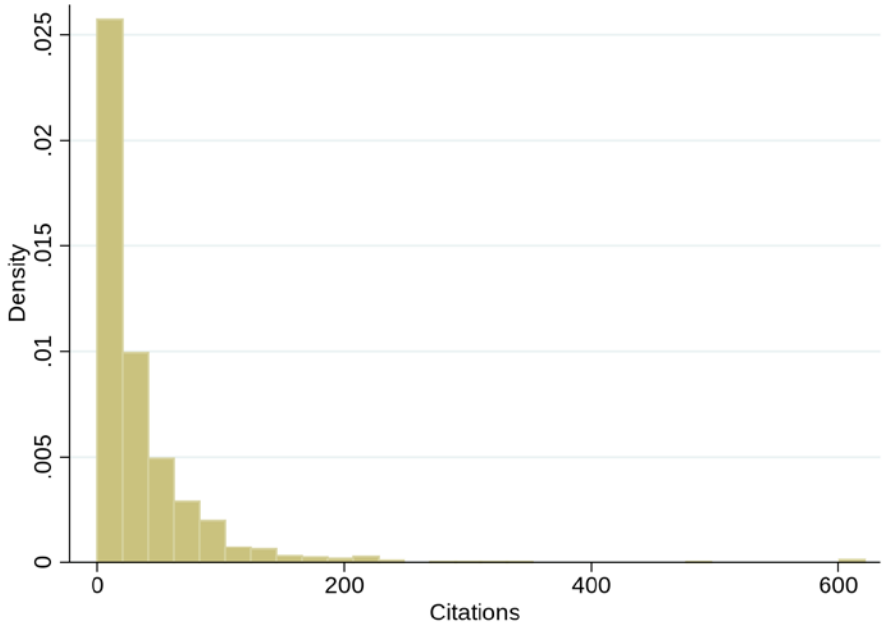


Figure XX: Kaplan-Meier hazard estimates: All self-citations

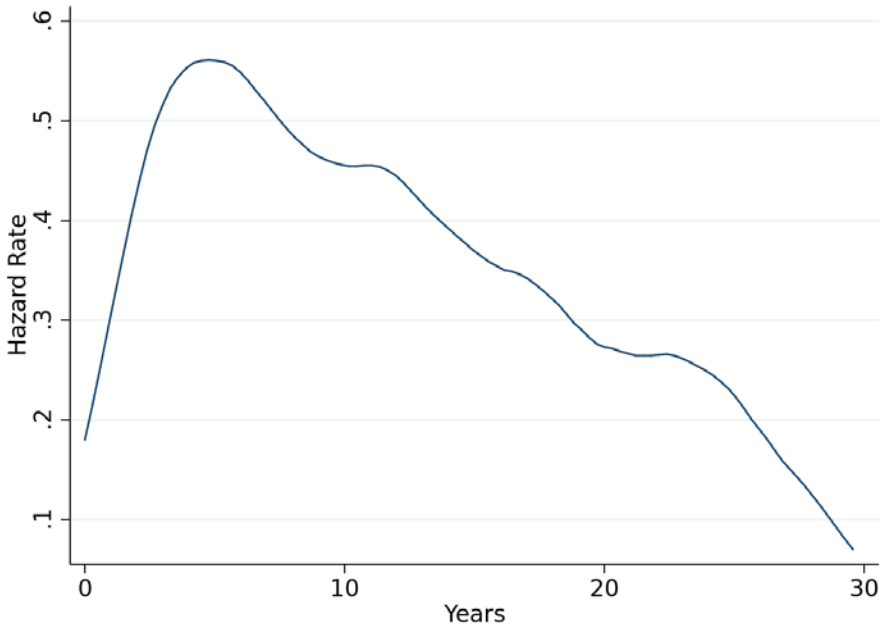


Figure XX: Kaplan-Meier hazard estimates: Citations

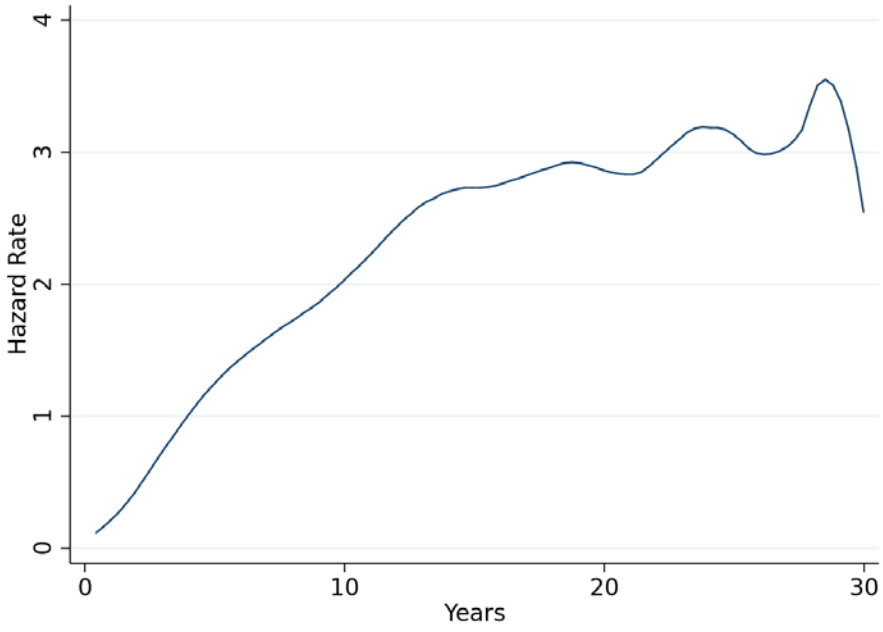


Table XX: Variables

| Main Variables | Definition | Full Sample | NDA Enter | NDA Exit |
|---|---|---|---|---|
| All Self-citations | Cumulative count of all patents citing the original patent assigned to the pioneer firm. | X | X | X |
| Citations | Cumulative count of all patents citing the original patent assigned to the other firms. | X | X | X |
| In Force | Percent of Orange Book patents that are unexpired. | X | | |
| Independent Claim Count | Number of Independent Claims per patent at grant | X | X | X |
| FDA Marketing Approval | Indicator of first NDA approval. | X | | |
| CDER | Center for Drug Evaluation and Research - Offices and Divisons (Therapeutic Class) | X | X | X |

Table XX: All self-citations (1983-2016)

| | (1) | (2) | (3) |
|---|---|---|---|
| | Cox | Weibull | Weibull frailty |
