The NBER Orange Book Dataset
A User's Guide

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

December 2022

Linking patents to products

Patents are widely used as measures of innovation

- But mappings from **patents** to commercialized **products** are often unclear
- In general: no straightforward linkage [exception: de Rassenfosse et al. (2018)]

Linking patents to products

Patents are widely used as measures of innovation

- But mappings from **patents** to commercialized **products** are often unclear
- In general: no straightforward linkage [exception: de Rassenfosse et al. (2018)]

Pharmaceuticals in the U.S. are different

Linking patents to products

Patents are widely used as measures of innovation

- But mappings from **patents** to commercialized **products** are often unclear
- In general: no straightforward linkage [exception: de Rassenfosse et al. (2018)]

Pharmaceuticals in the U.S. are different

- Drug makers are required to report all patent protection to the FDA
- Lists are published annually as the FDA's Orange Book
  - All patents (& regulatory exclusivities) associated with brand-name, small-molecule drugs

Linking patents to products

Patents are widely used as measures of innovation

- But mappings from **patents** to commercialized **products** are often unclear
- In general: no straightforward linkage [exception: de Rassenfosse et al. (2018)]

Pharmaceuticals in the U.S. are different

- Drug makers are required to report all patent protection to the FDA
- Lists are published annually as the FDA's Orange Book
  - All patents (& regulatory exclusivities) associated with brand-name, small-molecule drugs

This paper: a user's guide to the challenges and opportunities of this unique resource

This paper

We introduce a newly digitized, open-access dataset of Orange Book records [1985-2016]

- Annual editions provide snapshots at specific points in time
- New dataset provides a comprehensive portrait of legal protections over drug lifecycles

This paper

We introduce a newly digitized, open-access dataset of Orange Book records [1985-2016]

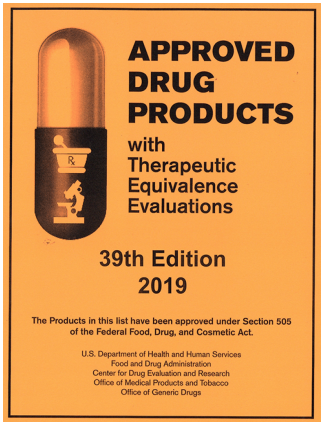
- Annual editions provide snapshots at specific points in time
- New dataset provides a comprehensive portrait of legal protections over drug lifecycles

Focus on two challenges:

1. Orange Book records are self-reported by firms and not audited by the FDA
   $\Longrightarrow$ Validation exercises against external benchmarks
2. The appropriate use of the Orange Book may differ across researchers and questions
   $\Longrightarrow$ Detailed use case: calculating market exclusivity

Outline

Background



Patents and regulatory exclusivities added under **Hatch-Waxman Act (1984)**

(Drug Price Competition and Patent Term Restoration Act)

- Modern regime for competition for brand-name and generic drugs
- Includes a variety of features intended to accelerate entry of generic drugs

Patents

Eligible: Patents with > 1 claim covering drug substances, products, and methods of use

- Must list any patents that "could reasonably be asserted" [21 U.S.C. § 355(b)(1)(A)(viii)]
- Patents issued after drug approval must be listed within 30 days of issuance

Patents

Eligible: Patents with > 1 claim covering drug substances, products, and methods of use

- Must list any patents that "could reasonably be asserted" [21 U.S.C. § 355(b)(1)(A)(viii)]
- Patents issued after drug approval must be listed within 30 days of issuance

Ineligible: Patents on manufacturing processes, packaging, metabolites, intermediates

Patents

Firms face strong incentives to list all eligible patents

- Potential generic entrants are required to *challenge* every Orange Book patent
  - Legal certification that patent is either invalid or not infringed
  - "Paragraph IV challenges" often lead to generic entry [21 U.S.C. § 355(j)(2)(A)(vii)(IV)]

Patents

Firms face strong incentives to list all eligible patents

- Potential generic entrants are required to *challenge* every Orange Book patent
  - Legal certification that patent is either invalid or not infringed
  - "Paragraph IV challenges" often lead to generic entry [21 U.S.C. § 355(j)(2)(A)(vii)(IV)]

Key feature: Infringement suits filed in response to Paragraph IV challenges

$\Longrightarrow$ Generic approval blocked for 30 months while suit is pending

Regulatory exclusivities

Exclusivities: Delays and prohibitions on approval of competitor drugs, awarded by FDA

