NBER WORKING PAPER SERIES

THE NBER ORANGE BOOK DATASET: A USER'S GUIDE

Maya Durvasula
C. Scott Hemphill
Lisa Larrimore Ouellette
Bhaven N. Sampat
Heidi L. Williams

Working Paper 30628
http://www.nber.org/papers/w30628

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2022, Revised May 2023

We thank AnneMarie Bryson, Antara Joardar, Casey Lincoln, Tamri Matiashvili, and Maya Roy for excellent research assistance. We gratefully acknowledge financial support from the National Institute on Aging through grant P01AG005842 to the National Bureau of Economic Research, the National Science Foundation through grant DGE-1656518, the Knight Hennessy Scholars Program, the Alfred P. Sloan Foundation, and the Smith Richardson Foundation. Hemphill has at various times served as an expert witness in litigation involving Orange Booklisted patents. Sampat has served as a consultant in litigation on issues relating to pharmaceutical patenting and innovation, and to the Medicines Patent Pool for work on patents and access to medicines. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The NBER Orange Book Dataset: A User's Guide
Maya Durvasula, C. Scott Hemphill, Lisa Larrimore Ouellette, Bhaven N. Sampat, and Heidi L. Williams
NBER Working Paper No. 30628
November 2022, Revised May 2023
JEL No. O0,O3

ABSTRACT

This paper introduces a newly digitized, open-access version of the Food and Drug Administration's "Orange Book"—a linkage between approved small-molecule drugs and the patents that protect them. The Orange Book also reports any applicable regulatory exclusivity that prevents competitive entry. We summarize the Orange Book's coverage and discuss the opportunities and challenges associated with using these data for research. Empirical validations against various administrative datasets suggest that Orange Book records are, largely, complete and accurate. We conclude with a specific use case—calculating legal exclusivity periods for drugs—to highlight the types of choices that researchers must make when using this resource.

Maya Durvasula
Stanford University
579 Jane Stanford Way
Stanford, CA 94305
mdurvas@stanford.edu

C. Scott Hemphill
New York University School of Law.
40 Washington Sq. South
New York, NY 10012
USA
hemphill@nyu.edu

Lisa Larrimore Ouellette
Stanford Law School
559 Nathan Abbott Way
Stanford, CA 94305
ouellette@law.stanford.edu

Bhaven N. Sampat
Department of Health Policy and Management
Columbia University
722 W 168th Street, Room 486
New York, NY 10032
and NBER
bns3@columbia.edu

Heidi L. Williams
Department of Economics
Stanford University
579 Jane Stanford Way
Office 323
Stanford, CA 94305
and NBER
hlwill@stanford.edu

1 Introduction

Patents are widely used as measures of innovation. However, the connection between a particular commercialized product and a particular patent is typically unclear. Indeed, a given product might lack any patent protection, and a given patent might be associated with no product. Or more commonly, a product might fall within the scope of tens or even thousands of patents, and a patent might be associated with many commercialized products. In general, there is no standard or straightforward way to link products and patents.[1]

Pharmaceuticals are different. For any brand-name small-molecule drug marketed in the United States, the drug maker is generally required to report applicable patent protection to the Food and Drug Administration (FDA).[2] The FDA makes this information available in a public document called the "Orange Book."[3] The Orange Book thus provides the tight product-patent linkage that is normally missing—enabling researchers to investigate, for example, the spillovers of public funding for research, changes in the effective exclusivity for brand-name pharmaceuticals over time, the interplay between patent examination and litigation, the behavior of patent examiners, and the landscape of drug patents in other countries (Azoulay et al., 2019; Hemphill and Sampat, 2012; Lietzan and Lybecker, 2020; Lemley and Tu, 2022; Frakes and Wasserman, 2022; Sampat and Shadlen, 2015).

Each edition of the Orange Book provides a snapshot of unexpired patent protection at a moment in time. As patents on a drug expire and new patents are issued, these changes are reflected in later editions. The Orange Book also provides a snapshot of unexpired regulatory exclusivity granted by the FDA. For example, certain novel drugs receive five years of regulatory exclusivity that blocks the entry of generic competition, even in the absence of any patents. Combining multiple editions reveals a comprehensive picture of patent and regulatory protection as it evolves over a drug's lifecycle.

In this paper, we introduce a newly digitized, open-access dataset of Orange Book records from 1985 to 2016, and provide a "user's guide" to its opportunities and challenges for researchers.[4] The dataset combines information from print editions in the early years of the Orange Book (1985–1999) with digital data in later years (2000–2016). The dataset includes both patent protection and regulatory exclusivity. This linked data between products, patents, and exclusivity enables the measurement and analysis of important features of the biomedical innovation ecosystem. For example, the Orange Book provides researchers examining the pharmaceutical industry with opportunities to assess the scope of patent and regulatory protection on new products, as well as to describe the geographic and financial origins of innovative technology.

In Section 2, we provide background information about the contents of the Orange Book and the construction of this dataset. In Section 3, we offer an overview of the data files. In Section 4, we assess the completeness and accuracy of the dataset by comparing it to various external datasets. Section 5 presents an Orange Book "use case"—comparing different measures of market exclusivity conferred by legal protection with a measure of generic approval—and Section 6 concludes.

[1] Some producers disclose a linkage by marking their product with the patent number or with the URL of a web page containing the relevant details. For recent work using online disclosure as a source of linkages, see de Rassenfosse (2018).

[2] Small-molecule drugs are relatively simple chemical compounds, which are derived by chemical synthesis. Biologic drugs, in contrast, are typically manufactured using living organisms. These two classes of drugs are regulated separately in the United States.

[3] The official name is *Approved Drug Products with Therapeutic Equivalence Evaluations*, and it is available at `https://www.fda.gov/drugs/drug-approvals-and-databases/approved-drug-products-therapeutic-equivalence-evaluations-orange-book`.

[4] The datasets—and associated documentation—are available here: `https://www.nber.org/research/data/orange-book-patent-and-exclusivity-data-1985-2016`.

2 Background on the Orange Book

Key Takeaways

- Since 1985, sponsors of brand-name small-molecule drugs have been required to provide a list of nearly all associated patents to the FDA. These lists are published annually as "The Orange Book."
- The NBER Orange Book Dataset includes Orange Book editions published between 1985 and 2016. Researchers interested in up-to-date versions can retrieve new annual editions from the FDA's website.

The Orange Book began as a collection of information about the substitutability of approved generic drugs for brand-name small-molecule drugs, following the repeal of anti-substitution laws in many states. The first edition was published on 31 October 1980—hence the pumpkin-colored cover. In 1984, Congress passed the Drug Price Competition and Patent Term Restoration Act—typically referred to as the Hatch-Waxman Act—which established the modern regime for competition between brand-name and generic drugs. Among several other changes, the Act provides a pathway for a generic firm to market a competing, therapeutically equivalent version of a brand-name drug without completing its own extensive and costly clinical trials. Instead, the generic firm can file an Abbreviated New Drug Application (ANDA) with the FDA, which requires a demonstration of "bioequivalence" to ensure that the generic product will be absorbed by the body at the same rate and to the same extent as the brand-name drug.

The ANDA provisions facilitate rapid generic entry.[5] To provide potential generic entrants with notice of patents that may stand in the way of entry, the Hatch-Waxman Act requires brand-name firms to file a list of these patents with the FDA. Beginning in 1985, this patent information—including patent numbers and expiration dates—was included as an addendum to the Orange Book.

To enable historical analyses, we obtained PDF versions of each annual Orange Book publication from 1985 to 2016 via a Freedom of Information Act request.[6] We hand-entered or parsed the patent tables from each of these annual publications to create digital files.

Researchers who would like to work with up-to-date versions of the Orange Book can do so by retrieving text files of annual editions from the FDA's website. The FDA's web archive allows individuals to access historical versions of the website, including links to past annual Orange Book files, beginning in 2016.[7] For each additional year of data, researchers can access the associated text file containing records of unexpired patents and regulatory exclusivities held on approved drugs. When appended, these records are sufficient to extend the coverage of the digitized Orange Book data described here through the end of the most recent calendar year (see Appendix A for details).

3 An overview of the data files

Key Takeaways

- Firms are required to list patents that cover drug substances, products, and methods of use, and the FDA lists additional non-patent regulatory exclusivity periods.
- The NBER Orange Book Dataset includes full listings of all patents (FDA-drug-patents.dta) and full listings of all exclusivities (FDA-drug-exclusivity.dta) as they appear in each edition.

[5] As the other half of the of Hatch-Waxman bargain, the Act also provides partial patent term restoration to firms marketing brand-name drugs for portions of their patent term that were lost on the way to FDA approval. See below for additional detail.

[6] No Orange Book was published in 1986.

[7] The web archive is available here: `https://www.fda.gov/about-fda/about-website/fdagov-archive`.

A final digitized version of our Orange Book patent and exclusivity files is available from the National Bureau of Economic Research. Four data files—in Stata .dta format—are available.

The patent data file, *FDA-drug-patents.dta*, contains full listings of all patents associated with each product as they appear in each edition of the Orange Book. Observations include a variable that captures the edition from which the record was drawn. Each observation also includes the New Drug Application (NDA) number issued by the FDA, one or more product "active ingredients" (chemical components that create pharmacological activity), and the "proprietary name" (product trade name). The full dataset includes 5,511 unique patents associated with 2,173 distinct NDAs.

FDA regulations require firms to list patents that cover drug substances, drug products, and drug methods of use.[8] Beginning in mid-2003, firms were required to flag submitted patents to indicate whether they protect drug substances (active ingredients) or drug products ("a finished dosage form, *e.g.*, tablet, capsule, or solution that contains a drug substance").[9] Accordingly, the data file includes flags for drug substances ("DS") and drug products ("DP") beginning in the 2004 edition. As for method-of-use patents, beginning with the 1998 edition, the Orange Book contains a flag ("use-code") and detailed use codes describing the approved indication or use covered by the patent. The use codes are defined in *FDA-patent-use-codes.dta*. These codes may be helpful in determining whether certain classes of patents affect generic entry differently (see Section 5.2). We caution, however, that to our knowledge neither compliance with listing requirements nor accuracy of firm-supplied flags has been validated.

The exclusivity data file, *FDA-drug-exclusivity.dta*, contains full listings of all regulatory exclusivities associated with each product. The observations are organized in a manner that is analogous to the patent data file. Each type of exclusivity is associated with a distinct code. The codes are defined in *FDA-drug-exclusivity-codes.dta*.

The FDA's Orange Book product listing database contains additional information not captured in our dataset.[10] A notable example is the Products file, which identifies, for every FDA-approved drug product, the existence and identity of therapeutically equivalent alternatives.

4 How complete and accurate are the patent and exclusivity records?

Key Takeaways

- Nearly all new small-molecule drugs approved since 1985 have some form of patent or exclusivity listed.
- Certain patents are ineligible for listing (such as manufacturing processes), and the Orange Book does not record information for most biological drugs.
- Nearly all "important" patents for small-molecule drugs—patents likely to constrain generic entry—are listed.
- Patent expiration dates are generally accurate and reflect changes to term length. Records are not, however, consistently updated to reflect expiration due to maintenance fee non-payment.

In this section, we provide an overview of the coverage of patent and exclusivity records in the Orange Book for drugs approved in the United States. We also describe several validation exercises, which are intended to assess the

[8] NDA holders are supposed to list any patents that "could reasonably be asserted" against manufacturers of the drug. 21 U.S.C. § 355(b)(1)(A)(viii). NDA holders are likely to list patents they own or exclusively license, but not patents owned by unrelated third parties.

[9] Drug substance and drug product flags are reported for patents listed after 18 August 2003. See `https://www.fda.gov/drugs/drug-approvals-and-databases/orange-book-data-files` for details.

[10] Documentation and data files are available for download: `https://www.fda.gov/drugs/drug-approvals-and-databases/orange-book-data-files`.

completeness and accuracy of Orange Book patent and exclusivity records. Table 1 summarizes these validation exercises.

4.1 Drug coverage

The NBER Orange Book Dataset includes records for all small-molecule drug products approved by the FDA under an NDA that have received any legal protections, including patents or regulatory exclusivities. Of the 2,458 NDAs approved between 1985 and 2015,[11] 1,682 (68%) have at least one patent, and 1,724 (70%) have at least one regulatory exclusivity. 1,958 NDAs (80%) include at least one patent or exclusivity reported in the Orange Book. The prevalence of legal protection is unsurprising, given the documented importance of intellectual property for R&D and product innovation in pharmaceutical markets (Cohen, Nelson, and Walsh (2000), Levin et al. (1987)).

The set of NDAs includes both new molecular entities (NMEs) and line extensions. An NME is a drug with a novel active ingredient. Line extensions include new dosage forms and formulations associated with previously approved products.[12] NMEs are particularly likely to have legal protection. For the 796 NMEs approved over the same period, 678 (85%) have at least one patent and 751 (94%) have at least one regulatory exclusivity in the Orange Book. 761 (96%) include at least one patent or exclusivity.

