PREDICTING THE "UNPREDICTABLE": AN EMPIRICAL ANALYSIS OF U.S. PATENT INFRINGEMENT AWARDS

12TH Annual Intellectual Property Scholars Conference
Stanford University Law School
9 August 2012

Michael J. Mazzeo Kellogg School of Management, Northwestern University

Jonathan Hillel Skadden, Arps, Slate, Meagher & Flom LLP

Samantha Zyontz Harvard University

Focus: Predicting Damage Awards

- Widespread concern exists about the "unpredictability" of patent damage awards and its effect on everything from litigation strategy to incentives for innovative activity.
  - 2011 FTC Report highlights "lottery ticket mentality" regarding litigation outcomes in some circles.

Focus: Predicting Damage Awards

- Widespread concern exists about the "unpredictability" of patent damage awards and its effect on everything from litigation strategy to incentives for innovative activity.
  - 2011 FTC Report highlights "lottery ticket mentality" regarding litigation outcomes in some circles.

- Our approach: assemble comprehensive data on damage awards and run straightforward regressions that use readily available, reasonable factors to predict award size.

Focus: Predicting Damage Awards

- Widespread concern exists about the "unpredictability" of patent damage awards and its effect on everything from litigation strategy to incentives for innovative activity.
  - 2011 FTC Report highlights "lottery ticket mentality" regarding litigation outcomes in some circles.

- Our approach: assemble comprehensive data on damage awards and run straightforward regressions that use readily available factors to predict award size.

- Findings: Infringement damages are highly predictable overall and are correlated with factors associated with economic value of patents, litigant size and case complexity.