| All Self-citations† | 1.13*** | 1.12*** | 1.02*** |
| | (0.0028) | (0.0025) | (0.0035) |
| Citations† | 0.99*** | 0.99*** | 0.99** |
| | (0.0014) | (0.0014) | (0.0021) |
| Independent Claim Count | 1.04*** | 1.04*** | 1.04** |
| | (0.0028) | (0.0028) | (0.013) |
| In force | 1.01*** | 1.02*** | 1.02*** |
| | (0.0010) | (0.00084) | (0.00090) |
| FDA Marketing Approval | 0.77*** | 0.71*** | 0.93** |
| | (0.021) | (0.018) | (0.026) |
| Constant | | 0.040*** | 0.070*** |
| | | (0.0046) | (0.020) |
| Log Rho (Weibull) | | 1.10*** | 1.09*** |
| | | (0.014) | (0.011) |
| Log Theta (frailty) | | | 1.13* |
| | | | (0.056) |
| AIC | 114257.7 | -6558.0 | -11991.7 |
| BIC | 114725.7 | -6072.4 | -11497.2 |
| Log-likelihood | -57075.9 | 3334.0 | 6051.9 |
| LR test (Chi2) | 5111.7 | 5610.3 | 776.4 |
| N | 50520 | 50520 | 50520 |
| Drugs | 1109 | 1109 | 1109 |
| Failures | 8833 | 8833 | 8833 |

Notes: *** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated.

† Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. All regressions include year and theraeutic class dummies.