The focus of the Orange Book is small-molecule drugs, which are drugs with relatively simple chemical structures that can be well characterized by scientists. Biologic drugs—more complex products that often lack fixed chemical structures, such as therapeutic proteins and vaccines—are largely excluded. Biologic drugs present a distinct set of scientific and regulatory challenges and have a different pathway to regulatory approval. Biologic drugs are approved under so-called biologics license applications (BLAs), not NDAs.[13] Prior to 2020, certain biologic drugs were approved under NDAs, rather than BLAs, and thus have listed patents and exclusivities in archival versions of the Orange Book.[14]

For researchers seeking to investigate patent protections for biologic drugs, we are unaware of any publicly available data source that would contain information comparable to the Orange Book. Since 2014, biologic drugs and "biosimilars" (analogous to small-molecule generics) have been reported in an FDA document called, in a nod to the Orange Book, the Purple Book.[15] However, the Purple Book does not include comprehensive listings of patents associated with biologic drugs. Recent legislation requires some patents to be listed in the Purple Book, somewhat improving transparency for this class of drugs.[16]

4.2 Patent coverage

A key question is which patents are, in fact, listed in the Orange Book.[17] The answer varies across drugs and reflects an interaction of legal requirements and firm strategy. A firm marketing a brand-name drug is required to list any patent containing at least one claim that covers the drug's active ingredient, its formulation, or any method of use pertaining to

[11] We obtained a count of NDAs using the Drugs@FDA database: `https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm`.

[12] Raw data on NDAs and NMEs can be downloaded via the FDA's Drugs@FDA database: `https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm`. A frequently updated compilation of NMEs can also be downloaded from the FDA: `https://www.fda.gov/drugs/drug-approvals-and-databases/compilation-cder-new-molecular-entity-nme-drug-and-new-biologic-approvals`.

[13] Between 1985 and 2015, 85% of all NMEs were approved under NDAs. Between 1985 and 1990, 6.3% of NMEs were approved under BLAs, a share that has grown over time, to 11% between 1995 and 2000, 18% between 2005 and 2010, and 26% between 2015 and 2020. Data on NMEs were drawn from the database linked in Footnote 12.

[14] In March 2020, some biologics that were initially approved as NDAs were reclassified as BLAs. Within this set of drugs that were reclassified, we identify 14 NMEs approved before 2017 with at least one patent and 22 NMEs with at least one regulatory exclusivity listed in the Orange Book. See details here: `https://www.fda.gov/media/119229/download`.

[15] The official name is *Lists of Licensed Biological Products with Reference Product Exclusivity and Biosimilarity or Interchangeability Evaluations*.

[16] See `https://www.bioprocessonline.com/doc/patent-transparency-for-biologics-biosimilars-the-revamped-purple-book-0001` for details.

[17] Note that we do not attempt to identify any drugs that, if all patents were properly listed, *should* appear in the Orange Book. As such, we focus on assessing the completeness of patent listings for drugs with at least one listed Orange Book patent.

an approved indication. Eligible patents issued after drug approval must be listed within thirty days of issuance.[18] Some types of patent are ineligible for listing, however, such as patents on manufacturing processes, packaging, metabolites, and intermediates.[19]

Firms face strong incentives to list all eligible patents because a generic firm that seeks to market a competing product prior to patent expiration is required to "challenge" every listed patent. The challenge takes the form of a legal certification that each patent is invalid or not infringed by the proposed generic product. These so-called "Paragraph IV" challenges are common, often result in patent litigation, and are a major source of generic entry.[20] Listing patents confers an additional advantage in the event of litigation. If the brand-name firm elects to file an infringement lawsuit within 45 days of receiving notice of a Paragraph IV challenge, FDA approval of the generic firm's ANDA is automatically blocked ("stayed") for 30 months while the suit is considered by a court.[21]

Notably, the FDA does not conduct audits of any kind to determine either the completeness or correctness of Orange Book patent listings; instead, they defer to firm self-reporting. Previous research uses Orange Book records to suggest a rise over time in "secondary" patenting listings for pharmaceutical products—those patents covering not the drug's active ingredient but rather other peripheral aspects of the molecule (Kapczynski, Park, and Sampat, 2012; Hemphill and Sampat, 2011). However, these secondary patents are more likely to be challenged and litigated, especially for high-sales drugs (Hemphill and Sampat, 2012), and when litigated to completion, the generic challenger typically prevails (Hemphill and Sampat, 2013).[22]

An important takeaway is that all patents listed in the Orange Book may not be equivalent in terms of their exclusionary force or their effects on competition. In some research contexts, distinguishing between primary and secondary patent protection may be important (Hemphill and Sampat, 2012; Frakes and Wasserman, 2022; Gupta, 2021; Kyle, Sampat, and Shadlen, 2020). In other contexts, there may be patents that are important to track for research purposes that are not listed (because of non-compliance with reporting requirements or because the patents fall into categories that cannot or need not be listed, as described above).[23] Below we report the results of several validation exercises in an effort to evaluate the completeness of the Orange Book dataset.

4.2.1 Comparison to IQVIA/Ark

Numerous firms manually construct records of the "landscape" of drug patents, as part of competitive intelligence and freedom-to-operate analyses. These efforts typically draw on the Orange Book and integrate supplemental expert searches. In addition to Orange Book data, some researchers have used data from the IQVIA/Ark Patent Intelligence database (formerly, IMS LifeCycle Patent Focus) to identify patents on particular drugs, which can be particularly useful for tracking patents on drugs not covered by the Orange Book (*e.g.*, biologic drugs not subject to Hatch-Waxman listing

[18] See 21 U.S.C. § 355(c)(2).

[19] See 21 C.F.R. § 314.53(b)(1).

[20] The name comes from the relevant "Paragraph" of the Hatch-Waxman statute. See 21 U.S.C. § 355(j)(2)(A)(vii)(IV). The FDA maintains an up-to-date list of Paragraph IV challenges. See `https://www.fda.gov/drugs/abbreviated-new-drug-application-anda/paragraph-iv-drug-product-applications-generic-drug-patent-challenge-notifications`. For method-of-use patents, instead of filing a Paragraph IV certification, a generic applicant may assert that the patent does not cover any use for which the applicant seeks approval. See 21 U.S.C. § 355(j)(2)(A)(viii).

[21] See 21 U.S.C. § 355(j)(5)(B)(iii). However, an ANDA can be approved within the 30-month stay period if, before the stay's expiration, a court enters a decision in favor of the generic entrant on the question of patent infringement or validity. Additionally, the length of the stay may be altered by the court if either party fails to cooperate in advancing litigation.

[22] In response to concern about improper Orange Book listings, the FDA has created a process to challenge the accuracy of these patent listings. These changes were prompted in part by the Supreme Court's holding (in *Caraco v. Novo Nordisk*, 566 U.S. 399 (2012)) that a generic drug manufacturer may file a counterclaim in litigation to require a brand-name drug manufacturer to correct Orange Book patent "use codes" that inaccurately describe the coverage of a particular patent. A list of Orange Book patent listing disputes is available here: `https://www.fda.gov/drugs/drug-approvals-and-databases/orange-book-patent-listing-dispute-list`.

[23] See this editorial on the consequences of unlisted process patents: `https://patentlyo.com/patent/2006/05/editorial_the_o.html`.

requirements), non-listable patents, non-U.S. patents on drugs, and non-granted patent applications. IQVIA/Ark also provides a range of value-added information about patents beyond what the Orange Book provides, including specific information about what the patents cover (consistent with the point made above that not all drug patents are created equal). Ark Patent Intelligence also performs "constraint analysis" to determine, in its own view, "whether or not the claims in the patent will prevent a generic or biosimilar version of [a drug] from entering a commercial market whilst enforced" (Ark Patent Intelligence, 2019). The firm bases its determination on data collected from "product labels, . . . press releases and clinical trial[s]." Each patent is classified into one of three categories. A patent is *constraining* if it completely blocks the "launch [of a] generic equivalen[t]." A patent is *not constraining* if it completely fails to do so. And a patent is *partially constraining* if "it is possible to develop a generic equivalent but it cannot be marketed for . . . particular product attributes"—for example, a particular indication.

For another paper (Kyle, Sampat, and Shadlen, 2020), one of the authors (Sampat) obtained IQVIA/Ark data for 550 NMEs approved between 1995 and 2017 that had at least one Orange Book patent. There are 6,522 unique patents in IQVIA/Ark for these drugs, but only 2,692 listed in the Orange Book (in any edition through 2021). Orange Book patents are almost always listed in IQVIA/Ark as well, which is reasonable as the Orange Book is one of the inputs for IQVIA/Ark. However, approximately three-fifths of the time, an IQVIA/Ark patent associated with a given drug appears nowhere in the Orange Book. These unlisted patents appear to be disproportionately from patent categories that are explicitly not allowed in the Orange Book; for example, 93% of what IQVIA/Ark categorizes as "process" patents are explicitly excluded. IQVIA/Ark may also include patents on unapproved uses for the molecule, or formulations or molecular forms not on the market. Of the IQVIA/Ark patents listed as constraining entry for a given drug, 75% are listed as protecting the same drug in the Orange Book. By contrast, only 26% of non-constraining patents are linked to the associated drug in the Orange Book. Taken together, comparison with the IQVIA/Ark data suggests that there are many patents associated with FDA-approved drugs that are not listed in the Orange Book. However, the vast majority of patents that may be viewed as mattering for competition—in the sense of constraining generic entry—appear in the Orange Book.

4.2.2 Litigation records

Next, we investigate the extent to which pharmaceutical patents involved in litigation are reported in the Orange Book. In this context, a similar picture emerges. As noted above, a common form of litigation associated with pharmaceutical patents involves Paragraph IV challenges, in which a generic firm challenges active Orange Book-listed patents held by brand-name firms in the process of preparing to enter with a competing product. A brand-name firm need not limit its litigation strategy to listed patents, however. For example, it might assert infringement of a manufacturing process patent, which is not eligible for listing.

To examine the prevalence of Orange Book-listed patents in brand-generic litigation, we draw on records from ParagraphFour.com, which compiles information about both Paragraph IV and other patent disputes. Our data cover cases resolved (adjudicated or settled) since 1 November 2003.[24] There were more than 700 drugs associated with these cases, and 2,286 unique litigated patents. Of the litigated patents, 214 (9.4%) were not listed in the Orange Book (in any edition through 2021). Overall, these data suggest that for brand-generic disputes that are litigated via the Paragraph IV channel, unlisted patents play a relatively minor role.

[24] Last accessed: 3 March 2023.

4.2.3 Extended patents under Hatch-Waxman

The Hatch-Waxman Act allows a drug sponsor to delay patent expiration on a drug containing a new active ingredient. Such patent term extension (PTE) permits the sponsor to recoup some of the time lost during clinical trials and the FDA review process.[25] A firm may seek PTE for one patent per approved drug.[26] A firm can receive a PTE of up to five years, and the remaining patent term after this extension has been applied cannot exceed 14 years beyond the date of FDA approval. Extension time is calculated based on time spent in regulatory review plus half of the time spent in clinical trials, up to the five-year maximum.[27] Patents eligible for PTE must be in force when the drug is approved by the FDA and must cover the product or substance, a method of using the product, or a method of manufacturing the product.

In principle, a brand-name drug maker might choose to extend a patent not listed in the Orange Book, such as a manufacturing process patent. Eisenberg (2012) notes that, in theory, this choice presents a strategic dilemma for a firm. A patent filed on a method of manufacture might be filed later than (say) an active ingredient patent, and thus confer more years of patent protection. On the other hand, the claims may be weaker and less likely to impede generic entry. In practice, active ingredient patents appear to be a primary target of PTE. Hemphill and Sampat (2012) find that for a sample of drugs that experienced generic entry between 2001 and 2010, 79% of the active ingredient patents were extended, compared to 13% of non-active ingredient patents. We compare the universe of Orange Book patents to a list of granted PTEs maintained by the USPTO, with grant dates listed through 2012. We examine the set of 397 drugs that received PTE and that have at least one listed Orange Book patent. For these 397 drugs, 382 (96%) elected to extend a patent that is listed in the Orange Book.[28]

4.3 Patent expiration

For each record, the Orange Book provides information about product-patent linkages and the expiration dates of all linked patents. To assess accuracy, we investigate whether legally mandated changes in the expiration date of certain sets of patents are reflected in the data file. We examine the following four sources of change:

1. Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS): To comply with the 1994 TRIPS Agreement, the standard U.S. patent term was shifted from a 17-year term that began upon patent issuance to a 20-year term than begins at patent filing. This change took effect in June 1995. For patents filed before 08 June 1995—including previously issued patents that were in force at the time of this change—patent term was set to be the later of the two potential expiration dates. For patents that had spent less than three years under review at the USPTO at the time of this change, TRIPS increased their patent terms. As a result, the maximum increase in a patent's term due to TRIPS is three years.

2. Patent Term Extension (PTE): As noted above, PTE allows firms to restore up to five years of patent protection for time that a product has spent during clinical trials and under regulatory review. In addition, patents that will expire before product approval, or that are under review for PTE but may expire during the PTE determination

[25] PTE is codified at 35 U.S.C. § 156. In Section 4.3, we describe validation of the quality of Hatch-Waxman/PTE data using USPTO expiration records.

[26] This general rule has a narrow exception: where a drug has separate NDAs approved on a single day, the firm may be eligible to extend multiple patents.

[27] The statute and corresponding regulations at 37 C.F.R. § 1.775 contain additional details. For example, the extension may be reduced if the applicant did not act with due diligence during the regulatory review period.

