

The UGA Patent Litigation Datafile

CODEBOOK FOR MASTER LITIGATION DATA

John L. Turner[†]

Department of Economics, University of Georgia

March 2013

1. OVERVIEW

These are the data captured as part of NSF SES-0751661, “A Comprehensive Data Set of Published US Patent Litigation Decisions: 1929-2006.” The basic framework for the data is set out by Henry, McGahee and Turner (2013, p. 2),¹ who state that the UGA Patent Litigation Datafile is

“...a set of virtually all patents found “invalid,” “not infringed” or “valid and infringed” by a district court (or the International Trade Commission) in a case where at least one decision (district or appellate) is recorded in the *United States Patents Quarterly* (USPQ) and where validity and/or infringement is a decisive issue.”

Additional variables are matched to the patents and cases satisfying these criteria. The original idea for this project was to precisely estimate rates of validity and infringement, and the data capture reflects this. Henry, McGahee and Turner (2013, subsection 3.1) includes a complete list of the criteria used.²

[†]Department of Economics, University of Georgia. Brooks Hall 5th Floor, Athens, GA 30602-6254. Phone: 706-542-3682. E-mail: jltturner@uga.edu. This project was generously supported by National Science Foundation grant SES-0751661.

¹Henry, Matthew D.; McGahee, Thomas P.; Turner, John L. 2013. “Dynamics of Patent Precedent and Enforcement: An Introduction to The UGA Patent Litigation Datafile,” University of Georgia Working Paper.

²Generally, we exclude district court decisions based on the conduct of the parties notwithstanding the underlying validity and/or infringement of the patent, such as fraud or misuse, as well as decisions based on validity and/or infringement being moot by virtue of either patentee behavior (e.g. the “on sale” provision, laches, etc.) or prior litigation (e.g. dismissals based on collateral estoppel).

2. PATENT-SPECIFIC VARIABLES

Column: A

Variable: MasterIndex

Description: Index for listing patent cases. Sorting is by patent number, then district court (DC) decision date, then appellate court (AC) decision date.

Type: Alphanumeric String.

Column: B

Variable: Patent

Description: This is the main US patent number collected from the USPQ.

Type: String

3333333	US Utility Patent number 3,333,333.
D333333	US Design Patent number 333,333.
PP03333	US Plant Patent number 3,333.
RE33333	US Reissue Patent number 33,333.
unknown	The information is not available.

Note: The above lists examples of how patent numbers look. There are 29 patent cases such that the USPQ opinion did not reveal the patent number. For these we record “unknown” for all patent characteristics.

Column: C

Variable: IssueDate

Description: The date that the USPTO issued the patent.

Type: MM/DD/YYYY

unknown	The information is not available.
---------	-----------------------------------

Column: D

Variable: AppDate

Description: The date that the patentee applied to the USPTO for the granted patent.

Type: MM/DD/YYYY

unknown	The information is not available.
---------	-----------------------------------

Column: E

Variable: OrigAppDate

Description: The oldest application date that relates directly to the patented technology.

Type: MM/DD/YYYY

unknown	The information is not available.
---------	-----------------------------------

Note: If the application was a renewal, continuation, division, etc., the OrigAppDate will be older than the AppDate. For example, for Patent 3,338,233, the application reads “Con-

tinuation of application Ser. No. 353,153, Mar. 19, 1964. This application Dec.28, 1966, Ser. No. 605,274,...” In this case AppDate is 12/28/1966 and OrigAppDate is 3/19/1964.

Column: F

Variable: ReissueOf

Description: The original patent number (if patent was a reissue patent).

Type: String

n/a	Not applicable (not a reissue).
unknown	The information is not available.

Column: G

Variable: Renewal

Description: Was the patent application a renewal of a previous application?

Type: Dummy

1	The patent application was a renewal.
0	The patent application was not a renewal.
unknown	The information is not available.

Column: H

Variable: Continuation

Description: Was the patent application a continuation of a previous application?

