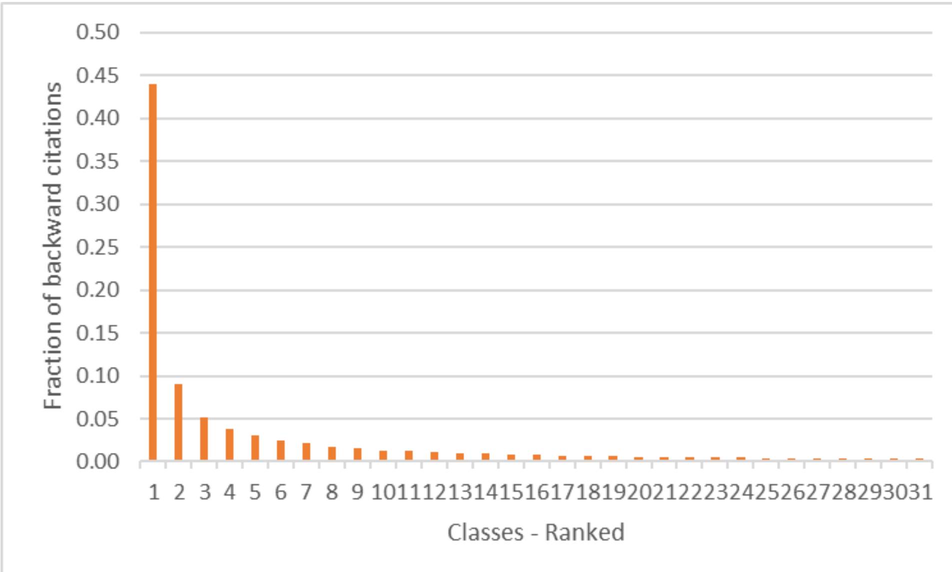
Appendices for

"What is the Price of Spanning Domain Boundaries? Distant Recombination and the Market Valuations of Firm Inventions"

Appendix A. List of Supplemental Figures and Analyses.

Figure A1. Distribution of Backward Citations between Local and Distant Classes



Note: Local class is ranked the first on the horizontal axis. Distant classes are ranked by their proximity to a focal domain. Only the top 30 distant classes (2nd to 31st in the graph) are shown in the figure.

Figure A2. Distribution of Domain Connectivity Score within Local Classes and between a Local class and Its Most Connected Distant Class.