[28] Of the 803 unique trade names associated with extensions, 397 correspond to Orange Book-listed drugs, and of the 809 patents extended, 510 are in the Orange Book. Based on our review of the remaining extended patents, nearly all correspond to biologic drugs (including vaccines) and medical devices.

process, are eligible to apply for an interim extension.[29] Interim PTE may be followed by full PTE, though this is not necessarily the case. If the patent does receive full PTE, interim PTE is counted toward this extension.

3. Patent Term Adjustment (PTA): Since May 2000, patents have been eligible to receive a Patent Term Adjustment (PTA), which extends patent term to account for delays caused by the USPTO during patent examination.[30]

4. Maintenance Fee Non-Payment: Patent owners are required to pay fees to the USPTO to maintain their patent rights at 3.5-, 7.5-, and 11.5-year intervals after patent issuance.[31] Non-payment results in termination of patent rights.

TRIPS To determine whether changes associated with TRIPS are reflected in Orange Book records, we examine the two Orange Book editions in which this change is likely to be most pronounced: 1995 and 1996, which reflect listings from the years 1994 and 1995 respectively. If the TRIPS changes impacted Orange Book listings, we would expect that patent expiration dates should shift forward in 1996, relative to the 1995 expiration dates, by no more than three years. Of the 415 patents that were listed in the Orange Book in both 1995 and 1996, 159 (38%) changed their listed expiration dates between 1995 and 1996. 154 patents reported a later expiration date in 1996 compared to 1995. Among these 154 extensions, only 1 patent reported an expiration extension that was greater than three years. The remaining expiration extensions are less than three years and, thus, are consistent with the TRIPS adjustment to patent terms.

We conduct the same comparison exercise with two other sets of years to confirm that these changes between 1995 and 1996 are due to TRIPS, rather than standard year-to-year modifications. We consider 1990-1991 and 1999-2000, five years before and after the 1994-1995 editions. In 1989-1990, of the 338 patents appearing in both editions, only 12 (4%) had any change in expiration. In years 2000-2001, 55 (6%) of 849 patents had different expiration dates in the two editions.

Patent term extension To determine whether the Orange Book reflects PTE and interim PTE, we compare Orange Book records to a list of all patents that received PTE between 1984 and 2016.[32]

Of the 4,027 patents listed in the Orange Book from 1985 to 2016, 402 (10%) appear in the PTE data and have non-missing patent expiration dates. For these patents, we compare listed patent expiration dates across the Orange Book and the USPTO Patent Center database. 357 (89%) have identical expiration dates in the Orange Book patents dataset and the USPTO PTE data. For the remaining patents, we investigate the source of the discrepancy. For 17 (4%), the Orange Book expiration date is correct. The remaining 28 (7%) reflect a variety of Orange Book errors.[33] These results are reported in Table 2, and further details are available in Appendix B. Although researchers should take note of cases

[29] See 35 U.S.C. § 156(e)(2) and § 156(d)(5).

[30] See 35 U.S.C. § 154(b).

[31] There is a six-month grace period after each deadline, during which the patent owner must pay a surcharge in addition to the maintenance fee. Data on maintenance fee payment are drawn from USPTO records downloaded from `https://bulkdata.uspto.gov/data/patent/maintenancefee/MaintFeeEvents.zip`.

[32] We accessed a version of this list current as of 18 January 2016 using the Internet Archive's Wayback Machine, here: `https://web.archive.org/web/20160118153156/https://www.uspto.gov/patent/laws-and-regulations/patent-term-extension/patent-terms-extended-under-35-usc-156`. Prior to the Hatch-Waxman Act, Congress extended individual patents on two products, Nutrasweet (aspartame) and Forane (an anesthetic) under 35 U.S.C. § 155 and 35 U.S.C. § 155A. See Patent Terms Extended Under 35 U.S.C. § 155; 35 U.S.C. § 155A, `https://web.archive.org/web/20170628065756/http://www.uspto.gov/patents/resources/terms/155.jsp` (listing patents for these two products). These products do not appear in our sample. Repeal of these statutory provisions became effective in 2012.

[33] The errors include a failure to take account of PTE (16 patents), typographical and similar clerical errors (7), miscalculation of TRIPS extensions (2), and errors in expiration date calculation that are included in both the Orange Book and USPTO data (2). Finally, one patent is associated with a drug that was later reclassified under a BLA application; while early editions of the Orange Book report accurate expiration dates, the drug is dropped from later editions and patent terms post-extension are never reported.

in which Orange Book listings are inaccurate, our findings suggest that Orange Book expiration dates affected by PTE are correct in 93% of cases.

Patent term adjustment Next, we collect records of PTA from the USPTO Historical Masterfile, which contains information on patents issued through 2014.[34]

We focus on the set of patents that received only PTA. The previous subsection assessed the accuracy of all patents that received PTE, including those records that received both PTE *and* PTA. Any discrepancies associated with this set of patents are investigated in Table 2.

We observe 943 Orange Book patents with reported PTA, which did not also receive PTE. Of these 943 patents, 366 (38%) had correct expiration dates (expiration dates that appropriately reflected PTA) in the Orange Book. In 144 cases (15%), the discrepancy between USPTO records and Orange Book records is exactly equal to PTA length, suggesting that the Orange Book was not updated to reflect PTA.

For the remaining 46 percent of observations (N=434)—for which USPTO and Orange Book expiration dates differed by more or less than the duration of PTA—we hand-checked a random 10 percent sample (N=43). Within this set, Orange Book expiration dates were correct for 30 (70%) of records. Discrepancies across the two datasets were explained in all cases by differences in the calculation of patent expiration dates and inconsistent reporting of maintenance fee non-payment, terminal disclaimers, and Certificates of Correction amending earlier decisions about PTA.[35] Table 3 and Appendix B.2 provide details on these cases. As with PTE, our findings suggest that the Orange Book reports accurate expiration dates for the majority of records. Appendix Section B summarizes our process for determining correct expiration dates for patents and provides suggestions for researchers who may be interested in doing so for a larger set of observations.

Finally, we examine PTA recalculation. Due to a USPTO error in interpreting the relevant section of the federal code, a narrow set of patents became eligible for a reconsideration of the duration of PTA.[36] Eligibility was limited to patents filed no earlier than 29 May 2000 and issued no later than 2 March 2010. Patent applicants were given 180 days from the issuance of their patent or, if the patent had not yet been issued, two months from the initial determination of PTA to request PTA recalculation. However, because this notice was issued in January 2010, only patents issued between August 2009 and March 2010 were realistically eligible for this recalculation. Petitions associated with these recalculations are flagged in the USPTO's Patent Center database with the event code "PTGR." We observe zero patents that received PTA in our sample that also have the PTGR event code. We conclude that there are no Orange Book patents in our sample that received PTA that also received a PTA recalculation on account of this error.

Maintenance fee non-payment We successfully link 4,929 Orange Book patents to USPTO records on maintenance fee payment. Of these patents, 2,656 (54%) expire at full term (20 years), while 997 (20%) expire at 12 years, 1,212 (25%) expire at 8 years, and 64 (1%) expire at four years. Spot-checking of these records suggests that maintenance fee non-payment is not reported in the Orange Book, as patents continue to appear despite having expired. This is unsurprising, given that firms with expired patents face no incentive to continue reporting changes to the FDA. Researchers interested in accurate measures of actual patent term should note, then, that patent expiration dates in the Orange Book do not reflect this form of patent term truncation and that integrating additional USPTO records may be valuable.

[34]The Historical Masterfile can be accessed here: `https://developer.uspto.gov/product/historical-masterfile`. Records of PTA are also available via the USPTO's Patent Center database: `https://patentcenter.uspto.gov`.

[35]Terminal disclaimers are typically filed by inventors who hold patents on several variants of the same technology. When later-filed patents may be obvious extensions of earlier-filed patents, inventors can "disclaim" (or dedicate to the public) any remaining patent term on the later-filed patent after the earlier-filed patent expires.

[36]See *Wyeth v. Kappos*, 591 F.3d 1364 (Fed. Cir. 2010); `https://www.uspto.gov/sites/default/files/patents/law/aipa/pta/wyeth_faqs_20100422.pdf`.

4.4 Exclusivity records

The second major component of the data file is a comprehensive set of FDA exclusivity records associated with each approved drug. Four are most important, in terms of their effectiveness in blocking competitive entry, particularly from generics:

1. New chemical entity (NCE) exclusivity: The FDA awards a five-year exclusivity period to brand-name drugs containing new active ingredients. NCE protection takes the form of *data* exclusivity: a generic firm is prohibited from relying on the brand-name firm's clinical trial data.

2. Orphan drug exclusivity (ODE): The FDA awards a seven-year exclusivity period upon the approval of a treatment for a rare disease. ODE is an example of *market* exclusivity: a prospective entrant may not market the same drug, even if it conducts its own clinical trials.[37]

3. Pediatric extension: The FDA awards a six-month extension of existing exclusivities, including NCE and ODE, for conducting pediatric clinical trials. The extension also applies to unexpired patents.

4. Generic exclusivity: Generics are eligible for their own form of exclusivity. In particular, the first generic to file a Paragraph IV patent challenge is potentially eligible for a 180-day period of exclusivity that blocks the approval of other generics.

In section 5.1.1, we discuss these and other types of exclusivity in greater detail.

In contrast to Orange Book patent listings, which are self-reported to the FDA by drug manufacturers, exclusivity periods are granted by the FDA itself and hence recorded directly by the agency. We therefore expect Orange Book records of FDA-granted exclusivities to be both complete and accurate. We nonetheless conduct two validation exercises.

First, we confirm that all drugs reported (separately) by the FDA to have received ODE match Orange Book exclusivity records. Specifically, we make use of an FDA-compiled list of all drugs that have received an orphan drug designation.[38] Given that the Orange Book focuses on FDA-approved drugs, we restrict this list to orphan designations for drugs that have actually been approved for marketing; we likewise restrict attention to drugs that received orphan drug designations before 2016, as our Orange Book data is current through the end of 2015. For every such drug, we can identify a corresponding ODE award in the Orange Book exclusivity file.

Second, we similarly confirm that all drugs reported (separately) by the FDA to have received pediatric exclusivity match Orange Book exclusivity records.[39] We observe pediatric exclusivity codes in the Orange Book for all drugs on the pediatric exclusivity list provided separately by the FDA. Consistent with our expectations, product-level Orange Book exclusivity records thus appear to be both complete and accurate.

[37] See Thomas (2017) for a detailed discussion of the specific conditions of each form of exclusivity. Thomas notes: "For many firms the distinction between a data exclusivity and marketing exclusivity may be more apparent than real. The expense of generating clinical data and other information needed to obtain marketing approval from the FDA is prohibitive for many firms."

[38] Downloaded from `https://www.accessdata.fda.gov/scripts/opdlisting/oopd/listResult.cfm`.

[39] Downloaded from `https://www.fda.gov/drugs/development-resources/pediatric-exclusivity-granted`.

5 Using the Orange Book: Calculating market exclusivity

Key Takeaways

- Patent and regulatory exclusivity addenda can be used to construct various measures of market exclusivity for small-molecule drugs. The appropriate strategy may depend on the research question at hand.
- A case study of Eli Lilly's Gemzar illustrates the subtleties of overlapping patents and regulatory exclusivities.

The Orange Book's linkage between products and both patents and exclusivities provides a unique opportunity to measure and analyze features of the biomedical innovation ecosystem. Table 4 describes common, useful linkages for researchers using Orange Book records, though this summary is neither exhaustive nor definitive.

For every small-molecule drug approved by the FDA after 1985, one can use regulatory exclusivity and patent addenda in the Orange Book to measure the effective exclusivity period.[40] Here, we describe the patent and exclusivity data included in the Orange Book in more detail and note several features that need to be taken into account when working with these records. We then provide one example illustrating the ways in which researchers can make use of Orange Book resources to compute a drug's period of market exclusivity. Finally, we investigate more systematically how well various Orange Book measures capture a drug's experienced exclusivity period.

5.1 Additional details on Orange Book exclusivity and patent addenda

5.1.1 Exclusivity addendum

As noted above, small-molecule drugs approved by the FDA may receive various forms of regulatory protection in addition to patent protection. We describe any FDA-assigned exclusivity that bars entry of a generic competitor as a form of "generic-blocking exclusivity."

Across editions of the Orange Book, a total of 1,054 unique exclusivity codes appear in the exclusivity addendum, which are sorted into sixteen categories by the FDA and summarized in Table 5. Of these sixteen categories, five confer generic-blocking exclusivity in the sense defined above:[41]

1. New chemical entity (NCE) exclusivity under the Hatch-Waxman Act and 21 C.F.R. § 314.108(b)(2)
   - Granted to drugs with new active ingredients.[42]
   - Protects the active ingredient and runs concurrently with other patents/exclusivities (with exceptions noted below).
   - Duration: five years.[43]

[40] Note that the 1985 edition of the Orange Book contains only those patent and exclusivities that were in force at the time of publication. This implies that Orange Book records will not indicate whether a particular drug was covered by a patent or exclusivity that expired before 1985.

[41] Table 5 describes a sixth category of generic-blocking exclusivity: competitive generic therapy (CGT). The FDA Reauthorization Act of 2017 allows the FDA to designate a drug with "inadequate generic competition" as a CGT. Subsequent generic entrants may be eligible for expedited FDA review, as well as a 180-day exclusivity period upon ANDA approval.