Type: Dummy

1	The patent application was a continuation.
0	The patent application was not a continuation.
unknown	The information is not available.

Note: A continuation-in-part is recorded as a continuation.

Column: I

Variable: Division

Description: Was the patent application a division of a previous application?

Type: Dummy

1	The patent application was a division.
0	The patent application was not a division.
unknown	The information is not available.

Column: J

Variable: Class

Description: Patent technology class.

Type: Numeric String

unknown	The information is not available.
---------	-----------------------------------

Note: See <http://www.uspto.gov/web/patents/classification/>.

Column: K

Variable: Subclass

Description: Patent technology subclass.

Type: Numeric String

unknown The information is not available.

Note: See <http://www.uspto.gov/web/patents/classification/>.

Column: L

Variable: OtherClasses

Description: Other main patent classes.

Type: Numeric String

none There are no other classes.

unknown The information is not available.

Note: See <http://www.uspto.gov/web/patents/classification/>.

Column: M

Variable: IPCCodes

Description: International Patent Classification.

Type: Alpha-numeric String

unknown The information is not available.

omitted Information not included in the patent document.

Note: See <http://www.wipo.int/classifications/ipc/en/>.

Column: N

Variable: Claims

Description: Number of claims listed in the approved patent.

Type: Integer

unknown The information is not available.

Column: O

Variable: NumInventors

Description: Number of inventors listed in the approved patent.

Type: Integer

unknown The information is not available.

Column: P

Variable: FirstInventor

Description: The first inventor listed in the patent.

Type: String

unknown The information is not available.

Column: Q

Variable: InvCity

Description: City of the first inventor.

Type: String

unknown The information is not available.

Column: R

Variable: InvState

Description: State of the first inventor.

Type: String

unknown The information is not available.

Column: S

Variable: InvCountry

Description: Country of first inventor.

Type: String

unknown The information is not available.

Note: This records WIPO-standard two-letter codes: <http://www.uspto.gov/patft/help/helpctry.htm>.

Column: T

Variable: Assignee

Description: Entity/person to whom the patent was assigned in the patent documents.

Type: String

none There was no assignee listed in the patent.

unknown The information is not available.

Column: U

Variable: AssCity

Description: Assignee(s) city

Type: String

n/a Not applicable (no assignee).

omitted Information not included in the patent document.

unknown The information is not available.

Column: V**Variable:** AssState**Description:** Assignee(s) state**Type:** String

n/a	Not applicable (no assignee).
omitted	Information not included in the patent document.
unknown	The information is not available.

Column: W**Variable:** AssCountry**Description:** Assignee(s) Country**Type:** String

n/a	Not applicable (no assignee).
omitted	Information not included in the patent document.
unknown	The information is not available.

Note: This records WIPO-standard two-letter codes: <http://www.uspto.gov/patft/help/helpctry.htm>.

Column: X**Variable:** NumRefs**Description:** Total number of citations referencing other patents/inventions (i.e., “backward references”).**Type:** Integer

omitted	Information not included in the patent document.
unknown	The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references.

Column: Y**Variable:** NumUSRefs**Description:** Number of citations referencing US patents.**Type:** Integer

omitted	Information not included in the patent document.
unknown	The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references.

Column: Z**Variable:** NumUtilityRefs**Description:** Number of citations referencing US utility patents/inventions.**Type:** Integer

omitted Information not included in the patent document.
unknown The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references.

Column: AA

Variable: NumDesignRefs

Description: Number of citations referencing US design patents/inventions

Type: Integer

omitted Information not included in the patent document.
unknown The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references.

Column: AB

Variable: NumPlantRefs

Description: Number of citations referencing US plant patents/inventions.

Type: Integer

omitted Information not included in the patent document.
unknown The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references.

Column: AC

Variable: NumReissueRefs

Description: Number of citations referencing US reissue patents.

Type: Integer

omitted Information not included in the patent document.
unknown The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references.