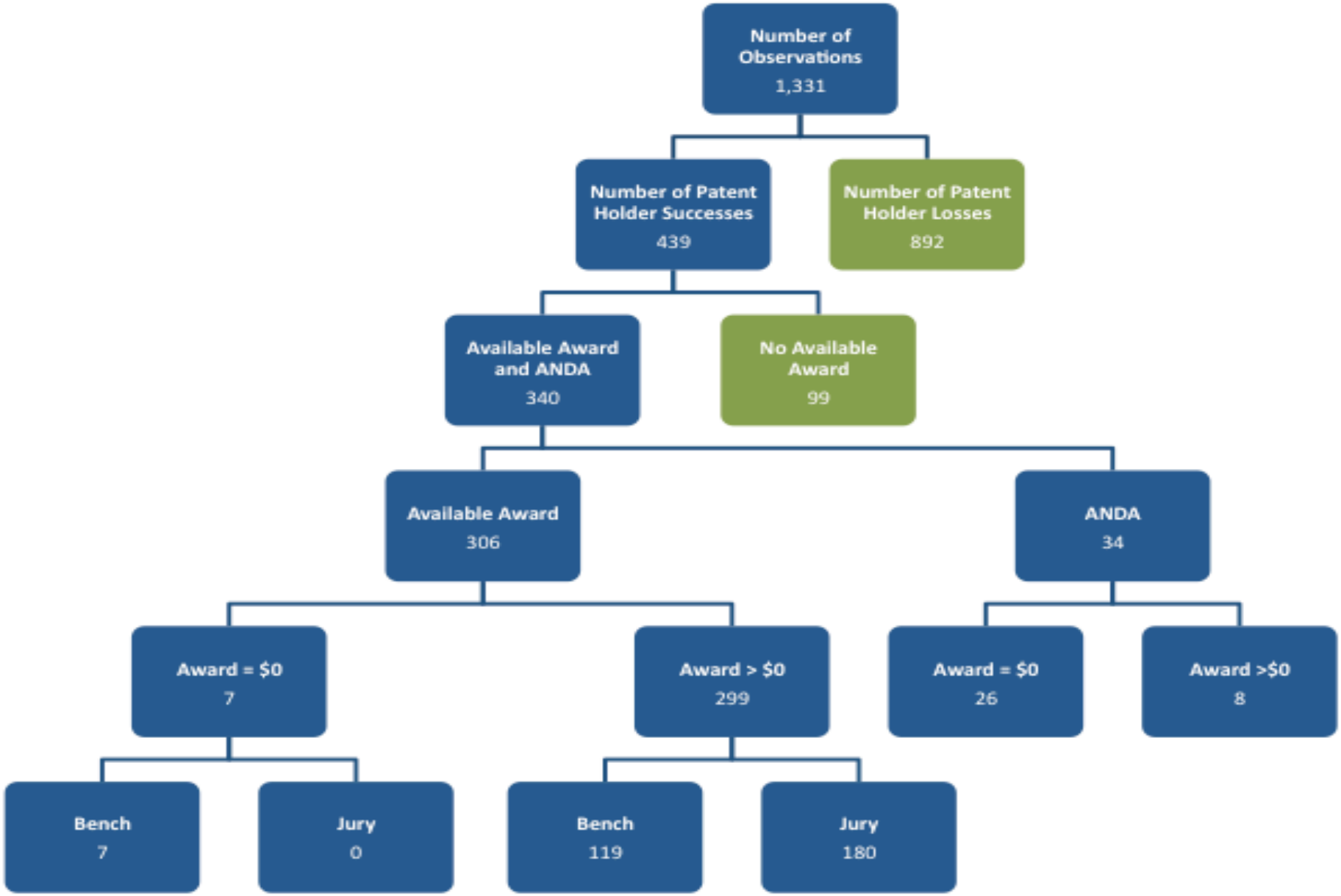
Prior Literature

- Studies by Lanjouw & Schankerman (1999-2004) described the predictors of patent litigation.
- Studies by consulting firm PwC (2007-2009) described the data (and caused considerable alarm).
- Lemley & Shapiro (2007) – demonstrated heterogeneity across industries in reasonable royalty rates.
- Allison, Lemley & Walker (2009) – described the characteristics of the "most litigated patents."
- Operdeck (2009) – finds no overriding patterns when trying to "explain" the size of awards statistically.