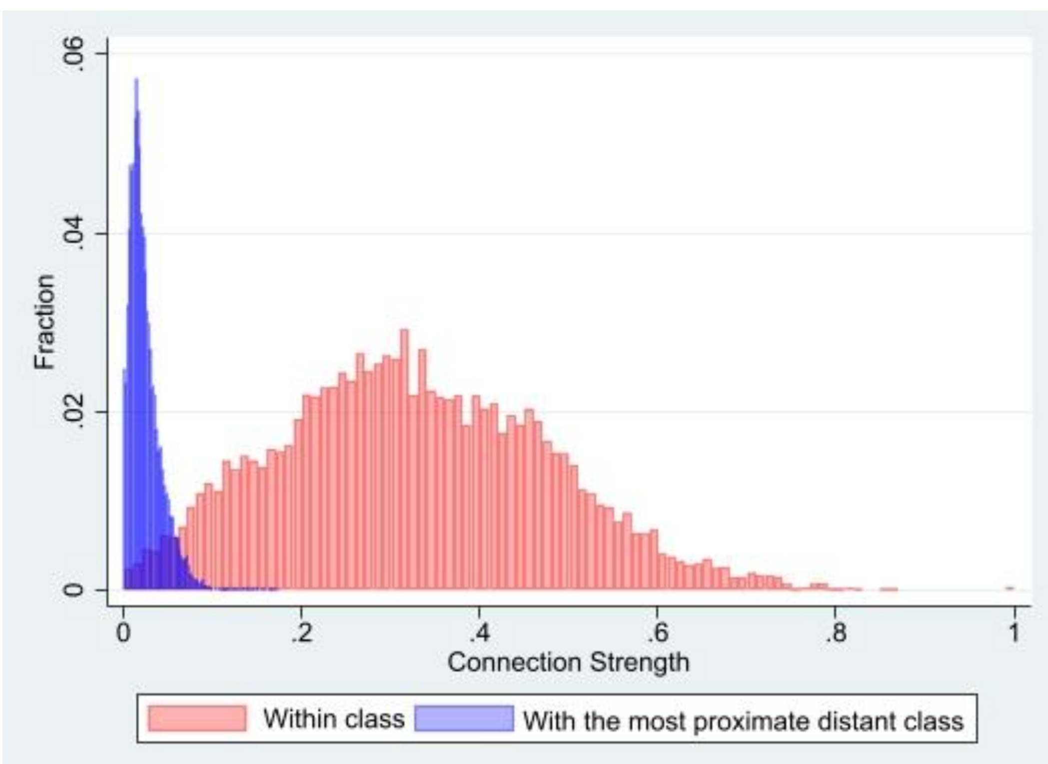
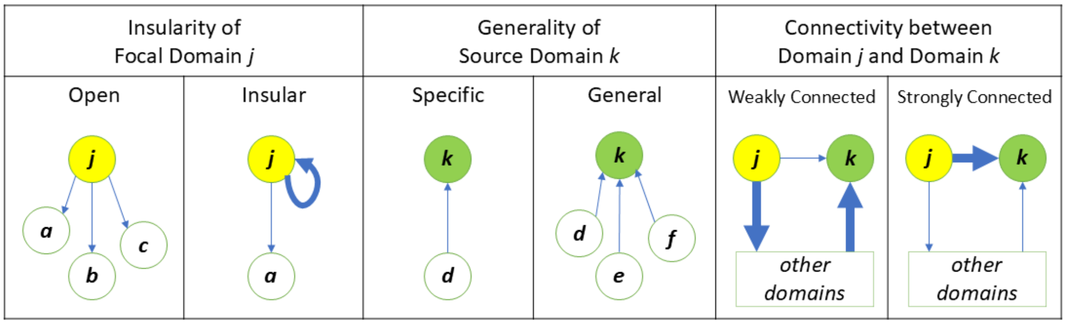
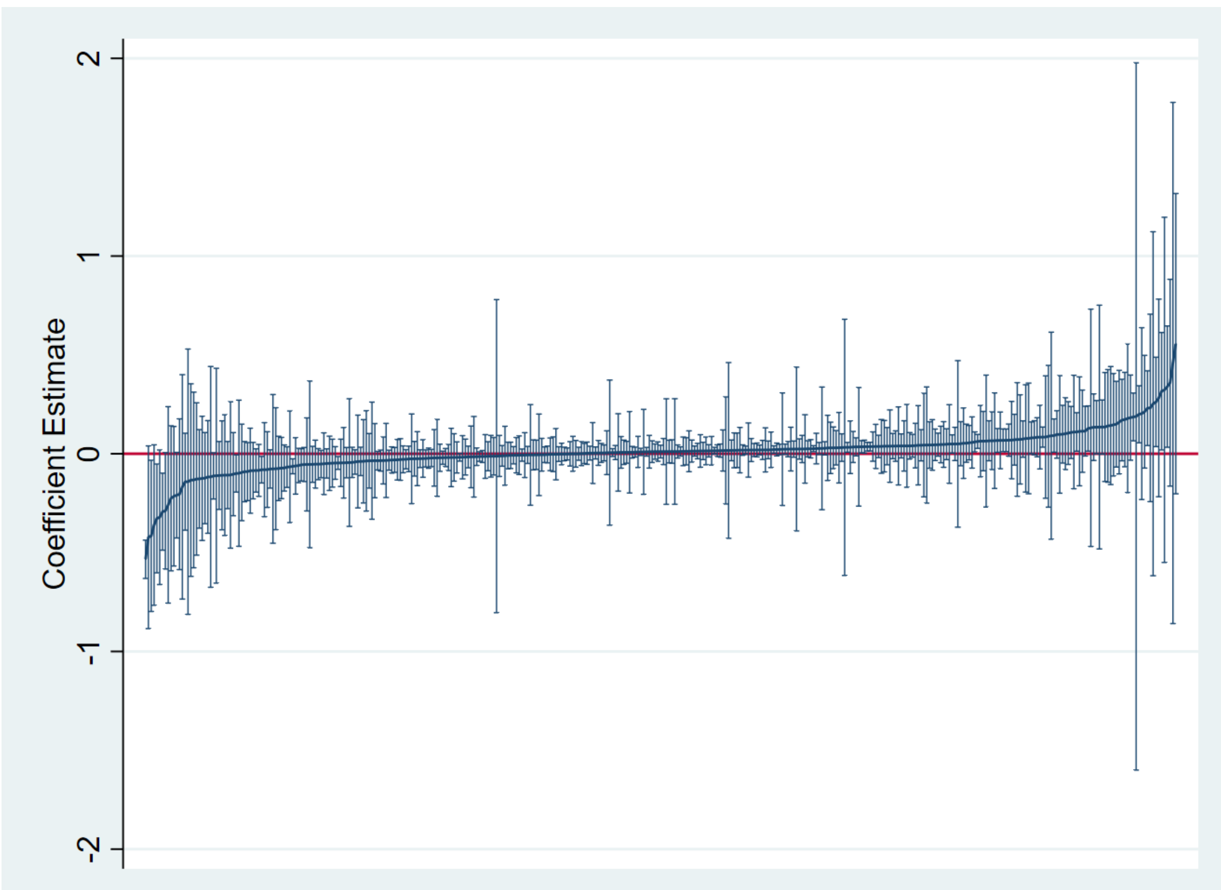