Column: AD

Variable: NumForReferences

Description: Number of citations referencing foreign patents.

Type: Integer

omitted Information not included in the patent document.
unknown The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information

about backward references.

Column: AE

Variable: OCPDate

Description: Oldest cited patent.

Type: MM/DD/YYYY

omitted Information not included in the patent document.

unknown The information is not available.

n/a Not applicable.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references. For a post-1947 patent with no references, we put “n/a” here.

Column: AF

Variable: MCPDate

Description: Median cited patent.

Type: MM/DD/YYYY

omitted Information not included in the patent document.

unknown The information is not available.

Note: US Patent documents issued prior to January 1, 1947 do not include information about backward references. For a post-1947 patent with no references, we put “n/a” here.

3. LITIGATION-SPECIFIC DATA

Column: AG

Variable: USPQDocketNum

Description: Docket number listed in the opinion published in the USPQ.

Type: String

notcapt Information may be available but has not yet been captured.

unknown The information is not available.

Column: AH

Variable: DJ

Description: Is the lawsuit for declaratory judgment?

Type: Dummy

1 The lawsuit is a declaratory judgment.

0 The lawsuit is not a declaratory judgment.

Column: AI

Variable: PatDef

Description: Is the patentee the defendant in the case?

Type: Dummy

- 1 The patentee is the defendant.
- 0 The patentee is not the defendant.

Column: AJ

Variable: DCPub

Description: Did the USPQ publish the opinion?

Type: Dummy

- 1 The USPQ published the opinion.
- 0 The USPQ did not publish the opinion.

Column: AK

Variable: DCEdition

Description: Edition of the federal report that recorded the patent litigation case.

Type: Alpha-numeric string

- 1 Edition 1 of the USPQ.
- 2 Edition 2 of the USPQ.
- F. Supp. Federal Supplement.
- F. Supp. 2d Federal Supplement (2nd edition).
- F.2d Federal Reporter (2nd edition).
- F.3d Federal Reporter (3rd edition).
- nocite DC decision not published or recorded.
- unpub Unpublished.
- WL Westlaw.

Note: The USPQ had two editions, 1 and 2. Edition 2 started at the beginning of 1987, with volume 1. Page numbers are four digits in Edition 2. Typically when there is no cite but we know the decision, we observe it from an opinion in another part of the case, e.g. the appellate decision is published and that opinion includes some information about the DC decision.

Column: AL

Variable: DCVolume

Description: Volume in which the DC case opinion was published.

Type: Numeric

- nocite DC decision not published or recorded.

Note: Typically when there is no cite but we know the decision, we observe it from an opinion in another part of the case, e.g. the appellate decision is published and that opinion includes some information about the DC decision.

Column: AM

Variable: DCPage

Description: USPQ page on which the DC case opinion was published.

Type: Numeric

nocite DC decision not published or recorded.

Note: Typically when there is no cite but we know the decision, we observe it from an opinion in another part of the case, e.g. the appellate decision is published and that opinion includes some information about the DC decision.

Column: AN

Variable: DCNumPatents

Description: Number of patents found INV, NINF or V&I in the DC decision.

Type: Integer

Note: In cases where there is a published AC decision but not a published DC decision, we record the number of patents in the AC decision here.

Column: AO

Variable: DCDecision

Description: The decision reached in the district court.

Type: String

INV Patent was found invalid.

NINF Patent was found not infringed.

V&I Patent was found valid and infringed.

Note: This variable was captured by careful reading of the DC decision. The data capture for this set restricts attention to patent cases where infringement and/or validity are at issue in the first DC decision. Hence, cases whose decisions were based on fraud, misuse, unenforceability, etc. were left out. See Henry and Turner (2006, pp. 95-97).

Column: AP

Variable: DCState

Description: State in which the District Court is located.

Type: String

n/a No applicable state.

unknown The information is not available.