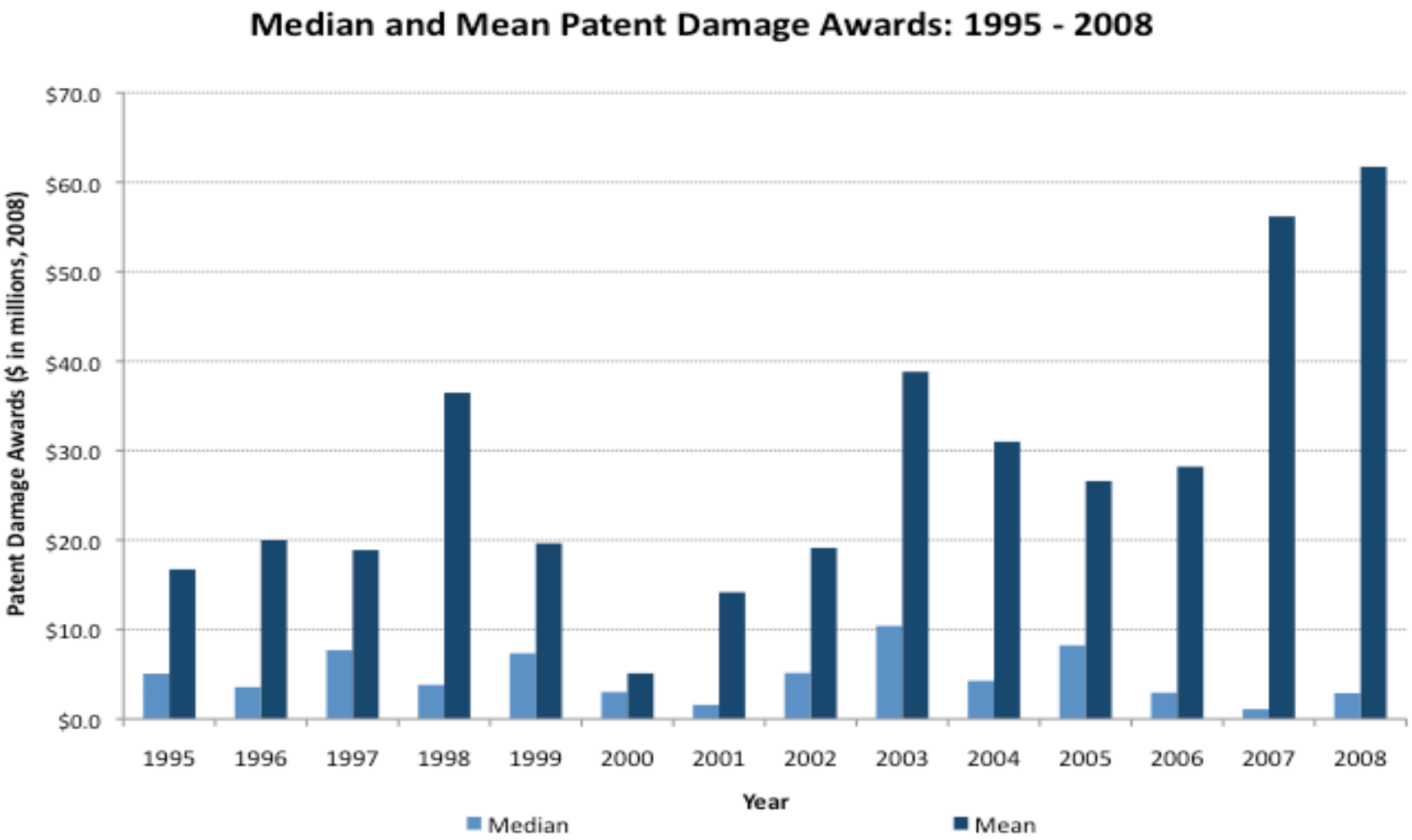
Analysis

- Dataset: comprehensive information from 340 cases decided in US federal courts between 1995 and 2008.