[42] The Ensuring Innovation Act of April 2021 amended the statutory language from "active ingredient" to "active moiety," codifying the FDA's interpretation of "active ingredient." Historically, although statutory language referred to "active ingredients," interpretations of this term differed. Compounds in drugs may undergo chemical reactions in the body in order to induce a therapeutic effect. Alternatively, several drugs with the same central compound may have different compounds linked to them in ways that are not always clinically significant. "Active moiety" refers to the core molecule or ion of a drug that is "responsible for the physiological or pharmacological action of a drug's substance." Additional details are available here: `https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=314.3`.

[43] A prospective generic entrant is allowed to file an ANDA with a Paragraph IV certification after four years. If the challenge is resolved quickly, generic entry prior to the 5-year mark is possible.

2. Orphan drug exclusivity (ODE) under the 1983 Orphan Drug Act and 21 C.F.R. § 316.31

    - Granted to drugs that either (i) treat diseases for which the relevant U.S. patient population is under 200,000 or (ii) involve research and development expenses that the firm does not reasonably expect to recover.
    - Protects the drug when it is prescribed to treat a qualifying disease or condition and runs concurrently with other patents/exclusivities (with exceptions noted below).
    - Duration: seven years.

3. Pediatric exclusivity (PED) under Section 505A of the FDA Modernization Act of 1997 and the Best Pharmaceuticals for Children Act (BPCA)

    - Granted to drugs that have been used in pediatric studies, following a written request by the FDA.
    - Extends existing exclusivity and unexpired patents. If a drug has been granted NCE, PED is added as an extension. If a drug has ODE, PED is added as an extension.
    - Duration: six months.

4. Generating Antibiotic Incentives Now (GAIN) exclusivity under Title VIII of the FDA Safety and Innovation Act (FDASIA)

    - Granted to products that have received a Qualified Infectious Disease Product designation (with some exceptions).
    - Extends existing exclusivity and does *not* extend any patents. If a drug has been granted NCE, GAIN is added as an extension. If a drug has ODE, GAIN is added as an extension.
    - Duration: five years.

5. 180-Day Exclusivity (also known as Generic Drug Exclusivity) under 21 U.S.C. § 355(j)(5)(B)(iv)

    - Granted to a (first) generic applicant submitting a Paragraph IV patent challenge and thereby risking an infringement lawsuit.
    - Protects against entry of other generic applicants that have filed Paragraph IV challenges.
    - Duration: 180 days from commercial marketing of the generic drug (subject to forfeiture if, for example, the generic applicant fails to market the drug within 75 days of a favorable court decision).

Note that we do not designate various "new clinical investigation" (NCI) exclusivities as blocking generic entry. Although these forms of protection provide three-year periods of exclusivity to incentivize new clinical research on existing drugs (*e.g.*, a new dosage schedule, clinical indication, or patient population),[44] generic competitors may be approved for any uses not covered by the exclusivity. In practice, these restrictions leave drugs vulnerable to generic competition.

Researchers wishing to calculate a drug's "legal protection period" empirically should be attentive to several details. First, not all regulatory exclusivities apply to every product within an FDA application. For example, NCI exclusivities may apply to one product form of a drug (*i.e.*, a specific dosage or strength), but not others. Pediatric extensions, by contrast, apply to every unexpired patent and exclusivity at the time when the extension is approved, regardless of whether every product is involved in pediatric clinical trials.

[44] In addition, the PED and GAIN extensions discussed above also apply to NCI exclusivity.

Second, most (but not all) regulatory exclusivities run concurrently. For example, if a firm receives a three-year NCI exclusivity and a seven-year ODE exclusivity period on the same date, the drug receives a total of seven years of protection. The exceptions are PED and GAIN, which run consecutively to existing FDA-assigned exclusivities. If a drug has no existing exclusivities, it receives no extensions.

Table 6 provides an example of the data structure for Rebetol, an antiviral drug.[45] Rebetol, in an oral solution dosage form, received FDA approval on 29 July 2003. The Orange Book lists an NCI exclusivity set to expire three years later, on 29 July 2006,[46] and ODE exclusivity set to expire on 29 July 2010. Both received PED extensions, which extended these expiration dates by six months, to 29 January 2007 and 29 January 2011, respectively.

One can construct an expiration date that takes into account PED and GAIN extensions by determining which other patents and regulatory exclusivities are in force at the time of the PED/GAIN exclusivity approval.[47]

5.1.2 Patent addendum

The Orange Book patent files list every patent record ever appearing in any Orange Book edition between 1985 and 2016. One important issue to be aware of is that patent expiration dates listed in the Orange Book data are unique within but not across years. This implies that researchers using the historical Orange Book records data need to take care to intentionally choose the conceptually appropriate expiration date for their application.

5.2 Calculating exclusivity: An example

Calculating market exclusivity can be complex, given the overlapping effects of patents and regulatory exclusivity. As an illustration, consider Gemzar, a treatment for pancreatic cancer approved by the FDA in 1996.[48] Gemzar was protected by multiple patents and regulatory exclusivities, which we consider in turn.

Patents The data file contains two listed patents: 4,808,614 and 5,464,826. The '614 patent, covering Gemzar's active ingredient (gemcitabine hydrochloride), was granted on 28 February 1989. The '826 patent, covering a method of use, was granted on 07 November 1995. Both patents were originally set to expire 17 years from issuance: 28 February 2006 and 07 November 2012, respectively. As shown in Table 7, both patents first appear in the 1997 edition of the Orange Book with their original expiration dates.

Later, the '614 patent received a PTE of 1,537 days,[49] resulting in a new expiration date of 15 May 2010. Table 7 shows that this extension first appeared in the 2000 edition of the Orange Book. Both patents subsequently received a six-month pediatric extension, resulting in effective expiration dates of 15 November 2010 (for the '614 patent) and 07 May 2013 (for the '826 patent).[50]

First generic entry occurred on 15 November 2010,[51] upon the expiration of the '614 patent. Notably, generic entry was not blocked by the later expiring '826 patent. This example illustrates the point made earlier that not all patents are created equal. In particular, the latest-expiring patent frequently does not determine when a brand-name firm's effective exclusivity period ends, given the prospect that a court may rule that it is invalid or not infringed. In general, in the

[45] NDA #021546 (trade name ribavirin).

[46] In particular, a new dosage form.

[47] Note that expiration dates associated with PED extensions may appear to be several days longer or shorter than six months. Statistical packages assign the number of days in a month differently, and researchers should correct for any such differences. Throughout our analysis, we confirmed that PED and GAIN extensions were accurately linked to the regulatory exclusivities that they extended.

[48] See NDA #020509.

[49] See `https://www.uspto.gov/patents/laws/patent-term-extension/patent-terms-extended-under-35-usc-156`.

[50] Because pediatric exclusivity—unlike patent term extensions—does not change the patent's actual expiration date, the new effective expiration dates are denoted by "*PED" in the Orange Book.

[51] See ANDA #079183.

language of Lemley and Shapiro (2015), patents are "probabilistic": the asserted property rights are uncertain because they are subject to challenge and litigation.

As discussed above, researchers should note that patent expiration dates change across editions of the Orange Book within a single application. In the Gemzar example, changes in patent expiration dates arose due to a Hatch-Waxman restoration extension as well as six-month pediatric exclusivity extensions, both of which extended patent expiration dates. However, changes can also occur in the opposite direction due, *e.g.*, to the non-payment of patent renewal fees.

As one descriptive exercise, we tabulate changes in Orange Book-listed patent expiration dates over time, summarized in Figure 1 as the number of months between the earliest and latest listed expiration date for every patent that appears in at least one edition of the Orange Book. This difference is 0 months for the median patent, six months at the 75th percentile of the distribution, and 165 months at the maximum (resulting from an apparent typo in the Orange Book).

Exclusivity The data file contains a number of regulatory exclusivities for Gemzar, which are summarized in Table 8. Most notably, Gemzar received protection as a new chemical entity thanks to its novel active ingredient. As discussed above, NCE exclusivity lasts five years from approval, and thus expired in 2001, long before the expiration of Gemzar's first expiring patent. In this case, NCE exclusivity did not impact generic competition. For other drugs with short-lived or weak patents, by contrast, exclusivity might be highly relevant, in part given its comparatively ironclad (as opposed to probabilistic) nature.

The data file also reports multiple instances of NCI exclusivity, which were lengthened by pediatric extensions. As noted above, NCI exclusivity is less effective in preventing generic competition. All NCI exclusivities expired before the first expiring patent.

Interaction of patents and exclusivity The full story of Gemzar is more complex than the foregoing summary suggests. First generic entry in November 2010 was confined to one strength of the drug: the 2-gram version. For two other strengths (1 gram and 200 mg), entry was delayed by two additional months. This result is due to a complex interaction between patents and exclusivity—in particular, between the later expiring '826 patent and generic exclusivity.[52]

As noted above, the first generic firm to file a Paragraph IV patent challenge is potentially eligible for a 180-day period of exclusivity that blocks the approval of other generics. Eligibility for generic exclusivity is considered on a strength-by-strength basis.[53] For Gemzar, Hospira was the first filer for the 2-gram strength, while Teva was the first filer for the other two strengths. By the time the '614 patent expired in November 2010, the '826 patent had been invalidated thanks to a successful lawsuit by Sun, another generic firm. One might therefore have expected immediate generic entry on all strengths. However, due to Teva's continued entitlement to the 180-day exclusivity period, approval of other generic products was blocked as to the 1-gram and 200-mg strengths. Moreover, Teva was not obliged to immediately launch its product, and did not do so. Teva waited until January 2011 to launch a generic version of the 1 gram and 200 mg strengths, with other generic competitors blocked by the 180-day exclusivity period until July 2011.[54]

As this example illustrates, the combined effect of patents and regulatory exclusivity is sometimes subtle. Here, the existence of a weak, late-expiring patent provided an opportunity for generic exclusivity to be put at play, ultimately delaying the onset of generic entry.

[52]The details in this discussion are drawn from searching the Drugs@FDA database for "gemcitabine hydrochloride," Eli Lilly's 10-Q SEC filings during the litigation, and the judicial opinions reported at *Sun Pharm. Indus. v. Eli Lilly & Co.*, 611 F.3d 1381 (Fed. Cir. 2010), *Eli Lilly & Co. v. Sicor Pharms.*, 705 F. Supp. 2d 971 (S.D. Ind. 2010), and *Eli Lilly & Co. v. Mayne Pharma*, 504 F. Supp. 2d 387 (S.D. Ind. 2007).

[53]*Apotex v. Shalala*, 53 F. Supp. 2d 454 (D.D.C. 1999), *aff'd*, No. 99-5231, 1999 WL 956686 (D.C. Cir. Oct. 8, 1999).

[54]Under the applicable statute, Teva was required to launch its product within 75 days of a final judgment of patent invalidity, or else forfeit its entitlement to generic exclusivity. In the end, final judgment was entered on 12 November 2010, and Teva launched just before forfeiture would have occurred.

5.3 Combining records to measure exclusivity

Building on the data described in Sections 5.1.1 and 5.1.2, we investigate how well combinations of these measures actually capture a drug's experienced exclusivity period as measured by observed generic approval.

We identify the set of drugs with generic approvals by linking Orange Book records with FDA administrative data recording drug approval dates as well as approval dates for generic competitors.[55] To overlap with the coverage of our Orange Book data, we focus on drugs approved by the FDA between 1985 and 2014.[56] We focus on a set of 1,673 NDAs that correspond to the first approval of a unique set of active ingredients by the FDA.[57] Any ANDA approved by the FDA for a given set of active ingredients is designated, for the purposes of this exercise, as a generic competitor.[58] Especially for drugs with small markets and products that are challenging to reverse-engineer, delays in generic entry are common (Hemphill and Sampat, 2012). A substantial number of drugs never experience generic entry even after expiration of all exclusivities, and these drugs are not captured by this exercise.[59]

Using this data combined with the Orange Book data, we construct four measures of nominal exclusivity to compare with actual time to generic approval.[60] Within our sample, the median drug receives 13 years of market exclusivity—time before either all legal protections expire or a generic competitor enters (whichever comes first).[61]

Measure #1: NCE exclusivity. We first compare the timing of generic approval to the expiration of NCE exclusivity. Because NCE protection is granted and enforced by the FDA—in contrast to patents, which are granted by the USPTO and challenged in courts—we expect that it should bind, in the sense of fully blocking generic approval. Figure 2 illustrates how we can test this idea on the 277 drugs in our sample that both have NCE exclusivity and any generic approval. Figure 2 plots the number of years between a drug's approval and NCE expiration (y-axis) against the number of years until the drug experiences generic approval (x-axis). Any observations below the 45^o line would represent cases in which generic approval occurred after NCE expiration. We include two-year bands around expiration.

Of these 277 drugs, 18 have an exclusivity period less than five years. These are the product of the FDA's "umbrella" policy, under which NCE exclusivity applies not only to the first drug eligible for exclusivity, but also to later-approved drugs containing the same active ingredient. The later-approved drugs are protected for the remainder of the five-year period, measured from the approval of the first drug.[62]

In addition, 29 drugs have an exclusivity period longer than five years. 26 of these correspond to drugs approved between 1 January 1982 and 24 September 1984, the effective date of the Hatch-Waxman Act. Under a now-repealed

[55] Specifically, we use Drugs@FDA records which are available for download from: `https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-data-files`.

[56] We aggregate all products approved under the same NDA into a single observation.

[57] This set includes NMEs – i.e., drugs with a novel active ingredient – as well as new combinations of previously approved active ingredients.