Note: We use standard two-letter abbreviations for US states. There is no applicable state when the court is the US Court of Claims (USCC) or the International Trade Commission (ITC). Typically, observations with “unknown” are cases where the appellate opinion does not provide information about district court.

Column: AQ

Variable: District

Description: District court district.

Type: String

CD	Central District.
ED	Eastern District.
ITC	International Trade Commission.
MD	Middle District.
ND	Northern District.
SD	Southern District.
TC	Tariff Commission.
USCC	US Court of Claims.
WD	Western District.
n/a	Not applicable.
unknown	The information is not available.

Note: Cases heard in Chicago, IL are in the northern district of Illinois, for example. When "n/a" appears, this typically means that the state has only one district (e.g. Delaware).

Column: AR

Variable: CAFCEra

Description: Case took place in the era when the CAFC was the relevant appellate court.

Type: Dummy

unknown	The information is not available.
---------	-----------------------------------

Note: This definition follows Atkinson, Marco and Turner (2009). If there was an appellate decision for the patent case, then the variable is a 1 if the CAFC heard the case and a 0 otherwise. If there was no appellate decision, then CAFCEra equals 1 if the district decision occurs on or after October 1, 1982, and CAFCEra equals 0 otherwise.

Column: AS

Variable: DistDate

Description: Date of the DC Decision.

Type: MM/DD/YYYY

unknown	The information is not available.
---------	-----------------------------------

Note: District Court decision dates are almost always listed in the DC decision, but are seldom listed in the AC decision.

Column: AT

Variable: Patentee

Description: Name of the patentee.

Type: String

Note: Identified from careful reading of the DC decision.

Column: AU

Variable: AllegInf

Description: Name of the first listed alleged infringer.

Type: String

Note: Identified from careful reading of the DC decision.

Column: AV

Variable: AllegInf2

Description: Name of the second and subsequent alleged infringers.

Type: String

unknown The information is not available.

Note: Identified from careful reading of the DC decision.

Column: AW

Variable: ACPub

Description: Did the USPQ publish the appellate decision?

Type: Dummy

1 The USPQ published the appellate decision.
0 The USPQ did not publish the appellate decision.
n/a Not applicable (no appellate case).

Column: AX

Variable: ACEdition

Description: Edition in which the AC decision was published.

Type: String

1 Edition 1 of the USPQ.
0 Edition 2 of the USPQ.
F.2d Federal Reporter (2nd edition).
F.3d Federal Reporter (3rd edition).
n/a Not applicable (no appellate case).
unpub Unpublished.
WL Westlaw.

Note: Typically when there is no cite but we know the decision, we observe it from an opinion in another part of the case, e.g. a Supreme Court decision.

Column: AY

Variable: ACVolume

Description: Volume in which the AC decision was published.

Type: String

n/a Not applicable (no appellate case).

unpub Unpublished.

Note: Typically when there is no cite but we know the decision, we observe it from an opinion in another part of the case, e.g. a Supreme Court decision.

Column: AZ

Variable: ACPage

Description: Page on which the AC decision was published.

Type: String

n/a Not applicable (no appellate case).

unpub Unpublished.

Note: Typically when there is no cite but we know the decision, we observe it from an opinion in another part of the case, e.g. a Supreme Court decision.

Column: BA

Variable: ACNumPatents

Description: Number of patents involved in the AC decision.

Type: Integer

n/a Not applicable (no appellate case).

unknown The information is not available.

Column: BB

Variable: FirstAC

Description: The first appellate court in which the case appears.

Type: Alpha-numeric string

1 First Circuit.

...

11 Eleventh Circuit.

DC DC Circuit.

CAFC Court of Appeals for the Federal Circuit.

CCPA Court of Customs and Patent Appeals.

n/a Not applicable (no appellate case).

Note: The 11th Circuit was created in 1982 (the old 5th Circuit was split in two), just prior to the creation of the CAFC. As a result, there are only a very small number of patent cases in the data that were heard by the 11th Circuit. The CAFC was created by combining the Court of Claims and the CCPA, so the CCPA does not hear any cases after 1982.