Table XX: All self-citations (1983-2016)
Exit at First NDA Approval

| | (1) | (2) | (3) |
|---|---|---|---|
| | Cox | Weibull | Weibull frailty |
| All Self-citations† | 1.31*** | 1.31*** | 1.13*** |
| | (0.011) | (0.011) | (0.013) |
| Citations† | 0.96*** | 0.96*** | 1.00 |
| | (0.0056) | (0.0055) | (0.0073) |
| Independent Claim Count | 1.03*** | 1.03*** | 1.05*** |
| | (0.0045) | (0.0044) | (0.014) |
| Constant | | 0.36*** | 0.37*** |
| | | (0.040) | (0.11) |
| Log Rho (Weibull) | | 1.13*** | 1.19*** |
| | | (0.018) | (0.020) |
| Log Theta (frailty) | | | 0.99 |
| | | | (0.067) |
| AIC | 43482.3 | -73.3 | -1981.0 |
| BIC | 43852.9 | 319.5 | -1580.7 |
| Log-likelihood | -21691.2 | 89.7 | 1044.5 |
| LR test (Chi2) | 1731.4 | 1771.5 | 266.7 |
| N | 12240 | 12240 | 12240 |
| Drugs | 934 | 934 | 934 |
| Failures | 3647 | 3647 | 3647 |

Notes: *** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated.

† Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. All regressions include year and theraeutic class dummies.

Table XX: All self-citations (1983-2016)
Enter at First NDA Approval

| | (1) | (2) | (3) |
|---|---|---|---|
| | Cox | Weibull | Weibull frailty |
| All Self-citations† | 1.15*** | 1.14*** | 1.02** |
| | (0.0034) | (0.0032) | (0.0069) |
| Citations† | 0.99*** | 0.99*** | 0.99* |
| | (0.0015) | (0.0015) | (0.0027) |
| Independent Claim Count | 1.04*** | 1.04*** | 1.02 |
| | (0.0039) | (0.0039) | (0.015) |
| Constant | | 0.64** | 4.73*** |
| | | (0.11) | (2.04) |
| Log Rho (Weibull) | | 0.52*** | 0.37*** |
| | | (0.027) | (0.032) |
| Log Theta (frailty) | | | 1.74*** |
| | | | (0.10) |
| AIC | 52504.1 | -8242.6 | -11855.3 |
| BIC | 52935.9 | -7793.8 | -11398.0 |
| Log-likelihood | -26201.0 | 4174.3 | 5981.6 |
| LR test (Chi2) | 3635.3 | 3479.2 | 166.6 |
| N | 35149 | 35149 | 35149 |
| Drugs | 934 | 934 | 934 |
| Failures | 4637 | 4637 | 4637 |

Notes: *** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated.

† Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. All regressions include year and theraeutic class dummies.

Table XX: Citations (1983-2016)

| | (1) | (2) | (3) |
|---|---|---|---|
| | Cox | Weibull | Weibull frailty |
| All Self-citations† | 0.98*** | 0.98*** | 0.99* |
| | (0.0020) | (0.0020) | (0.0041) |
| Citations† | 1.02*** | 1.02*** | 1.00 |
| | (0.00028) | (0.00026) | (0.00056) |
| Independent Claim Count | 1.02*** | 1.02*** | 1.04** |
| | (0.0015) | (0.0015) | (0.013) |
| In force | 1.01*** | 1.01*** | 1.00*** |
| | (0.00027) | (0.00021) | (0.00022) |
| FDA Marketing Approval | 1.03 | 1.00 | 1.05** |
| | (0.016) | (0.015) | (0.018) |
| Constant | | 0.017*** | 0.035*** |
| | | (0.0010) | (0.0086) |
| Log Rho (Weibull) | | 1.86*** | 1.86*** |
| | | (0.012) | (0.0096) |
| Log Theta (frailty) | | | 1.06 |
| | | | (0.046) |
| AIC | 479711.4 | -184948.8 | -211094.3 |
| BIC | 480179.4 | -184463.1 | -210599.8 |
| Log-likelihood | -239802.7 | 92529.4 | 105603.2 |
| LR test (Chi2) | 13075.5 | 13862.8 | 255.7 |
| N | 50520 | 50520 | 50520 |
| Drugs | 1109 | 1109 | 1109 |
| Failures | 37979 | 37979 | 37979 |

Notes: *** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated.

† Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. All regressions include year and theraeutic class dummies.

Table XX: Citations (1983-2016)
Exit at First NDA Approval

| | (1) | (2) | (3) |
|---|---|---|---|
| | Cox | Weibull | Weibull frailty |
| All Self-citations† | 0.87*** | 0.88*** | 0.95*** |
| | (0.0081) | (0.0080) | (0.015) |
| Citations† | 1.12*** | 1.11*** | 1.02*** |
| | (0.0026) | (0.0024) | (0.0034) |
| Independent Claim Count | 1.02*** | 1.02*** | 1.04** |
| | (0.0035) | (0.0035) | (0.016) |
| Constant | | 0.084*** | 0.055*** |
| | | (0.0093) | (0.017) |
| Log Rho (Weibull) | | 1.68*** | 1.97*** |
| | | (0.018) | (0.022) |
| Log Theta (frailty) | | | 1.23*** |
| | | | (0.075) |
| AIC | 86234.0 | -22391.7 | -26934.5 |
| BIC | 86612.1 | -21998.8 | -26534.3 |
| Log-likelihood | -43066.0 | 11248.9 | 13521.3 |
| LR test (Chi2) | 4124.1 | 3903.3 | 132.5 |
| N | 12240 | 12240 | 12240 |
| Drugs | 934 | 934 | 934 |
| Failures | 7659 | 7659 | 7659 |

Notes: *** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated.

† Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. All regressions include year and theraeutic class dummies.

Table XX: Citations (1983-2016)
Enter at First NDA Approval

| | (1) | (2) | (3) |
|---|---|---|---|
| | Cox | Weibull | Weibull frailty |
| All Self-citations† | 0.99*** | 0.99*** | 0.99 |
| | (0.0022) | (0.0022) | (0.0059) |
| Citations† | 1.02*** | 1.02*** | 1.00*** |
| | (0.00030) | (0.00027) | (0.00062) |
| Independent Claim Count | 1.02*** | 1.02*** | 1.02 |
| | (0.0018) | (0.0018) | (0.013) |
| Constant | | 0.063*** | 0.067*** |
| | | (0.0049) | (0.018) |
| Log Rho (Weibull) | | 1.55*** | 1.69*** |
| | | (0.016) | (0.017) |
| Log Theta (frailty) | | | 1.08 |
| | | | (0.052) |
| AIC | 330004.5 | -156819.3 | -176434.8 |
| BIC | 330436.3 | -156370.5 | -175977.5 |
| Log-likelihood | -164951.2 | 78462.6 | 88271.4 |
| LR test (Chi2) | 10569.9 | 10749.7 | 210.8 |
| N | 35149 | 35149 | 35149 |
| Drugs | 934 | 934 | 934 |
| Failures | 28027 | 28027 | 28027 |

Notes: *** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated.

† Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. All regressions include year and theraeutic class dummies.

Table XX: Fixed Effects Models (1983-2016)

| | (1) Self-Citations | (2) Citations |
|---|---|---|
| All Self-citations† | 1.15*** | 0.99*** |
| | (0.00) | (0.00) |
| Citations† | 0.98*** | 1.02*** |
| | (0.00) | (0.00) |
| Independent Claim Count | 1.03*** | 1.01*** |
| | (0.00) | (0.00) |
| In force | 1.02*** | 1.01*** |
| | (0.00) | (0.00) |
| FDA Marketing Approval | 0.73*** | 1.47*** |
| | (0.02) | (0.02) |
| Constant | 2.8e-45*** | 6E-49 |
| | (0.00) | (0.00) |
| Log Rho (Weibull) | 33.9*** | 30.7*** |
| | (0.92) | (0.49) |
| Year Fixed Effects | X | X |
| Therapeutic Class Fixed Effects | X | X |
| Year & TC Fixed Effects | X | X |
| AIC | -36398 | -272434 |
| BIC | -31604.9 | -270308 |
| Log-likelihood | 18722 | 136449 |
| LR test (Chi2) | 12502.3 | 26821.9 |
| N | 70600 | 70600 |
| Drugs | 1109 | 1109 |
| Failures | 8566 | 37260 |

Notes: *** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated.

† Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. All regressions include year and theraeutic class dummies.

| office_division* | Office Name | Division Name | Regression** |
|---|---|---|---|
| **CDER/OAP/DAIP** | Office of Antimicrobial Products | Division of Anti-Infective Products | 1 |
| **CDER/OAP/DAVP** | Office of Antimicrobial Products | Division of Antiviral Products | 2 |
| **CDER/OAP/DTOP** | Office of Antimicrobial Products | Division of Transplant and Ophthalmology Products | 3 |
| **CDER/ODEI/DCRP** | Office of Drug Evaluation I | Division of Cardiovascular and Renal Products | 4 |
| **CDER/ODEI/DNP** | Office of Drug Evaluation I | Division of Neurology Products | 5 |
| **CDER/ODEI/DPP** | Office of Drug Evaluation I | Division of Psychiatry Products | 6 |
| **CDER/ODEII/DAAAP** | Office of Drug Evaluation II | Division of Anesthesia, Analgesia, and Addiction Products | 7 |
| **CDER/ODEII/DMEP** | Office of Drug Evaluation II | Division of Metabolism and Endocrinology Products | 8 |
| **CDER/ODEII/DPARP** | Office of Drug Evaluation II | Division of Pulmonary, Allergy, and Rheumatology Products | 9 |
| **CDER/ODEIII/DBRUP** | Office of Drug Evaluation III | Division of Bone, Reproductive and Urologic Products | 10 |
| **CDER/ODEIII/DDDP** | Office of Drug Evaluation III | Division of Dermatology and Dental Products | 11 |
| **CDER/ODEIII/DGIEP** | Office of Drug Evaluation III | Division of Gastroenterology and Inborn Errors Products | 12 |
| **CDER/ODEIV/DMIP** | Office of Drug Evaluation IV | Division of Medical Imaging Products | 13 |
| **CDER/ODEIV/DNDP** | Office of Drug Evaluation IV | Division of Nonprescription Drug Products | 14 |
| **CDER/OHOP/DHP** | Office of Hematology and Oncology Drug Products | Division of Hematology Products | 15 |
| **CDER/OHOP/DOP1** | Office of Hematology and Oncology Drug Products | Division of Oncology Products 1 | 16 |
| **CDER/OHOP/DOP2** | Office of Hematology and Oncology Drug Products | Division of Oncology Products 2 | 17 |

** Table provided by the FDA's Center for Drug Evaluation and Research (CDER). ** Numbers correspond to specific regressions in Tables XX and XX.*