Regulatory exclusivities

Exclusivities: Delays and prohibitions on approval of competitor drugs, awarded by FDA

Generic-blocking exclusivities

1. New chemical entity (NCE) exclusivity: **5 years**
2. Orphan drug exclusivity (ODE): **7 years**
3. Pediatric (PED) exclusivity: **6 months**
4. Generating Antibiotic Incentives Now (GAIN) exclusivity: **5 years**
5. 180-day exclusivity (Generic Drug Exclusivity): **180 days**

Outline

Data Resources

Data Resources

Four data files:

1. **FDA Drug Patents:** Full listing of patent records
   [edition, FDA application no., generic name, trade name, patent no., patent expiration]
2. **FDA Patent Use Codes**: Details on approved indications/uses [1998-present]
3. **FDA Drug Exclusivity**: Full listings of all exclusivities
4. **FDA Drug Exclusivity Codes**: Details on exclusivity types

Data Resources

Four data files:

1. **FDA Drug Patents:** Full listing of patent records
   [edition, FDA application no., generic name, trade name, patent no., patent expiration]
2. **FDA Patent Use Codes**: Details on approved indications/uses [1998-present]
3. **FDA Drug Exclusivity**: Full listings of all exclusivities
4. **FDA Drug Exclusivity Codes**: Details on exclusivity types

Dataset does not include all information in FDA Orange Book

- *e.g.*, the Products file, which identifies therapeutically equivalent alternatives for drugs
- Paper provides directions to expand coverage to more recent years

Outline

Drug coverage

Of the **2,458** new (small molecule) drugs approved between 1985 and 2015:

- $\sim$80% of drugs have at least one form of legal protection

Drug coverage

Of the **2,458** new (small molecule) drugs approved between 1985 and 2015:

- ∼80% of drugs have at least one form of legal protection

Of the **796** new molecular entities approved between 1985 and 2015:

- ∼96% of *innovative* drugs have at least one form of legal protection

Drug coverage

Of the **2,458** new (small molecule) drugs approved between 1985 and 2015:

- ∼80% of drugs have at least one form of legal protection

Of the **796** new molecular entities approved between 1985 and 2015:

- ∼96% of *innovative* drugs have at least one form of legal protection

$\Longrightarrow$ Consistent with documented importance of IP in pharmaceutical markets

[Cohen, Nelson, Walsh 2000; Levin et al. 1987]

Drug coverage

Of the **2,458** new (small molecule) drugs approved between 1985 and 2015:

- ∼80% of drugs have at least one form of legal protection

Of the **796** new molecular entities approved between 1985 and 2015:

- ∼96% of *innovative* drugs have at least one form of legal protection

$\Longrightarrow$ Consistent with documented importance of IP in pharmaceutical markets
[Cohen, Nelson, Walsh 2000; Levin et al. 1987]

* Note: does not include patent listings for biologics (some are listed in FDA's Purple Book)

Exclusivity coverage

Exclusivity is granted by the FDA and recorded directly by the agency

- Expect no errors in Orange Book records

Exclusivity coverage

Exclusivity is granted by the FDA and recorded directly by the agency

- Expect no errors in Orange Book records

Two simple validation exercises:

1. Confirm that Orange Book records of ODE match an FDA list of orphan designations
2. Confirm that Orange Book records of PED match an FDA list of pediatric studies

Exclusivity coverage

Exclusivity is granted by the FDA and recorded directly by the agency

- Expect no errors in Orange Book records

Two simple validation exercises:

1. Confirm that Orange Book records of ODE match an FDA list of orphan designations
2. Confirm that Orange Book records of PED match an FDA list of pediatric studies

$\Longrightarrow$ Exclusivity records are **complete** and **accurate**

Patent coverage

Patent coverage reflects a mix of legal requirements and firm strategy

- Expect coverage to vary across drugs and firms
- May also vary over time: documented increase in listing of "secondary patents"
  [Kapczynski, Park, and Sampat 2012; Hemphill and Sampat 2011]

Patent coverage

Three validation exercises:

1. Comparison to IQVIA/Ark [Kyle, Sampat, Shadlen 2020]
   - Many drug patents (on the order of 60%) are not listed in the Orange Book
   - But the vast majority that are "likely to constrain generic entry" are listed

Patent coverage

Three validation exercises:

1. Comparison to IQVIA/Ark [Kyle, Sampat, Shadlen 2020]
   - Many drug patents (on the order of 60%) are not listed in the Orange Book
   - But the vast majority that are "likely to constrain generic entry" are listed
2. Litigation records [ParagraphFour.com]
   - Most patents ($\sim 92\%$) litigated under "Paragraph IV" are listed