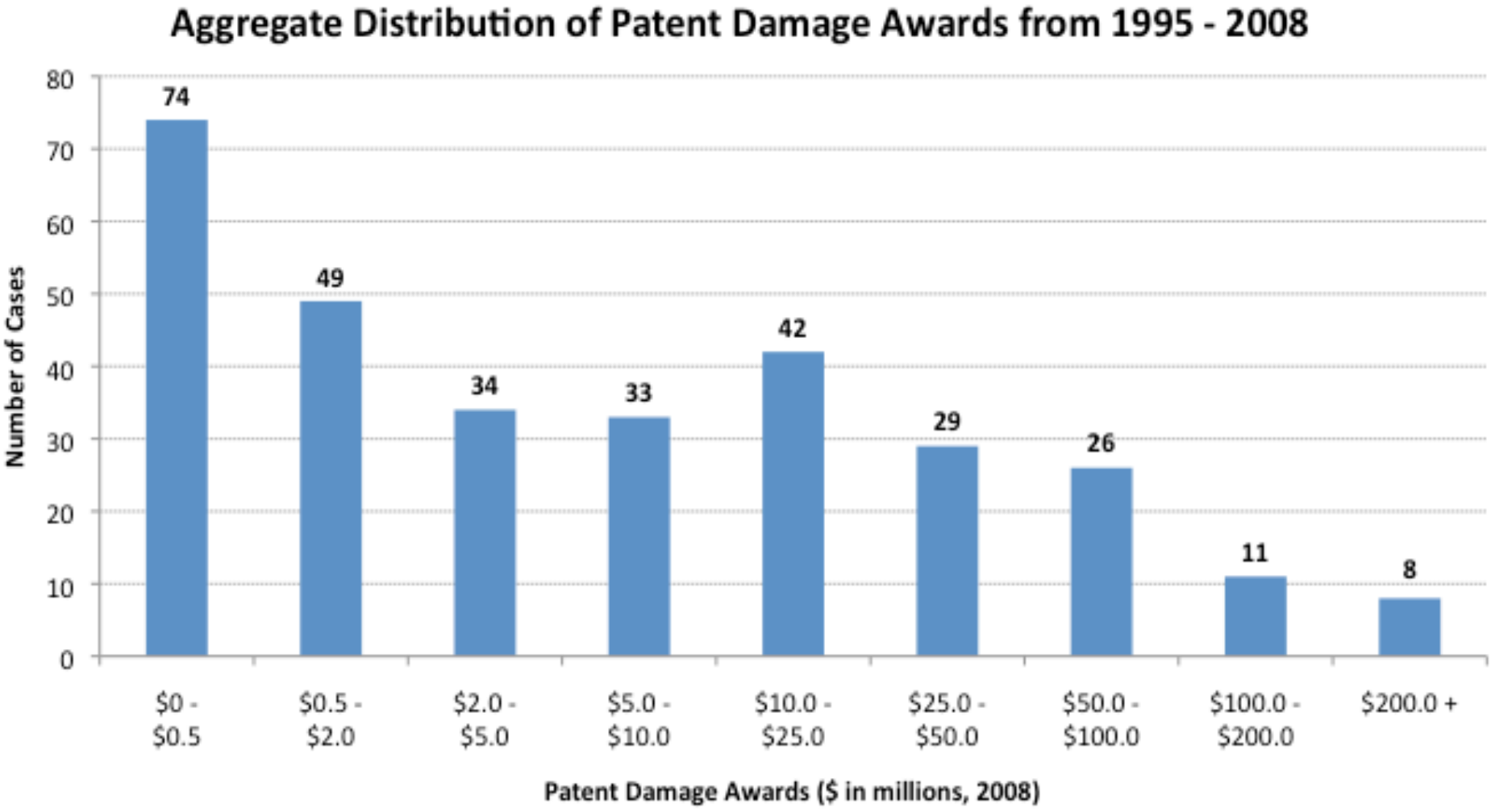
Evolving the PwC Dataset



Dataset: Size distribution of damage awards in patent infringement cases, 1995-2008



Almost the Entire Iceberg: the top eight cases represent 47.6 percent of collective damages



Analysis

- Dataset: comprehensive information from 340 cases decided in US federal courts between 1995 and 2008.

- Controls: assembled a detailed set of case characteristics, matched to the damage award levels, to act as potential explanatory variables.

Variable Groups	Description	Sources
Category 1: Case Information		
Identifiers	Variables including a unique ID assigned by the authors, the docket number of the case, and the full names of the first listed plaintiff and defendant in the case.	PwC database, Google, Westlaw, and PACER
Dates	Variables including the year of the original award in district court, date the complaint for case was filed, the earliest start date of trial on validity, infringement, or damages, and the number of days between the trial start date and the complaint date.	PwC database, Google, Westlaw, and PACER
Location	Variables including where the case was litigated, including state, circuit, and court.	PwC database, Google, Westlaw, and PACER
Other Case Information	Variables determining if the case contained a summary judgment for the patent holder on validity and/or infringement, if the case involved an invalidated patent-at-issue, and if the patent holder was successful in its patent claims.	PwC database, Google, Westlaw, and PACER
Damage Awards	If the patent holder was successful, variables for the total award amount, lost profits, reasonable royalties, prejudgment interest, enhanced damages, price erosion damages, and other damages. Also included are whether or not the case settled before damages were awarded, whether or not the case resulted in only an injunction, and whether or not the case was an ANDA filing.	PwC database, Google, Westlaw, and PACER
Category 2: Litigant Information		
General Assignee	Includes number of patent assignees associated with the patents-at-issue in the case, the names of the assignees, if one of the assignee(s) is the first named plaintiff or defendant in the case (can be both), if the plaintiff name listed is an assignee (patent holder), and if the patent holder markets or manufactures its technology covered by the patent.	PwC database, Google, Westlaw, PACER, and NBER patent database
NBER Assignee	Dummy variables from the 2002 NBER database which coded the Assignee(s) as "Unassigned," "US, Non-Government," Non-US, Non-Government,", "US, Individual," "Non-US, Individual," "US Government," or "Non-US, Government."	NBER patent database
Assignee Identifiers	Includes the variables determining whether or not the first named plaintiff or defendant are an individual, private entity, public entity, university, part of the U.S. government, a domestic entity, foreign entity, part of the 2009 Fortune 500 list, part of the 2009 Fortune 1000 list, a subsidiary of a parent company.	EDGAR, Manta, Hoover's Online, Westlaw, and Fortune 1000
Assignee Parent Identifiers	Variables for the parent companies of the plaintiff or defendant listed if it was a subsidiary that include whether or not the parent company is a private entity, public entity, domestic entity, foreign entity, part of the 2009 Fortune 500 list, part of the 2009 Fortune 1000 list, if the first named plaintiff or defendant is owned by a joint venture (2 parents or more).	EDGAR, Manta, Hoover's Online, Westlaw, and Fortune 1000
SIC Codes	Variables identifying the 2-, 3-, and 4- digit SIC codes for the potential infringers.	NBER patent database, Google, and Westlaw
Category 3: Patent(s)-at-Issue Information		
General Patent	Variables identifying the number of patent(s) at issue in the case and their type as either utility, reissue, design, or application number.	NBER patent database, Google, and Westlaw
Patent Classification	Includes variables for all patents-at-issue such as application year calculated for minimum and maximum (minimums and maxima differ for cases with multiple patents-at-issue and are the same for cases with only one patent-at-issue); grant date year calculated for minimum and maximum; grant date calculated for minimum and maximum; age of the oldest and youngest patent-at-issue in a case calculated for minimum and maximum; number of claims calculated for minimum, maximum, average and total; number of forward citations through 2002 from the NBER 2002 data, calculated for minimum, maximum and average; number of forward citations through 2010 if the 2002 forward citations were not available, calculated for minimum, maximum and average; the IPC4 classification listed first on the patent; and the PTO main classification for each patent listed in the case.	NBER patent database, Google, and Westlaw