Column: BC

Variable: ACDate

Description: The date of the first appellate decision in the case.

Type: MM/DD/YYYY

n/a	Not applicable (no appellate case).
unknown	The information is not available.

Column: BD

Variable: Affirmed

Description: Was the district court decision affirmed?

Type: Dummy

1	The district court decision was affirmed.
0	The district court decision was not affirmed.
n/a	Not applicable (no appellate case).

Note: This variable captures the least-confusing and cleanest way to measure the performance of appellate courts.

Column: BE

Variable: ACDecision

Description: Details about the nature of the appellate decision

Type: String

affirmed	The DC decision was affirmed.
n/a	Not applicable (no appellate case).
ninf*	The AC vacated invalidity and replaced it with not infringed.
R&R	The AC reversed the DC and remanded the case back to the DC.
rem	The AC remanded the case back to the DC.
rev(INV)	The AC reversed the DC and replaced it with invalid.
rev(NINF)	The AC reversed the DC and replaced it with not infringed.
rev(V&I)	The AC reversed the DC and replaced it with valid and infringed.
rev(Valid)	The AC reversed the DC decision of invalidity, did not address infringement.
V&R	The AC vacated the DC decision and remanded the case back to the DC.

Note: The ninf* decision pertains to a period between 1987-93 when the CAFC made it standard practice to vacate district court decisions relating to validity (for any and all claims found valid or invalid) whenever it found a patent not infringed. The Supreme Court decision in *Cardinal Chemical Co. v. Morton Int'l, Inc.* 508 U.S. 83 (1993) ended this practice. Decisions starting with a tilde(~) involve things like estoppel, fraud, misuse, etc., that relate to conduct of the parties notwithstanding the validity and/or infringement of the patent.

Column: BF

Variable: DC2ndDecDate

Description: Decision date for 2nd DC decision in the case.

Type: MM/DD/YYYY

n/a no 2nd decision.

unknown The information is not available.

Column: BG

Variable: DC2ndCite

Description: Citation for 2nd DC Decision

Type: Vol-Pg

n/a no 2nd decision.

unknown The information is not available.

Note: If there are just numbers, this is a USPQ cite.

Column: BH

Variable: DC2ndDec

Description: The 2nd DC decision.

Type: String

INV Patent was found invalid.

n/a Not applicable.

NINF Patent was found not infringed.

unknown The information is not available.

V&I Patent was found valid and infringed.

Note: Decisions starting with a tilde(~) involve things like estoppel, fraud, misuse, etc., that relate to conduct of the parties notwithstanding the validity and/or infringement of the patent. These patent cases would be left out of the data if this was the first decision.

Column: BI

Variable: AC2ndDecDate

Description: Decision date for 2nd AC decision in the case.

Type: MM/DD/YYYY

n/a no 2nd AC decision.

unknown The information is not available.

Column: BJ

Variable: AC2ndCite

Description: Citation for 2nd AC Decision

Type: Vol-Pg

n/a no 2nd AC decision.

unknown The information is not available.

Note: If there are just numbers, this is a USPQ cite.

Column: BK

Variable: AC2ndDec

Description: The decision in the 2nd AC decision.

Type: String

affirmed	The DC decision was affirmed.
n/a	Not applicable (no appellate case).
ninf*	The AC vacated invalidity and replaced it with not infringed.
R&R	The AC reversed the DC and remanded the case back to the DC.
rem	The AC remanded the case back to the DC.
rev(INV)	The AC reversed the DC and replaced it with invalid.
rev(NINF)	The AC reversed the DC and replaced it with not infringed.
rev(V&I)	The AC reversed the DC and replaced it with valid and infringed.
rev(Valid)	The AC reversed the DC decision of invalidity, did not address infringement.
V&R	The AC vacated the DC and remanded the case back to the DC.