Patent coverage

Three validation exercises:

1. Comparison to IQVIA/Ark [Kyle, Sampat, Shadlen 2020]
   - Many drug patents (on the order of 60%) are not listed in the Orange Book
   - But the vast majority that are "likely to constrain generic entry" are listed
2. Litigation records [ParagraphFour.com]
   - Most patents ($\sim 92\%$) litigated under "Paragraph IV" are listed
3. Extended patents under Hatch-Waxman
   - Most patents ($\sim 96\%$) extended under H-W are listed

Patent coverage

Three validation exercises:

1. Comparison to IQVIA/Ark [Kyle, Sampat, Shadlen 2020]
   - Many drug patents (on the order of 60%) are not listed in the Orange Book
   - But the vast majority that are "likely to constrain generic entry" are listed
2. Litigation records [ParagraphFour.com]
   - Most patents ($\sim 92\%$) litigated under "Paragraph IV" are listed
3. Extended patents under Hatch-Waxman
   - Most patents ($\sim 96\%$) extended under H-W are listed

$\Longrightarrow$ Records of all "important" drug patents are **complete**

Patent expiration

Patents may not always expire 20 years after filing...

Patent expiration

Patents may not always expire 20 years after filing...

We investigate whether OB records reflect four changes to expiration dates:

1. Agreement on Trade-Related Aspects of IP Rights (TRIPS)
2. Patent Term Extension (PTE)
3. Patent Term Adjustment (PTA)
4. Maintenance Fee Non-Payment

Patent expiration

Patents may not always expire 20 years after filing...

We investigate whether OB records reflect four changes to expiration dates:

1. Agreement on Trade-Related Aspects of IP Rights (TRIPS)
2. Patent Term Extension (PTE)
3. Patent Term Adjustment (PTA)
4. Maintenance Fee Non-Payment

$\Longrightarrow$ **High-level:** all changes are **accurate** except maintenance fees

- Substantial problem: $\sim$45 percent of OB patents expire before full-term
- Accurate measures of patent term require additional USPTO records

Outline

Calculating exclusivity: combining records

The "right" calculation of market exclusivity is a function of the research question

- *Nominal*: How much market exclusivity do various legal protections confer?
- *Expected*: How long can a firm expect to be shielded from competition?
- *Realized*: How much time actually elapsed before generic entry?

Calculating exclusivity: combining records

The "right" calculation of market exclusivity is a function of the research question

- *Nominal*: How much market exclusivity do various legal protections confer?
- *Expected*: How long can a firm expect to be shielded from competition?
- *Realized*: How much time actually elapsed before generic entry?

Exercise: compare measures of exclusivity to observed generic entry

- Sample: **1,673** new drugs approved between 1985 and 2014
- Median drug: 13 years of market exclusivity

Calculating exclusivity: combining records

The "right" calculation of market exclusivity is a function of the research question

- *Nominal*: How much market exclusivity do various legal protections confer?
- *Expected*: How long can a firm expect to be shielded from competition?
- *Realized*: How much time actually elapsed before generic entry?

Exercise: compare measures of exclusivity to observed generic entry

- Sample: **1,673** new drugs approved between 1985 and 2014
- Median drug: 13 years of market exclusivity

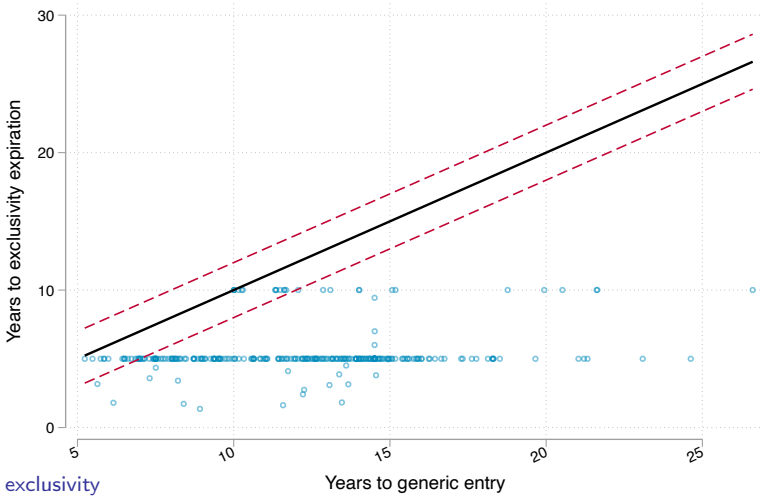
Paper provides detailed advice for calculation of "legal protection periods"