Analysis

- Dataset: comprehensive information from 340 cases decided in US federal courts between 1995 and 2008.

- Controls: assembled a detailed set of case characteristics, matched to the damage award levels, to act as potential explanatory variables.

- Regressions:
  1. Overall predictability of damage award amounts.
  2. Analysis of explanatory power of particular significant factors.

Regressions (1): Overall predictability

Dependent Variable = Patent Damage Awards in 2008 $	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)	Model (7)
R-Squared	0.6399	0.7340	0.7403	0.7427	0.7561	0.7702	0.4457
Adjusted R-Squared	0.5368	0.6566	0.6621	0.6599	0.6618	0.6696	0.2030
F (k-1, N-k)	5.88 (75, 262)	15.15 (76, 261)	14.40 (78, 259)	20.44 (82, 255)	20.12 (94, 243)	19.50 (95, 217)	2.54 (95, 217)
Sample Size (N)	338	338	338	338	338	313	313
Standard Errors	Robust	Robust	Robust	Robust	Robust	Robust	Robust
Dependent Variable Type	Log	Log	Log	Log	Log	Log	Linear
Independent Variables	Base Controls	Model (1) + ANDA Dummy	Model (2) + Interactions	Model (3) + non-parametric total patents	Model (4) + Year Dummies	Model (5) + Avg. Forward Citations	Model (6)

Regressions (1): Overall predictability

Dependent Variable = Patent Damage Awards in 2008 $	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)	Model (7)
R-Squared	0.6399	0.7340	0.7403	0.7427	0.7561	0.7702	0.4457
Adjusted R-Squared	0.5368	0.6566	0.6621	0.6599	0.6618	0.6696	0.2030
F (k-1, N-k)	5.88 (75, 262)	15.15 (76, 261)	14.40 (78, 259)	20.44 (82, 255)	20.12 (94, 243)	19.50 (95, 217)	2.54 (95, 217)
Sample Size (N)	338	338	338	338	338	313	313
Standard Errors	Robust	Robust	Robust	Robust	Robust	Robust	Robust
Dependent Variable Type	Log	Log	Log	Log	Log	Log	Linear
Independent Variables	Base Controls	Model (1) + ANDA Dummy	Model (2) + Interactions	Model (3) + non-parametric total patents	Model (4) + Year Dummies	Model (5) + Avg. Forward Citations	Model (6)

Regressions (2): What matters?