[58] Actual substitution of generic drugs is more complicated. State law and private payers facilitate substitution by therapeutically equivalent generics, which as discussed in Section 2 requires a showing of bioequivalence. See the preface to the Orange Book for a more detailed discussion: `https://www.fda.gov/drugs/development-approval-process-drugs/orange-book-preface`. The concept of equivalence employed here is less restrictive.

[59] The FDA maintains a list of off-patent, off-exclusivity drugs without an approved generic at `https://www.fda.gov/drugs/abbreviated-new-drug-application-anda/list-patent-exclusivity-drugs-without-approved-generic`.

[60] Note that Hemphill and Sampat (2012) find that the correlation between time to first generic approval and time to first generic launch, for a set of 112 new molecular entities, is 0.97.

[61] This estimate is consistent with evidence reported in Kesselheim, Sinha, and Avorn (2017), who find that most new drugs receive between 12 and 16 years of market exclusivity.

[62] For example, NDA #019537, covering the antibiotic Cipro (ciprofloxacin) in a tablet form, was approved on 22 October 1987. Its five-year NCE exclusivity expired on 22 October 1992. NDA #019847, for Cipro in an injectable form, was approved on 26 December 1990, and shares the same NCE expiration date (22 October 1992). The umbrella policy also applies to fixed-dose combination drugs that contain the same active ingredient. See `https://www.fda.gov/files/drugs/published/New-Chemical-Entity-Exclusivity-Determinations-for-Certain-Fixed-Combination-Drug-Products.pdf`.

provision of the Act, these drugs received 10-year "transitional" exclusivity terms.[63] Two more drugs received pediatric extensions, which lengthened the duration of NCE exclusivity to five-and-a-half years.[64] For the remaining drug, the expiration date appears to be an Orange Book error.[65]

As expected, we observe no generic approval before NCE expiration. As such, we interpret Measure #1 as a lower bound on a drug's exclusivity period.

Measure #2: NCE, ODE, GAIN, and PED exclusivities. Next, we compare the timing of generic approval to the latest of the following exclusivities: NCE, ODE, and GAIN, with any associated PED extensions (see Section 5.1.1 for details). As was the case for NCE exclusivity, we expect to observe no generic approval before exclusivity expiration, as these exclusivity forms should fully bar generic approval.

Figure 3 plots time to exclusivity expiration against time to generic approval. Any observations below the 45^o line would represent cases in which generic approval occurred after exclusivity expiration. We once again include two-year bands around expiration; 13% of observations fall within the bands. Nearly all generic approval occurs after the expiration of generic-blocking exclusivities. In 8 cases in which generic approval occurs before exclusivity expiration, at least one product under the drug's NDA was not protected by the relevant exclusivity. In many cases, these observations correspond to drugs that were available for use before the passage of the Hatch Waxman Act (*e.g.*, atenolol, approved in 1981) but which received protections as "new" drugs following the introduction of regulatory exclusivities.[66]

Of the 280 drugs with at least one exclusivity in this set and observed generic approval, there are four cases for which the exclusivity period appears to exceed 15 years. In each of these cases, the drug received ODE. For example, Novantrone was approved on 23 December 1987;[67] however, it received ODE in 2000 that expired on 13 October 2007. This appears as a 20-year exclusivity term in our data.

Figure 3 suggests that—in nearly all cases—this set of exclusivities does, in fact, block generic approval. We interpret this measure as a small extension of the "lower-bound" in Measure #1.

Measure #3: Earliest patent expiration. We next compare the timing of generic approval to the latest of (i) all generic-blocking exclusivities assigned by the FDA and (ii) the earliest expiring Orange Book patent. Generic-blocking regulatory exclusivities, as in the case explored in Measure #1, should fully protect against generic entry. Existing literature suggests that early expiring patents on drugs may provide stronger property rights than those filed later in a drug's lifecycle, as they protect "primary" features of the technology (Hemphill and Sampat, 2011; Branstetter, Chatterjee, and Higgins, 2016; Gupta, 2021). We can construct this measure for the 454 drugs in our sample period that both have at least one Orange Book patent and experience generic approval.

Note, however, that the earliest expiring patent is not necessarily the patent that protects a drug's active ingredient. Information on the active ingredient ("primary") patent is not reported in the Orange Book. Some teams, including Kapczynski, Park, and Sampat (2012) and Hemphill and Sampat (2012), have reviewed patent texts by hand to identify active ingredient patents. Similar efforts, perhaps drawing on novel machine learning and natural language processing tools, may allow researchers to integrate such information into their analyses.

Figure 4 plots time to expiration against time to approval. 46% of observations fall within the two-year bands. Again, we observe a set of drugs with long nominal exclusivity periods (based on the combination of regulatory exclusivities and

[63] This provision, 21 U.S.C. § 355(j)(5)(F)(i), was repealed in 2021.

[64] See NDAs #018831 and #020192.

[65] Zocor (NDA #019766) was approved on 23 December 1991 and appears, first, in the 2002 Orange Book, with a listed NCE exclusivity expiration date of 23 December 1997. Approval documents from the FDA confirm that NCE exclusivity expired in this case in 1996.

[66] See NDA #018240.

[67] See NDA #019297.

the earliest-expiring patent); specifically, 35 of the 454 observations in this case have exclusivity periods that exceed 20 years. In 26 cases, the drug was approved before Hatch-Waxman's effective date (*i.e.*, approved before September 1984). These drugs received patents, in some cases, decades after products were initially approved, resulting in what appear to be extremely long exclusivity periods. For the remaining nine cases, we observe a similar pattern in which initial patents were not obtained until many years after the drug had been approved. Consider NDA #018613, for Ovide. The drug was approved on 02 August 1982, but the earliest record of an Orange Book patent appears in 2010 for U.S. Patent No. 7,560,445, which expires on 01 February 2027.

Figure 4 suggests that generic approval is common both before and after earliest patent expiration. When generic approval occurs after this expiration date, this may be due to a combination of secondary patents (see Measure #4 below) and market forces that delay entry.

Measure #4: Latest patent expiration. Finally, we compare the timing of generic approval to the latest of all generic-blocking exclusivities assigned by the FDA and the latest expiring patent in any Orange Book edition. To the extent that late-expiring patents are more likely to either be invalidated through litigation or unenforced, we expect to observe generic approval before market exclusivity expiration in more cases here than with measure #3.

In Figure 5, for a sample of 454 drugs, we observe that roughly half of drugs in our sample have generics approved before exclusivity expiration, as defined using the latest expiring patent. 34% of observations fall within the two-year bands around expiration. This is consistent with the idea that late-expiring patents are less effective at delaying generic approval and aligns with the findings in Hemphill and Sampat (2012), in which the authors show that, especially for high-sales drugs, late-expiring secondary patents are disproportionately targeted in litigation. When cases are litigated to completion, generics typically prevail.

In Figure 6, we replicate Figure 5 and highlight with red triangular markers any drug for which, as of June 2020, the FDA has recorded at least one Paragraph IV challenge. Note that Paragraph IV challenges appear to explain many of the cases in which generic approval is "early" as compared to the date of last patent expiration. Weakly, this is consistent with the idea that our measure #4 is a reasonable "upper bound" on the duration of a drug's exclusivity period.

Taken together, these four measures make clear that researchers using the Orange Book exclusivity and patent measures must be intentional in their decisions about how to aggregate available data into effective exclusivity measures. Beyond the measures available in the Orange Book itself, distinguishing between primary and secondary patents may be important, and the effects of these patents on generic approval may vary depending on the incentives firms face to initiate patent challenges, including projected sales.

6 Conclusion

The pharmaceutical industry is a unique context where, due to FDA regulatory requirements, it is possible to link patents to products. Since 1985, the FDA's Orange Book has listed patent and exclusivity information for small-molecule drugs approved by the FDA. In this paper, we introduce a digitized version of these records for the years 1985 to 2016 (Appendix A describes how these data can be updated to include later years).

Like many administrative datasets, the Orange Book was not created for researchers. Nonetheless, it includes many variables—including cleaned patent numbers, FDA drug application numbers, and both brand and generic drug names—that facilitate linkages with other datasets that are commonly used in empirical analyses.

Whether the data are sufficiently accurate and complete for research purposes will depend on the specific research question being asked. By design, the Orange Book focuses on small-molecule drugs and does not record information for most biological drugs. We find that, for 77% of NDAs and 96% of NMEs, the Orange Book reports at least one patent or

exclusivity. In other words, nearly all new drugs approved since 1985 have some form of patent or exclusivity listed in the Orange Book.

Conditional on a drug appearing in the Orange Book, how complete are patent records? There is minimal review associated with the listing process; the FDA reports information provided by firms marketing brand-name drugs. As we note, rules associated with Hatch-Waxman challenges and litigation create incentives for firms to report all eligible Orange Book patents (covering active ingredients, formulations, and methods of use). Comparison to a commercial dataset, IQVIA/Ark, indicates that there are many patents associated with specific drugs that are not listed; our review, however, suggests that these are mainly from categories not eligible for Orange Book listing. Of patents that IQVIA/Ark suggests should "constrain" generic entry, roughly three-quarters are listed in the Orange Book. Litigation records indicate that more than 90 percent of patents asserted in litigation against potential generic entrants are listed in the Orange Book. Additionally, 96 percent of patents that firms select for term extension—arguably, a revealed preference measure of which patents matter to the firm—are listed in the Orange Book. Taken together, we interpret these results as suggesting that the Orange Book does seem to include nearly all of the "important" patents on listed drugs.

We also find that patent expiration dates listed in the Orange Book are generally accurate and reflect changes to patent terms due to TRIPS; Orange Book expiration dates coincide with expiration dates in patent term extension letters in roughly 90 percent of cases and also reflect patent term adjustment in the majority of cases. While firms do seem to report these extensions to patent terms diligently, they do not appear to provide updates on maintenance fee non-payment. This is an important note for researchers, as nearly half of Orange Book listed patents are not maintained to full term.

Patent numbers reported in the Orange Book are accurate and, as such, can be merged with a variety of other patent datasets (including the USPTO's PatentsView and the OECD Triadic Patent Dataset). NDA numbers match cleanly to Drugs@FDA, which can be used to link drug-specific information. Other fields, including sponsoring firm name, are less standardized. Brand and generic drug names are available from the Orange Book but, in our experience, direct string matches against other datasets perform poorly and typically require manual matching procedures to improve accuracy.[68]

In addition to patent information, we also examine the completeness of Orange Book exclusivity information. As compared to other FDA sources, Orange Book records are both complete and accurate. This is unsurprising: unlike patent information, which is provided and updated by firms, exclusivity information is drawn directly from FDA records.

Our final exercise—combining patent and exclusivity data to calculate legal exclusivity periods—is intended to capture the complexities associated with using these data and to highlight the types of choices that researchers must make. Especially given concerns about the proliferation of irrelevant or invalid drug patents, the gap between "nominal" and "actual" legal exclusivity may continue to grow. While the Orange Book dataset provides a rich resource for understanding the pharmaceutical market, researchers will need to make choices based on specific questions of interest and, in some cases, to bring in outside information when determining exactly which Orange Book patents "count."

There have been many calls to reform Orange Book listing practices to achieve various policy goals. Hemphill and Sampat (2022) suggest modifications that would prevent firms from listing irrelevant patents (which may later be the target of costly patent challeneges) and collect more information from applicants about what, exactly, listed patents are intended to cover. Such changes would likely make the Orange Book an even more valuable resource for researchers.

Moving forward, efforts to construct new product-patent linkages will likely require new statutes—as were in place when listing requirements for the FDA's Purple Book were expanded in 2020. New authority would also put to rest any argument that this information should be treated as a trade secret. Settings where listing requirements can be tied to some corresponding benefit, as in the case of other products that require premarket clearance, are likely to be the most straightforward to implement.

[68] Both firm names and drug names can be obtained from Drugs@FDA by merging on the NDA number.

References

Ark Patent Intelligence. 2019. "Constraint Analysis: Introduction of Concepts and Analysis."

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat. 2019. "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules." *The Review of Economic Studies* 86 (1):112–152.

Branstetter, Lee G., Chirantan Chatterjee, and Matthew Higgins. 2016. "Regulation and Welfare: Evidence from Paragraph IV Generic Entry in the Pharmaceutical Industry." *RAND Journal of Economics* 47 (4):857–890.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh. 2000. "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)." *NBER Working Paper # 7552* .

de Rassenfosse, Gaetan. 2018. "Notice Failure Revisited: Evidence on the Use of Virtual Patent Marking." *NBER Working Paper # 24288* .

Eisenberg, Rebecca. 2012. "Patents and Regulatory Exclusivity." In *Oxford Handbook on the Economics of the Biopharmaceutical Industry*. Oxford: Oxford University Press.

Frakes, Michael D. and Melissa F. Wasserman. 2022. "Investing in Ex Ante Regulation: Evidence from Pharmaceutical Patent Examination." *American Economic Journal: Economic Policy* .

Gupta, Charu N. 2021. "One Product, Many Patents: Imperfect Intellectual Property Rights in the Pharmaceutical Industry." *Working Paper* .

Hemphill, C. Scott and Bhaven N. Sampat. 2011. "When Do Generics Challenge Drug Patents?" *Journal of Empirical Legal Studies* 8 (4):613–649.

———. 2012. "Evergreening, Patent Challenges, and Effective Market Life in Pharmaceuticals." *Journal of Health Economics* 31 (2):327–339.