Table XXa: Self-Citations by Therapeutic Class (1983-2016)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| All Self-citations† | 1.14*** | 1.03 | 0.99 | 1.27*** | 1.21*** | 1.27*** | 1.31*** | 1.09*** | 1.01 |
| | (0.036) | (0.018) | (0.047) | (0.046) | (0.028) | (0.028) | (0.050) | (0.012) | (0.012) |
| Citations† | 1.04* | 0.97*** | 1.01 | 0.98* | 0.99 | 0.90*** | 0.98 | 0.99 | 1.02 |
| | (0.018) | (0.0080) | (0.0064) | (0.0075) | (0.013) | (0.016) | (0.014) | (0.0038) | (0.011) |
| Independent Claim Count | 1.19*** | 1.18*** | 0.97 | 1.01 | 0.96 | 1.05*** | 1.04*** | 1.05*** | 0.96* |
| | (0.036) | (0.032) | (0.032) | (0.024) | (0.027) | (0.013) | (0.011) | (0.0098) | (0.014) |
| In force | 1.04*** | 1.05*** | 1.02*** | 1.03*** | 1.03*** | 1.00 | 1.01 | 1.02*** | 1.04*** |
| | (0.010) | (0.0097) | (0.0031) | (0.0049) | (0.0051) | (0.0032) | (0.0032) | (0.0024) | (0.0065) |
| FDA Marketing Approval | 0.54*** | 0.88 | 0.96 | 0.66*** | 0.78* | 1.35* | 0.53*** | 0.67*** | 0.72*** |
| | (0.073) | (0.10) | (0.13) | (0.081) | (0.096) | (0.17) | (0.058) | (0.047) | (0.061) |
| Constant | 0.0016*** | 0.0071*** | 0.029*** | 0.013*** | 0.0097*** | 0.042*** | 0.0092*** | 0.030*** | 0.00022*** |
| | (0.0017) | (0.0069) | (0.018) | (0.0071) | (0.0063) | (0.024) | (0.0096) | (0.011) | (0.00026) |
| Log Rho (Weibull) | 0.92 | 0.99 | 1.04 | 1.07 | 1.05 | 1.11 | 1.10 | 1.21*** | 1.22*** |
| | (0.057) | (0.053) | (0.062) | (0.060) | (0.064) | (0.070) | (0.055) | (0.042) | (0.049) |
| AIC | 22.4 | -471.2 | 63.1 | 130.3 | -145.0 | -815.6 | -229.4 | -2159.3 | -1759.6 |
| BIC | 223.6 | -310.3 | 276.0 | 346.6 | 20.8 | -625.5 | -7.37 | -1918.2 | -1538.3 |
| Log-likelihood | 24.8 | 262.6 | 3.43 | -30.2 | 101.5 | 440.8 | 149.7 | 1115.7 | 914.8 |
| LR test (Chi2) | 371.6 | 367.2 | 225.5 | 344.6 | 313.5 | 406.2 | 530.7 | 800.7 | 1162.3 |
| N | 1974 | 2865 | 3234 | 3567 | 2247 | 2346 | 4201 | 5988 | 4114 |
| Drugs | 67 | 53 | 73 | 89 | 63 | 58 | 88 | 97 | 72 |
| Failures | 404 | 537 | 440 | 468 | 420 | 518 | 580 | 1214 | 858 |

*** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated. † Normalized independent variable counts are normalized by the number of backwards citations made by the original patent.

Regressions include year/theraeutic class dummies.

Table XXb: Self-Citations by Therapeutic Class (1983-2016)

| | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) |
|---|---|---|---|---|---|---|---|---|
| All Self-citations† | 1.04*** | 1.24*** | 1.22*** | 1.08 | 1.30*** | 1.12*** | 1.50*** | 1.00 |
| | (0.0045) | (0.036) | (0.047) | (0.098) | (0.036) | (0.011) | (0.12) | (0.042) |
| Citations† | 0.85*** | 0.95** | 1.00 | 1.02 | 0.91*** | 0.93*** | 0.97** | 1.00 |
| | (0.014) | (0.017) | (0.0080) | (0.016) | (0.022) | (0.015) | (0.0097) | (0.011) |
| Independent Claim Count | 1.14*** | 1.16*** | 1.03 | 1.74*** | 0.93 | 1.05* | 0.89 | 0.89 |
| | (0.036) | (0.031) | (0.030) | (0.14) | (0.043) | (0.021) | (0.058) | (0.068) |
| In force | 1.02*** | 1.01*** | 1.02*** | 1.06*** | 1.00 | 1.02** | 1.02* | 1.02*** |
| | (0.0024) | (0.0029) | (0.0049) | (0.011) | (0.0023) | (0.0053) | (0.0064) | (0.0046) |
| FDA Marketing Approval | 1.19 | 0.55*** | 1.07 | 0.42*** | 0.48*** | 1.00 | 0.94 | 1.24 |
| | (0.12) | (0.081) | (0.14) | (0.083) | (0.073) | (0.13) | (0.17) | (0.24) |
| Constant | 0.20*** | 5.3e-09 | 0.012*** | 0.00013*** | 0.000000095 | 0.0031*** | 0.039** | 0.054*** |
| | (0.065) | (0.0000054) | (0.0084) | (0.00015) | (0.000084) | (0.0034) | (0.039) | (0.034) |
| Log Rho (Weibull) | 1.30*** | 1.12 | 1.00 | 0.87 | 1.03 | 1.38*** | 0.79** | 1.03 |
| | (0.056) | (0.072) | (0.058) | (0.075) | (0.069) | (0.086) | (0.070) | (0.093) |
| AIC | -2741.6 | -330.4 | -247.2 | -74.9 | -460.1 | -1146.4 | 78.9 | -167.4 |
| BIC | -2522.9 | -125.7 | -34.4 | 58.1 | -307.5 | -969.0 | 236.6 | -13.9 |
| Log-likelihood | 1404.8 | 200.2 | 159.6 | 61.5 | 258.0 | 604.2 | -11.4 | 110.7 |
| LR test (Chi2) | 1865.9 | 662.8 | 393.2 | 404.2 | 350.1 | 880.9 | 191.0 | 147.9 |
| N | 4597 | 2561 | 2726 | 1888 | 1717 | 2253 | 2063 | 2179 |
| Drugs | 97 | 78 | 69 | 36 | 45 | 52 | 39 | 33 |
| Failures | 929 | 411 | 477 | 224 | 380 | 501 | 229 | 243 |

*** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated. † Normalized independent variable counts are normalized by the number of backwards citations made by the original patent.

Regressions include year/theraeutic class dummies.

Table XXa: Citations by Therapeutic Class

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| All Self-citations† | 0.97 | 0.93*** | 1.03 | 0.99 | 0.95*** | 0.96* | 0.97* | 1.02** | 0.91*** |
| | (0.017) | (0.0100) | (0.020) | (0.017) | (0.014) | (0.017) | (0.013) | (0.0071) | (0.013) |
| Citations† | 1.09*** | 1.02*** | 1.01*** | 1.00** | 1.05*** | 1.04*** | 1.06*** | 1.01*** | 1.05*** |
| | (0.0068) | (0.0024) | (0.0011) | (0.0013) | (0.0043) | (0.0067) | (0.0029) | (0.0012) | (0.0030) |
| Independent Claim Count | 1.03* | 1.15*** | 1.15*** | 1.10*** | 1.04*** | 1.01 | 1.01** | 1.06*** | 1.02 |
| | (0.013) | (0.016) | (0.014) | (0.0099) | (0.011) | (0.0090) | (0.0046) | (0.0052) | (0.0099) |
| In force | 1.01*** | 1.01*** | 1.00 | 1.00*** | 1.00*** | 1.01*** | 1.00*** | 1.01*** | 1.00*** |
| | (0.0014) | (0.0012) | (0.00073) | (0.00071) | (0.00097) | (0.0015) | (0.00074) | (0.00068) | (0.00086) |
| FDA Marketing Approval | 0.93 | 1.40*** | 0.78*** | 0.92 | 0.96 | 1.25* | 0.86** | 0.67*** | 0.79*** |
| | (0.089) | (0.11) | (0.053) | (0.062) | (0.073) | (0.11) | (0.050) | (0.029) | (0.040) |
| Constant | 0.019*** | 0.086*** | 0.0088*** | 0.045*** | 0.034*** | 0.0029*** | 0.030*** | 0.020*** | 0.0050*** |
| | (0.0045) | (0.019) | (0.0039) | (0.0081) | (0.0094) | (0.0010) | (0.0057) | (0.0037) | (0.0019) |
| Log Rho (Weibull) | 1.63*** | 1.57*** | 1.88*** | 1.85*** | 1.81*** | 2.34*** | 2.13*** | 2.02*** | 1.83*** |
| | (0.063) | (0.045) | (0.051) | (0.047) | (0.062) | (0.078) | (0.053) | (0.040) | (0.048) |
| AIC | -5199.8 | -10123.3 | -13471.5 | -13190.9 | -7386.2 | -8564.3 | -19645.2 | -24580.2 | -17655.5 |
| BIC | -4998.6 | -9962.4 | -13258.6 | -12974.6 | -7220.4 | -8374.2 | -19423.2 | -24339.1 | -17434.3 |
| Log-likelihood | 2635.9 | 5088.7 | 6770.7 | 6630.4 | 3722.1 | 4315.1 | 9857.6 | 12326.1 | 8862.8 |
| LR test (Chi2) | 751.4 | 690.2 | 1875.8 | 973.7 | 872.4 | 975.6 | 3041.3 | 2224.2 | 3529.5 |
| N | 1974 | 2865 | 3234 | 3567 | 2247 | 2346 | 4201 | 5988 | 4114 |
| Drugs | 67 | 53 | 73 | 89 | 63 | 58 | 88 | 97 | 72 |
| Failures | 1370 | 2121 | 2541 | 2795 | 1622 | 1644 | 3361 | 4428 | 2969 |

*** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated. † Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. Regressions include year/theraeutic class dummies.

Table XXb: Citations by Therapeutic Class

| | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) |
|---|---|---|---|---|---|---|---|---|
| All Self-citations† | 1.00 | 0.92*** | 1.03 | 0.97 | 1.04** | 1.00 | 1.34*** | 0.92*** |
| | (0.0050) | (0.017) | (0.019) | (0.026) | (0.014) | (0.0081) | (0.041) | (0.011) |
| Citations† | 1.05*** | 1.07*** | 1.01*** | 1.02*** | 1.04*** | 1.04*** | 0.99** | 1.03*** |
| | (0.0035) | (0.0028) | (0.0018) | (0.0047) | (0.0049) | (0.0027) | (0.0023) | (0.0013) |
| Independent Claim Count | 1.10*** | 1.10*** | 1.08*** | 1.18*** | 1.20*** | 1.02 | 0.74*** | 1.20*** |
| | (0.016) | (0.014) | (0.018) | (0.020) | (0.026) | (0.013) | (0.022) | (0.034) |
| In force | 1.00*** | 1.01*** | 1.00 | 1.01*** | 1.00 | 1.01*** | 1.00* | 1.00** |
| | (0.00069) | (0.00096) | (0.0011) | (0.0011) | (0.0010) | (0.0017) | (0.0011) | (0.0012) |
| FDA Marketing Approval | 1.05 | 1.11 | 0.97 | 1.95*** | 0.77** | 1.12 | 1.16 | 1.61*** |
| | (0.063) | (0.093) | (0.075) | (0.19) | (0.075) | (0.086) | (0.096) | (0.15) |
| Constant | 0.091*** | 0.044*** | 0.014*** | 0.085*** | 0.098*** | 0.0063*** | 0.12*** | 0.24*** |
| | (0.016) | (0.011) | (0.0043) | (0.019) | (0.027) | (0.0035) | (0.040) | (0.059) |
| Log Rho (Weibull) | 1.68*** | 1.60*** | 1.96*** | 0.88** | 1.59*** | 1.67*** | 1.70*** | 1.48*** |
| | (0.044) | (0.056) | (0.068) | (0.039) | (0.065) | (0.063) | (0.064) | (0.060) |
| AIC | -16602.5 | -9308.6 | -12446.1 | -6635.6 | -4938.8 | -8477.9 | -9696.6 | -13135.4 |
| BIC | -16383.8 | -9103.9 | -12233.3 | -6502.5 | -4786.2 | -8300.6 | -9538.9 | -12981.9 |
| Log-likelihood | 8335.2 | 4689.3 | 6259.0 | 3341.8 | 2497.4 | 4269.9 | 4876.3 | 6594.7 |
| LR test (Chi2) | 1582.1 | 1488.9 | 2793.9 | 1291.2 | 349.2 | 1821.1 | 1440.6 | 1658.9 |
| N | 4597 | 2561 | 2726 | 1888 | 1717 | 2253 | 2063 | 2179 |
| Drugs | 97 | 78 | 69 | 36 | 45 | 52 | 39 | 33 |
| Failures | 3336 | 1942 | 2039 | 1502 | 1167 | 1593 | 1704 | 1845 |

*** significant at the 0.01 level. ** significant at the 0.05 level. * significant at the 0.1 level. Standard errors are clustered by drug and coefficients are exponentiated. † Normalized independent variable counts are normalized by the number of backwards citations made by the original patent. Regressions include year/theraeutic class dummies.