Measure #1: NCE

Test: NCE is granted & enforced by FDA $\Longrightarrow$ expect to bind

Measure #1: NCE

Test: NCE is granted & enforced by FDA $\Longrightarrow$ expect to bind



Measure #2: Earliest patent

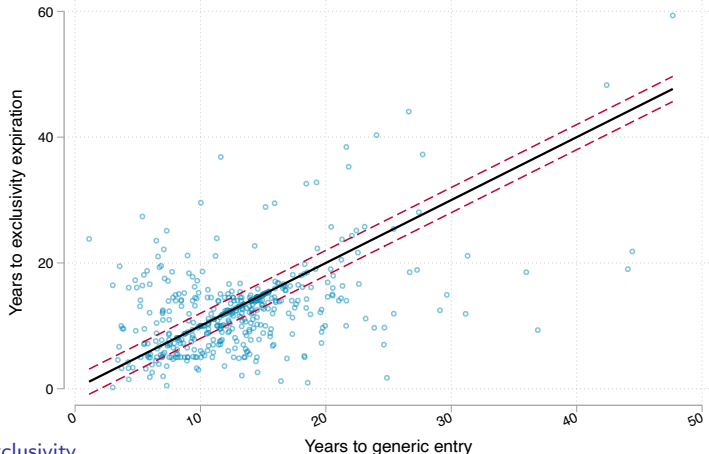
Test: Early expiring patents likely stronger than later expiring$\Longrightarrow$ expect minimal generic entry

[Hemphill and Sampat, 2011; Branstetter, Chatterjee, and Higgins, 2016; Gupta, 2021]

Measure #2: Earliest patent

Test: Early expiring patents likely stronger than later expiring$\Longrightarrow$ expect minimal generic entry

[Hemphill and Sampat, 2011; Branstetter, Chatterjee, and Higgins, 2016; Gupta, 2021]



Measure #3: Latest patent

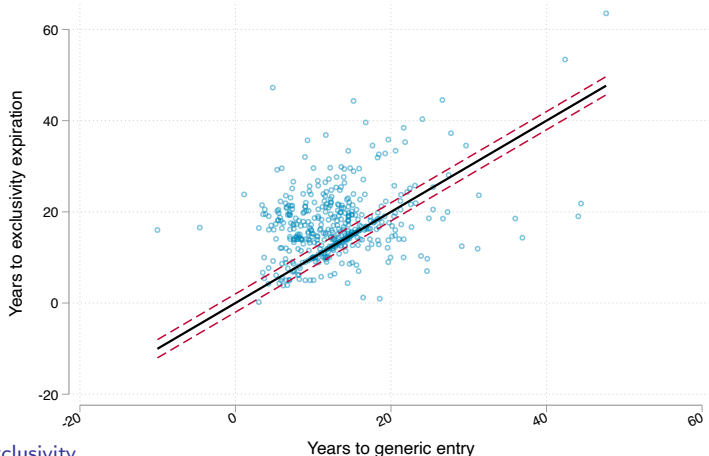
Test: Latest expiring patents are often invalidated/unenforced $\Longrightarrow$ expect more generic entry

[Hemphill and Sampat, 2011; Branstetter, Chatterjee, and Higgins, 2016; Gupta, 2021]

Measure #3: Latest patent

Test: Latest expiring patents are often invalidated/unenforced $\Longrightarrow$ expect more generic entry

[Hemphill and Sampat, 2011; Branstetter, Chatterjee, and Higgins, 2016; Gupta, 2021]



Outline

Using the Orange Book for research

Key advantages:

- Unique opportunity to study linkages between patents and products
  - Rich resource for understanding pharmaceutical markets
- Variables facilitate a wide variety of linkages to other datasets
  - Drug name, sponsor, patent number, FDA application number
- *Good news:* Records are generally complete and accurate
  - Exception: maintenance fee non-payment

Using the Orange Book for research

Key advantages:

- Unique opportunity to study linkages between patents and products
  - Rich resource for understanding pharmaceutical markets
- Variables facilitate a wide variety of linkages to other datasets
  - Drug name, sponsor, patent number, FDA application number
- *Good news:* Records are generally complete and accurate
  - Exception: maintenance fee non-payment

"Calculating exclusivity" $\Longrightarrow$ Appropriate use of data is based on questions of interest