———. 2013. "Drug Patents at the Supreme Court." *Science* 339 (6126):1386–1387.

Kapczynski, Amy, Chan Park, and Bhaven N. Sampat. 2012. "Polymorphs and Prodrugs and Salts (Oh My!): An Empirical Analysis of 'Secondary' Pharmaceutical Patents." *PLOS ONE* 7 (12). URL `https://papers.ssrn.com/abstract=2191728`.

Kesselheim, Aaron S., Michael S. Sinha, and Jerry Avorn. 2017. "Determinants of Market Exclusivity for Prescription Drugs in the United States." *JAMA Internal Medicine* 177 (11):1658. URL `http://archinte.jamanetwork.com/article.aspx?doi=10.1001/jamainternmed.2017.4329`.

Kyle, Margaret, Bhaven N. Sampat, and Kenneth C. Shadlen. 2020. "TRIPS, Patents, and Drug Prices in India." *Working Paper* .

Lemley, Mark A. and Carl Shapiro. 2015. "Probabilistic Patents." *Journal of Economic Perspectives* 19:75–98.

Lemley, Mark A. and Sean Tu. 2022. "What Litigators Can Teach the Patent Office About Pharamaceutical Patents." *Washington University Law Review* 99.

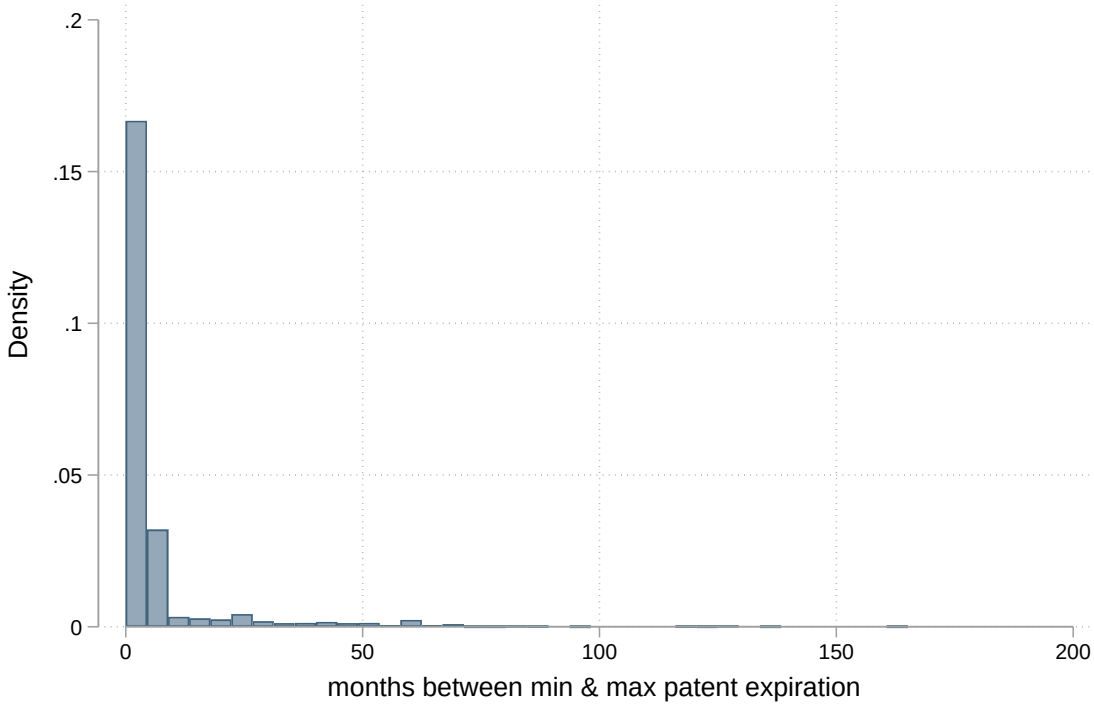
Levin, Richard, Alvin Klevorick, Richard Nelson, Sidney Winter, Richard Gilbert, and Zvi Griliches. 1987. "Appropriating the returns from industrial research and development." *Brookings Papers on Economic Activity* (3):783–831.

Lietzan, Erika and Kristina Lybecker. 2020. "Distorted Drug Patents." *Washington Law Review* 95.

Sampat, Bhaven N. and Kenneth C. Shadlen. 2015. "TRIPS Implementation and Secondary Pharmaceutical Patenting in Brazil and India." *Studies in Comparative International Development* 50 (2):228–257. URL http://link.springer.com/10.1007/s12116-015-9181-7.

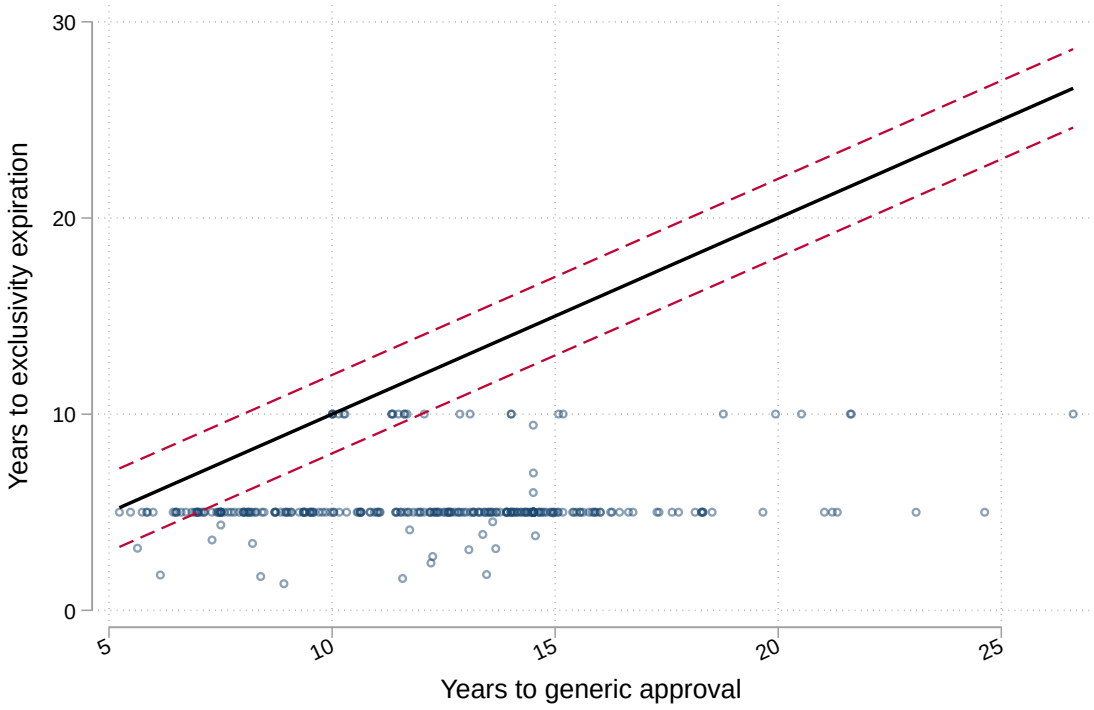
Thomas, John R. 2017. "Regulatory Exclusivity Reform in the 115th Congress." Report R44951, Congressional Research Service.

Figure 1: Changes in Orange Book-listed patent expiration dates across editions



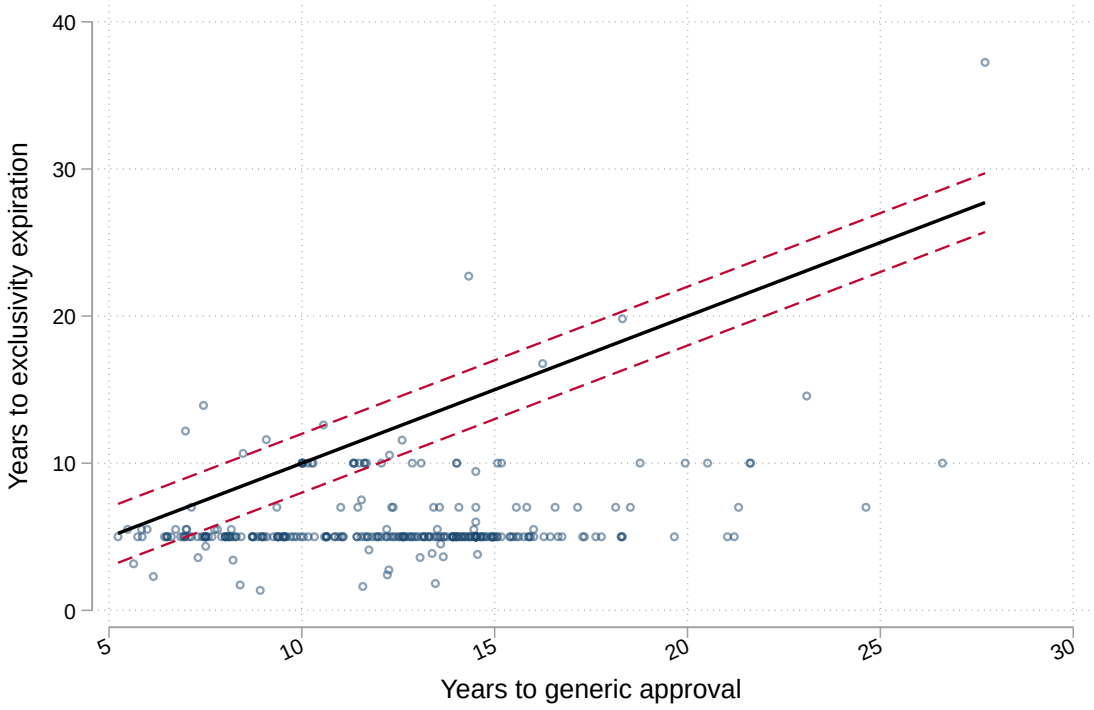
Notes: Figure plots the number of months between the earliest and latest expiration dates listed for each patent in the Orange Book, based on records reported in the Orange Book patent addendum. $N = 5,511$.

Figure 2: NCE exclusivity vs. generic approval



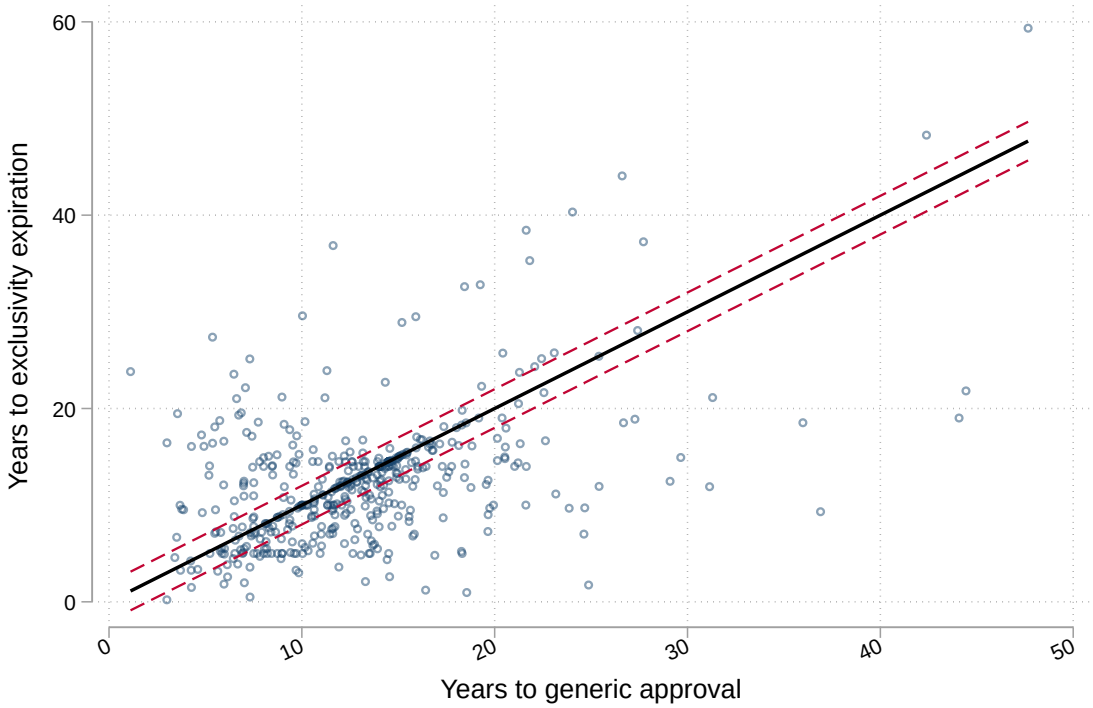
Notes: Figure plots the relationship between approval of a generic competitor and the following measure of exclusivity: NCE exclusivity, assigned by the FDA. Approval of a generic competitor is measured using Drugs@FDA administrative data. NCE is reported in the Orange Book exclusivity addendum. Both the x- and y-axes are measured as time since the drug's initial approval. The 45^o line is plotted in black, with two-year bands plotted as red dotted lines. An observation above the 45^o line represents a drug for which a generic competitor was approved before exclusivity expired. The sample includes drugs with both NCE exclusivity and observed generic approval ($N = 277$).