- Focus the analysis on exactly which critical factors help to explain the size of awarded damages:
  - Underlying "value" of the patents in the case:
    - Number of patents
    - Number of claims
    - Forward citations
    - Patent Age
  - Litigant information:
    - Status of patent holders as practicing entities
    - Proxies for size/income of defendants
  - Case strategy information:
    - Judge vs. Jury
    - Time-to-trial

Regressions (2): What matters?

Number of obs	240
F(10, 229)	15.710
Prob > F	0.000
R-squared	0.362
Root MSE	88629.000

Dependent = *Log of patent damage awards in 2008 dollars*	Coef.	Robust Std. Error	t	P>t	[95% Conf. Interval]	
Average Number of Patent Claims	0.00418	0.00169	2.47	0.014	0.00849	0.00751
Number of Patents	0.07319	0.01466	4.99	0.000	0.04431	0.10208
Average Number of Forward Citations	0.00526	0.00182	2.89	0.004	0.00168	0.00884
Average Age of Patent	0.00009	0.00004	2.31	0.022	0.00001	0.00016
Dummy for "Practicing" Patent Holder	0.18153	0.13329	1.36	0.175	0.08111	0.44417
Defendant is a Fortune 500 Comp. (or sub)	0.25912	0.18626	1.39	0.166	0.10788	0.62613
Defendant is a Public Comp. (or sub)	0.63925	0.13479	4.74	0.000	0.37367	0.90482
Dummy for Trial by Jury	0.77575	0.15008	5.17	0.000	0.48003	1.07146
Time-to-Trial (days)	0.00032	0.00008	4.06	0.000	0.00017	0.00048
Year of Decision (time trend)	-0.05784	0.01557	-3.72	0.000	0.08851	0.02717
Constant	120.59220	31.11397	3.88	0.000	59.28595	181.89850

Regressions (2): What matters?

Number of obs	240
F(10, 229)	15.710
Prob > F	0.000
R-squared	0.362
Root MSE	88629.000

Dependent = *Log of patent damage awards in 2008 dollars*	Coef.	Robust Std. Error	t	P>t	[95% Conf. Interval]	
Average Number of Patent Claims	0.00418	0.00169	2.47	0.014	0.00849	0.00751
Number of Patents	0.07319	0.01466	4.99	0.000	0.04431	0.10208
Average Number of Forward Citations	0.00526	0.00182	2.89	0.004	0.00168	0.00884
Average Age of Patent	0.00009	0.00004	2.31	0.022	0.00001	0.00016
Dummy for "Practicing" Patent Holder	0.18153	0.13329	1.36	0.175	0.08111	0.44417
Defendant is a Fortune 500 Comp. (or sub)	0.25912	0.18626	1.39	0.166	0.10788	0.62613
Defendant is a Public Comp. (or sub)	0.63925	0.13479	4.74	0.000	0.37367	0.90482
Dummy for Trial by Jury	0.77575	0.15008	5.17	0.000	0.48003	1.07146
Time-to-Trial (days)	0.00032	0.00008	4.06	0.000	0.00017	0.00048
Year of Decision (time trend)	-0.05784	0.01557	-3.72	0.000	0.08851	0.02717
Constant	120.59220	31.11397	3.88	0.000	59.28595	181.89850

Regressions (2): What matters?

Number of obs	240
F(10, 229)	15.710
Prob > F	0.000
R-squared	0.362
Root MSE	88629.000

Dependent = *Log of patent damage awards in 2008 dollars*	Coef.	Robust Std. Error	t	P>t	[95% Conf. Interval]	
Average Number of Patent Claims	0.00418	0.00169	2.47	0.014	0.00849	0.00751
Number of Patents	0.07319	0.01466	4.99	0.000	0.04431	0.10208
Average Number of Forward Citations	0.00526	0.00182	2.89	0.004	0.00168	0.00884
Average Age of Patent	0.00009	0.00004	2.31	0.022	0.00001	0.00016
Dummy for "Practicing" Patent Holder	0.18153	0.13329	1.36	0.175	0.08111	0.44417
Defendant is a Fortune 500 Comp. (or sub)	0.25912	0.18626	1.39	0.166	0.10788	0.62613
Defendant is a Public Comp. (or sub)	0.63925	0.13479	4.74	0.000	0.37367	0.90482
Dummy for Trial by Jury	0.77575	0.15008	5.17	0.000	0.48003	1.07146
Time-to-Trial (days)	0.00032	0.00008	4.06	0.000	0.00017	0.00048
Year of Decision (time trend)	-0.05784	0.01557	-3.72	0.000	0.08851	0.02717
Constant	120.59220	31.11397	3.88	0.000	59.28595	181.89850

Regressions (2): What matters?