Figure 3: Generic-blocking regulatory exclusivities vs. generic approval



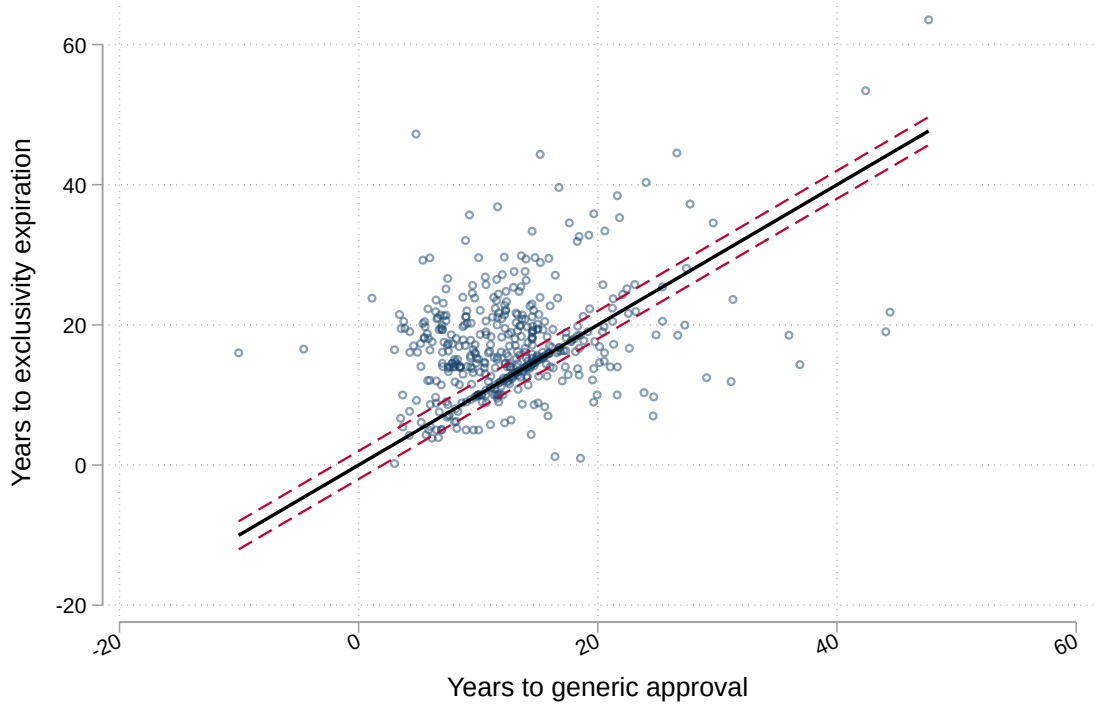
Notes: Figure plots the relationship between approval of a generic competitor and the following measure of exclusivity: all generic-blocking exclusivities assigned by the FDA. Approval of a generic competitor is measured using Drugs@FDA administrative data. Generic-blocking exclusivities are reported in the Orange Book exclusivity addendum. Both the x- and y-axes are measured as time since the drug's initial approval. The 45^o line is plotted in black, with two-year bands plotted as red dotted lines. An observation above the 45^o line represents a drug for which a generic competitor was approved before exclusivity expired. The sample includes drugs with at least one generic-blocking exclusivity and observed generic approval ($N = 280$).

Figure 4: Earliest expiring patent vs. generic approval



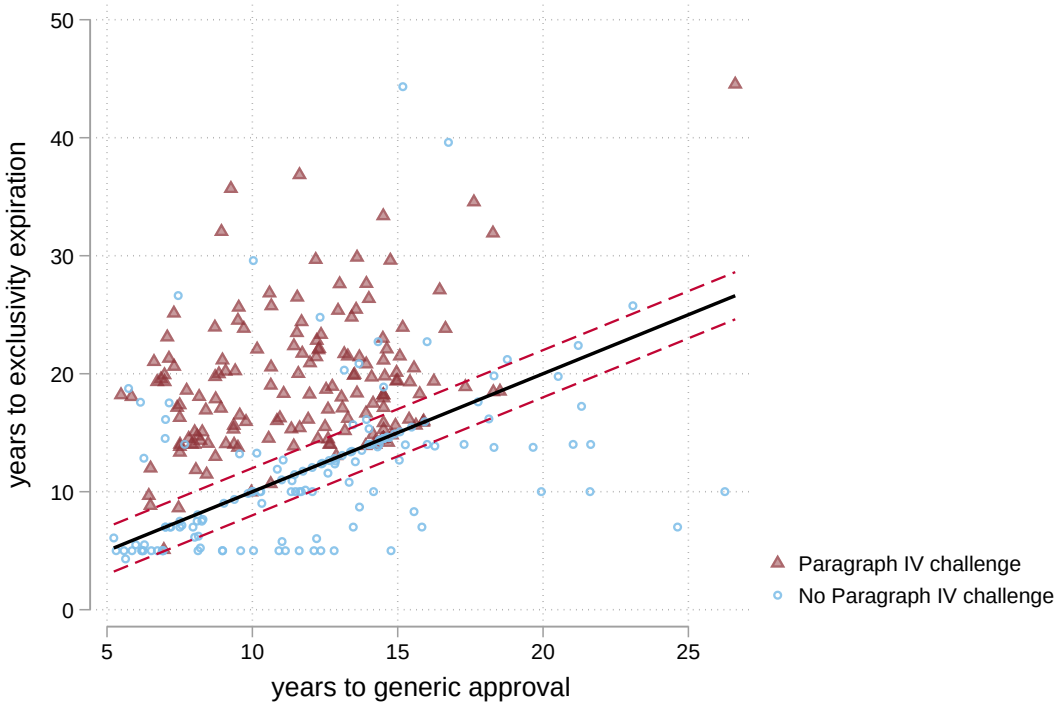
Notes: Figure plots the relationship between approval of a generic competitor and the following measure of exclusivity: the latest of all generic-blocking exclusivity expiration dates and the first expiring patent on a particular drug. Approval of a generic competitor is measured using Drugs@FDA administrative data. Exclusivity and patent records are reported in Orange Book addenda. Both the x- and y-axes are measured as time since the drug's initial approval. The 45^o line is plotted in black, with two-year bands plotted as red dotted lines. An observation above the 45^o line represents a drug for which a generic competitor was approved before exclusivity expired. The sample includes drugs with at least one Orange Book patent and observed generic approval ($N = 454$).

Figure 5: Latest expiring patent vs. generic approval



Notes: Figure plots the relationship between approval of a generic competitor and the following measure of exclusivity: the latest of all generic-blocking exclusivity expiration dates and the last expiring patent on a particular drug. Approval of a generic competitor is measured using Drugs@FDA administrative data. Exclusivity and patent records are reported in Orange Book addenda. Both the x- and y-axes are measured as time since the drug's initial approval. The 45^o line is plotted in black, with two-year bands plotted as red dotted lines. An observation above the 45^o line represents a drug for which a generic competitor was approved before exclusivity expired. The sample includes drugs with at least one Orange Book patent and observed generic approval ($N = 454$).

Figure 6: Generic approval & Paragraph IV challenges



Notes: Figure plots the relationship between approval of a generic competitor and the following measure of exclusivity: the latest of all generic-blocking exclusivity expiration dates and the last expiring patent on a particular drug. Drugs with competitors that entered via Paragraph IV challenge are marked with red triangles. Both the x- and y-axes are measured as time since the drug's initial approval. The 45^o line is plotted in black, with two-year bands plotted as red dotted lines. An observation above the 45^o line represents a drug for which a generic competitor was approved before exclusivity expired. The sample includes drugs with at least one Orange Book patent and observed generic approval ($N = 454$).

Table 1: Patent Data Validation Exercises

Validation Exercise	Sample	Findings
Patent Coverage		
Comparison to IQVIA/Ark	NMEs approved between 1995 and 2017	76% of IQVIA/Ark "constraining" patents listed in OB
Comparison to ParagraphFour.com	Cases resolved (adjudicated or settled) since 1 November 2003	91% of litigated patents listed in OB
Comparison to Hatch-Waxman patent term extensions (PTE)	Drugs with PTE granted as of 2012	96% of drugs had extended patents listed in OB
Expiration Date Accuracy		
Inspection of records affected by TRIPS	Records in OB editions 1994-1995 and 1995-1996	Only one change in OB expiration date inconsistent with TRIPS
Inspection of patents with PTE	Patents granted PTE as of 2016	89% of patents have identical expiration in OB and PTO records
Inspection of patents with PTA	Patents granted PTA as of 2014	OB records are accurate for majority of cases; see Section 4.3 for details
Comparison to maintenance fee payment dataset	Patents with maintenance fee data as of 2022	46% of OB patents expire early; OB does not capture early expiration

Table 2: Discrepancies Between USPTO Patent Term Extension and Orange Book Records

Issue	Affected Patents	Is OB Correct?
OB does not report PTE	3998790; 4085225; 4154839; 4670444*; 4689320; 4808605; 4898724; 4911920; 5008256; 5180668; 5234404; 5494903; 5510353; 5639443; 5418226; 5843901	NO
OB miscalculates TRIPS extension	4284647; 4244946	NO
Miscellaneous OB errors	4105783*; 4234571*; 4567264*; 4585597*; *4687659**; *4845075**; 4990517*; 5387603*; 5446194*	NO
USPTO miscalculates TRIPS extension	3962432; 3927046; 3985758; 4001323; 4014986; 4056635	YES
USPTO reports only interim PTE, not full PTE	4977138; 5254556; 5407914	YES
Patent included on USPTO list, but no listed extended expiration date	4695578; 4762856; 5196444	YES
Terminal disclaimer omitted from PTE term calculation	4886812*; 5656667*	YES
USPTO does not reflect a Certificate of Correction that amended the term extension	4695590*	YES
Miscellaneous USPTO errors	3857952*; 4661491*	YES
Drug approved under a BLA and subsequently reclassified	5681818*	N/A

Notes: Table reports discrepancies between patent expiration dates reported in USPTO records (USPTO Patent Center and Hatch-Waxman extension list) and dates listed in the Orange Book. Hatch-Waxman extensions are drawn from a 18 January 2016 copy of the patent term extension list downloaded from the USPTO website (obtained using Internet Archive's Wayback Machine). For each patent appearing in multiple editions of the Orange Book, we use the latest applicable edition. For two patents (italicized), both OB and USPTO records are incorrect. * denotes patents for which we report additional details on discrepancies in Appendix B.1.

Table 3: Discrepancies Between USPTO Patent Term Adjustments and Orange Book Records

Issue	Affected Patents	Is OB Correct?
OB does not reflect Certificate of Correction	8460704; 8470347; 8476010; *8545884*; 8637512; *8664215*; 8748573; *8802142*	NO
Maintenance fee non-payment	6923983; 7387793; 7405223; 7686786	NO
OB uses incorrect filing date to compute term	6669948	NO
USPTO calculates expiration using *actual* filing date; OB calculates expiration using *effective* filing date	6767901; 7700128; 7932268; 7964592; 8003673; 8168616; 8193196; 8318802; 8324189; 8410131; 8415345; 8653061; 8746242; 8759350; 8790641; 8846650	YES
Terminal disclaimer only recorded in OB	6960577; 7144861; 7320999; 7348362; 7384980; 7419973; 7452874; 7566705; 7683037; 7709444; 7816379; 7838027; 8672898; 8829017	YES

Notes: Table reports sources of discrepancies between patent expiration dates reported in USPTO patent term adjustment data and dates listed in the Orange Book. Patent term adjustment information is drawn from a 2014 copy of the USPTO's Historical Masterfile, downloaded from the USPTO website. Table reports sources of discrepancies for a randomly drawn 10 percent sample of all patents with different expiration dates across the two datasets ($N = 43$). For three patents (italicized), USPTO records are also incorrect because expiration is calculated using actual instead of effective filing date. Additional details are discussed in Appendix B.2.

Table 4: Common Orange Book Linkages

Dataset	OB Variable	Linked Variable	Notes for Use
Drugs@FDA	NDA number	NDA number	Multiple distinct NDAs may be associated with each OB patent, and multiple patents may be associated with each NDA.
USPTO administrative records	Patent number	Patent number	With considerable cleaning and manual matching, it may be possible to link firm names in OB data to USPTO assignee data.
PatentsView	Patent number	Patent number	
Medical Expenditure Panel Survey	NDA number	National Drug Code (NDC) number and drug name (brand-name, generic)	A crosswalk from NDC numbers to NDA numbers is available from the FDA's Structured Product Labeling system. Generic and brand-name merges are often messy and may require manual matching.
OECD Triadic Patent Families Database	Patent number	Patent number	
Hatch-Waxman patent extensions	Patent number	Patent number	
Paragraph IV challenges	NDA number or drug name	NDA number or drug name	

Notes: Table describes common linkages between the Orange Book and other datasets that capture aspects of biomedical innovation. Drugs@FDA data are available at `https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-data-files`. USPTO administrative records are available in many forms at `https://www.uspto.gov/ip-policy/economic-research/research-datasets`. PatentsView data are available at `https://patentsview.org/download/data-download-tables`. Waves of the Medical Expenditure Panel Survey are available at `https://meps.ahrq.gov/data_stats/download_data_files.jsp`. OECD Triadic Patent Families data are available at `https://data.oecd.org/rd/triadic-patent-families.htm`. A list of Hatch-Waxman patent extensions maintained by the USPTO may be downloaded from `https://www.uspto.gov/patents/laws/patent-terms-extended`. A list of Paragraph IV challenges maintained by the FDA may be downloaded from `https://www.fda.gov/drugs/abbreviated-new-drug-application-anda/patent-certifications-and-suitability-petitions`.

Table 5: Orange Book Exclusivity Code Categories

Code	**Exclusivity Type**	**Duration**	**Blocks Generic Entry**
NCE	New Chemical Entity	Five years	Yes
ODE	Orphan Drug Exclusivity	Seven years	Yes
PED	Pediatric Exclusivity	Six months	Yes
GAIN	Generating Antibiotic Incentives Now	Five years	Yes
PC*	Patent Challenge	180 days	Yes
CGT*	Competitive Generic Therapy	180 days	Yes
D	New Dosage Schedule	Three years	No
I	New Indication	Three years	No
M	Miscellaneous	Three years	No
NC	New Combination	Three years	No
NDF	New Dosage Form	Three years	No
NE	New Ester / Salt	Three years	No
NP	New Product	Three years	No
NPP	New Patient Population	Three years	No
NR	New Route	Three years	No
NS	New Strength	Three years	No

Table 6: NDA #021546 (Rebetol/Ribavirin), PED extensions

Application Number	Exclusivity Type	Exclusivity Expiration	Linked PED Extension	Corrected Expiration
021546	NDF	29-Jul-2006	yes	29-Jan-2007
021546	PED	29-Jan-2007	—	—
021546	ODE	29-Jul-2010	yes	29-Jan-2011
021546	PED	29-Jan-2011	—	—

Notes: Table reports pediatric exclusivity extensions associated with NDA #021546 (Rebetol/Ribavirin). Columns 1-3 report application numbers, exclusivity types, and exclusivity expiration dates for Rebetol, as they appear in the Orange Book exclusivity addendum. Column 4 flags whether a particular non-PED exclusivity has an associated pediatric extension. Column 5 lists the corrected expiration date for the non-PED extension, which matches that listed for the linked PED extension in the Orange Book.

Table 7: Orange Book patents associated with Gemzar (NDA #020509)

Patent	Edition	Expiration	Use Code
4808614	1997	28-Feb-2006	
	1998	28-Feb-2006	
	1999	28-Feb-2006	
	2000	15-May-2010	
	2001	15-May-2010	
	2002	15-May-2010	
	2003	15-May-2010	
	2004	15-May-2010	
	2005	15-Nov-2010*	
	2006	15-Nov-2010*	
	2007	15-Nov-2010*	
	2008	15-Nov-2010*	
	2009	15-Nov-2010*	
	2010	15-Nov-2010*	
	2011	15-Nov-2010*	
5464826	1997	07-Nov-2012	U-146
	1998	07-Nov-2012	U-146
	1999	07-Nov-2012	U-146
	2000	07-Nov-2012	U-146
	2001	07-Nov-2012	U-146
	2002	07-Nov-2012	U-146
	2003	07-Nov-2012	U-146
	2004	07-Nov-2012	U-146
	2005	07-May-2013*	U-146
	2006	07-May-2013*	U-146
	2007	07-May-2013*	U-146
	2008	07-May-2013*	U-146
	2009	07-May-2013*	U-146
	2010	07-May-2013*	U-146
	2011	07-May-2013*	U-146
	2012	07-May-2013*	U-146
	2013	07-May-2013*	U-146

Notes: Table shows Orange Book patents associated with Gemzar (NDA #020509). Use code U-146 indicates that the patent is for "a method of treating susceptible neoplasms in mammals." An asterisk (*) indicates that the patent expiration date was affected by a pediatric exclusivity patent term extension, as denoted by "*PED" in the Orange Book.

Table 8: Orange Book exclusivities associated with Gemzar (NDA #020509)

Exclusivity Code	Definition of Code	Exclusivity Expiration	First Edition	Last Edition
NCE	New chemical entity	15-May-2001	1997	2001
I-234	For use in combination with cisplatin for the first-line treatment of patients with inoperable locally advanced (stage IIIA or IIIB) or metastatic (stage IV) non-small cell lung cancer	26-Aug-2001	1999	2001
I-428	For use in combination with paclitaxel for the first-line treatment of patients with metastatic breast cancer after failure of prior anthracycline containing adjuvant chemotherapy unless anthracyclines were clinically contraindicated	19-May-2007	2005	2007
PED	Pediatric exclusivity	19-Nov-2007	2005	2008
M-40	Additional information regarding clinical studies performed in pediatric patients with leukemia added to precautions	26-Apr-2008	2006	2008
PED	Pediatric exclusivity	26-Oct-2008	2006	2009
I-499	Use of Gemzar in combination with carboplatin for the treatment of patients with advanced ovarian cancer that has relapsed at least six months after completion of platinum-based therapy	14-Jul-2009	2007	2010

Notes: Table shows Orange Book exclusivities associated with Gemzar (NDA #020509). First and last editions listed indicate the first and last years of Orange Book editions in which these exclusivity codes and expiration dates appeared in connection with NDA #020509. The exclusivity expiration dates for all entries are constant across editions.

A Updating data

The NBER Orange Book Data includes records from 1985 to 2016. Researchers may find it useful to append data from more recent years to these historical records. Annual editions, in PDF form, are available for download from the FDA.[69] Note that considerable processing may be necessary in order to turn these PDFs into "research-ready" data files.

Alternatively, researchers may prefer to download text files containing Orange Book patent and exclusivity records that are updated monthly. The FDA makes the most recent version of the electronic Orange Book available for download.[70] Researchers can use the Internet Archive's Wayback Machine and the FDA's own web archive, which covers October 2016 to present, to access historical versions of these monthly files. By collecting and appending "snapshots" of the database taken at regular intervals, it is possible to reconstruct the complete listing of patents and exclusivities—including records that may appear for only short periods of time before being de-listed.[71]

B PTE and PTA discrepancies

Here, we outline the process that we use to determine the correct expiration dates for Orange Book patents with PTE and PTA. Scaling up such a procedure may be valuable for future efforts to audit the expiration dates of all Orange Book patents, beyond the samples examined here.

To determine the correct expiration date in each case, we first collect the following for each patent: filing dates (actual and effective), issuance date, PTE records, and PTA records. Determining a patent's effective filing date can be complicated for child patent applications if, for example, the application in question is a continuation-in-part; researchers may need to determine effective filings dates on a case-by-case basis for such records.[72] We also confirm that there are no Certificates of Correction on file with the USPTO that alter PTE or PTA determinations.[73]

Next, we determine if maintenance fees were paid by 4, 8, or 12 years after the patent's issuance; if not, the correct expiration date is the date on which a maintenance payment was due. We also review USPTO Patent Center records to determine whether a terminal disclaimer is on file; if so, the correct expiration date is the terminal disclaimer date, plus any PTE and PTA extensions.

For the remaining cases, we determine whether the patent's actual filing date was before 08 June 1995, the date on which TRIPS changes to patent term calculation took effect. If a patent was filed before this date, its expiration date is the later of: { effective filing date + 20 years + PTA + PTE, issuance date + 17 years + PTA + PTE }. If a patent was filed after this date, its expiration date is its effective filing date plus 20 years and any associated PTA and PTE.

B.1 PTE discrepancies

We resolve 45 discrepancies in patent expiration dates between the USPTO Patent Term Extension (PTE) list and the Orange Book. Details on specific, complicated cases flagged in Table 2 are below:

[69] The most recent annual edition is available here: `https://www.fda.gov/drugs/drug-approvals-and-databases/orange-book-data-files`. The FDA.gov archive allows researchers to retrieve previous annual editions: `https://www.fda.gov/about-fda/about-website/fdagov-archive`. Research and advocacy organizations also maintain archives of annual Orange Book editions.

[70] The current version can be accessed here: `https://www.fda.gov/drugs/drug-approvals-and-databases/approved-drug-products-therapeutic-equivalence-evaluations-orange-book`.

[71] As of 05 September 2022, a Wayback Machine search for the following URL returned the FDA webpage "Approved Drug Products with Therapeutic Equivalence Evaluations (Orange Book)," which includes a link to download "Orange Book Data Files (compressed) (ZIP)", the monthly updated text file: `https://www.fda.gov/Drugs/InformationOnDrugs/ucm129662.htm`.

[72] For our small number of discrepancies, we use USPTO Patent Center records to make these determinations for each case. Scaling up this procedure may be possible using the USPTO's bulk data on continuation patents.

[73] A list of all patents with Certificates of Correction is available here: `https://www.uspto.gov/patents/search/authority-files/certificates-correction`. To the best of our knowledge, no machine-readable version of these corrections exists.

- Patent **3857952**: This patent was granted 31 December 1974. With a 17 year patent term and 2 year PTE, the expiration date should be 31 December 1993, as reported in the Orange Book. We are unable to determine how the PTE date of 17 December 1993 was obtained, as this patent is not available in the USPTO's Patent Center database. We assume that the PTE date is an error.

- Patent **4105783**: This patent has a terminal disclaimer of term after 03 May 1994. With a 541 day extension, this yields the PTE expiration date of 26 October 1995 – the correct expiration date. The Orange Book lists an expiration date of 15 January 1997, which reflects what we believe is an earlier effective filing date (which is somewhat unclear, due to the existence of a continuation-in-part) of 23 July 1975 + 20 years + 541 days = 14 January 1997. There are often one day discrepancies between the Orange Book and USPTO datasets.

- Patent **4234571**: The correct expiration date is 11 June 2001. The Orange Book lists 11 June 2011.

- Patent **4567264**: This patent received five years of PTE and 1 year of interim PTE. The four year extension in the Orange Book appears to be a typo.

- Patent **4585597**: The PTE list reports a 1,455 day term extension from 16 June 2003, which should be 16 June 2007. The OB reported date of 10 June 2007. Based on our review of available documentation, we believe the OB date is an error.

- Patent **4661491**: The PTE list miscalculates the proper extension. This patent had an initial expiration date of 27 May 2006 and received a 1,697 day extension (as confirmed by both the PTE list and a Certificate of Extension). However, the HW list adds only 365 days to the expiration. The Orange Book date correctly adds 1,697 days to the original expiration.

- Patent **4670444**: The Orange Book listing does not include PTE, but does include PED. The correct expiration date should be 09 June 2007 = 09 December 2006 + six months.

- Patent **4687659**: The PTE list reports an expiration date of 08 January 2007, which is the grant date (18 August 1987) + 17 years + 873 days of PTE. The Orange Book reports an expiration date of 04 May 2007, which we believe was obtained as follows: filing date (13 November 1984) + 17 years + 873 days = 05 April 2007; the listing then swaps month and day (05 April 2007 becomes 04 May 2007). The correct expiration date, based on our review of documentation associated with this record, is 05 April 2007.

- Patent **4695590**: A Certificate of Correction clarifies that PTE should be 712 days, instead of 122. The Orange Book is correct.

- Patent **4845075**: This patent's original expiration date was 04 July 2006, and it received a 1,578 day extension. The PTE list reports 29 October 2010 as an expiration date, and the Orange Book lists 11 April 2009. 29 October 2010 is 17 years after patent *grant* plus 1,578 days. 11 April 2009 is 18 years after patent *filing* plus 1,578 days. Based on our review of records, both 29 October 2010 and 11 April 2009 are incorrect. The Orange Book expiration date may reflect a typo. In this case, the patent's expiration date—calculated as 20 years after patent filing + PTE—should be 11 April 2011, *20* years after patent *filing* plus an extension.

- Patent **4886812**: A terminal disclaimer was filed. 08 October 2010 is the correct expiration date after the terminal disclaimer.

- Patent **4990517**: There are two different expiration dates for this patent reported in the last Orange Book edition in which it appears. One expiration date—08 December 2011 — appears to be a 10-day typo of the correct PTE date, 18 December 2011. With a 6-month pediatric extension, the correct Orange Book date should be 18 June 2012, not 08 June 2012.
- Patent **5387603**: The correct expiration date, listed in PTE data, is 12 December 2018. The Orange Book lists 01 December 2018.
- Patent **5446194**: 19 November 2013 is the correct expiration date, given the length of PTE. Based on our review of available documentation, we believe 19 October 2013 is an error in the Orange Book.
- Patent **5656667**: In 2009, a terminal disclaimer was filed, disclaiming the term after the expiration date of patent 5502077. The Orange Book expiration date is correct, based on this terminal disclaimer.
- Patent **5681818**: This patent is associated with a drug that was initially approved under a BLA. It was removed from the Orange Book before the Orange Book could reflect PTE.

B.2 PTA discrepancies

Table 3 investigates discrepancies between Orange Book and USPTO PTA records for a randomly drawn 10 percent sample ($N = 43$) of patents with different expiration dates across these two datasets. In 16 cases (38%), patent expiration dates were calculated differently in PTA data and the Orange Book: in these cases, the Orange Book reports an expiration date based on the *effective* filing date of a patent, associated with parent patents, and the USPTO reports an expiration date based on its *actual* filing date. Orange Book records are correct in these cases. In 5 cases (12%), Orange Book records—as of the 2015 edition—do not include Certificates of Correction to PTA issued by the USPTO. For three patents, both errors are present: USPTO and Orange Book data use different filing dates for patents, and Orange Book records do not record a relevant Certificate of Correction. In 14 cases (33%), only the Orange Book correctly reports an expiration date updated to reflect a terminal disclaimer. In the remaining five cases, Orange Book expiration dates are incorrect, as they either do not reflect maintenance fee non-payment (and associated patent expiration) or include a typographical error.