Number of obs	240
F(10, 229)	15.710
Prob > F	0.000
R-squared	0.362
Root MSE	88629.000

Dependent = *Log of patent damage awards in 2008 dollars*	Coef.	Robust Std. Error	t	P>t	[95% Conf. Interval]	
Average Number of Patent Claims	0.00418	0.00169	2.47	0.014	0.00849	0.00751
Number of Patents	0.07319	0.01466	4.99	0.000	0.04431	0.10208
Average Number of Forward Citations	0.00526	0.00182	2.89	0.004	0.00168	0.00884
Average Age of Patent	0.00009	0.00004	2.31	0.022	0.00001	0.00016
Dummy for "Practicing" Patent Holder	0.18153	0.13329	1.36	0.175	0.08111	0.44417
Defendant is a Fortune 500 Comp. (or sub)	0.25912	0.18626	1.39	0.166	0.10788	0.62613
Defendant is a Public Comp. (or sub)	0.63925	0.13479	4.74	0.000	0.37367	0.90482
Dummy for Trial by Jury	0.77575	0.15008	5.17	0.000	0.48003	1.07146
Time-to-Trial (days)	0.00032	0.00008	4.06	0.000	0.00017	0.00048
Year of Decision (time trend)	-0.05784	0.01557	-3.72	0.000	0.08851	0.02717
Constant	120.59220	31.11397	3.88	0.000	59.28595	181.89850

Regressions (2): What matters?

Number of obs	240
F(10, 229)	15.710
Prob > F	0.000
R-squared	0.362
Root MSE	88629.000

Dependent = *Log of patent damage awards in 2008 dollars*	Coef.	Robust Std. Error	t	P>t	[95% Conf. Interval]	
Average Number of Patent Claims	0.00418	0.00169	2.47	0.014	0.00849	0.00751
Number of Patents	0.07319	0.01466	4.99	0.000	0.04431	0.10208
Average Number of Forward Citations	0.00526	0.00182	2.89	0.004	0.00168	0.00884
Average Age of Patent	0.00009	0.00004	2.31	0.022	0.00001	0.00016
Dummy for "Practicing" Patent Holder	0.18153	0.13329	1.36	0.175	0.08111	0.44417
Defendant is a Fortune 500 Comp. (or sub)	0.25912	0.18626	1.39	0.166	0.10788	0.62613
Defendant is a Public Comp. (or sub)	0.63925	0.13479	4.74	0.000	0.37367	0.90482
Dummy for Trial by Jury	0.77575	0.15008	5.17	0.000	0.48003	1.07146
Time-to-Trial (days)	0.00032	0.00008	4.06	0.000	0.00017	0.00048
Year of Decision (time trend)	-0.05784	0.01557	-3.72	0.000	0.08851	0.02717
Constant	120.59220	31.11397	3.88	0.000	59.28595	181.89850

Applications & Extensions

- Model that "explains" awards can also be used to "predict" damage award levels based on available data (case, litigant and patent-at-issue information).

- Expand dataset to include information about:
  - More nuanced details regarding potential non-practicing entities
  - Cases lost at trial
  - Cases settled between infringement decision and damage awards

Summary

- Systematic empirical evidence suggests that the well-publicized, very large patent infringement damage awards are infrequent.
- Constructed regression model with detailed control variables explains considerable portion of the variation in observed damage awards.
- More targeted regressions suggest that patent "value," litigant size and case strategy affect the level of damage awards (in predictable ways).
- Future research: expanding the dataset on damage awards and exploring other datasets on patent value.