Figure A3. Domain-level Measure Illustrations



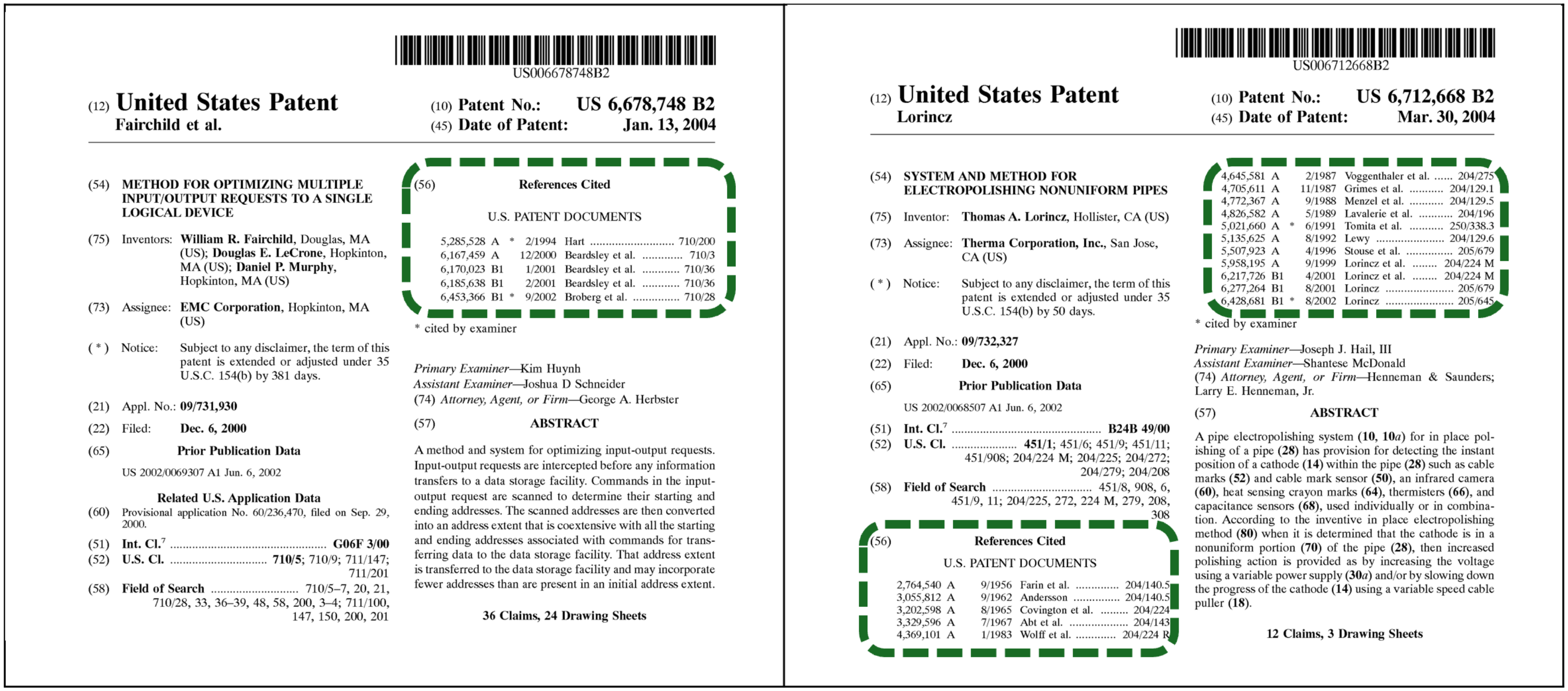
Note: Arrows point from citing domain to cited domain. Weight differences of arrows illustrate different frequencies of citations.

Figure A4. Heterogeneous Effects of Distant Recombination on Market Valuations Across Patent Classes in the Sample



Note: Each coefficient plot represents the estimate of the distant recombination coefficient in the baseline model (Model 1, Table 3) from a subsample of patents belonging to a specific class; classes with insufficient observations or yield outlier estimates (top/bottom 1%) are excluded for clearer comparisons; 95% CIs are reported; coefficient magnitude can be interpreted as the average percentage change in market valuation corresponding to the presence of distant recombination for patents in a given class.

Figure A5. Sample Patent Front Page Excerpts Showing Citation and Class Information


US006678748B2

(12) **United States Patent**
Fairchild et al.

(10) **Patent No.: US 6,678,748 B2**
(45) **Date of Patent: Jan. 13, 2004**

(54) **METHOD FOR OPTIMIZING MULTIPLE INPUT/OUTPUT REQUESTS TO A SINGLE LOGICAL DEVICE**

(75) Inventors: **William R. Fairchild**, Douglas, MA (US); **Douglas E. LeCrone**, Hopkinton, MA (US); **Daniel P. Murphy**, Hopkinton, MA (US)

(73) Assignee: **EMC Corporation**, Hopkinton, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 381 days.

(21) Appl. No.: **09/731,930**

(22) Filed: **Dec. 6, 2000**

(65) **Prior Publication Data**

US 2002/0069307 A1 Jun. 6, 2002

Related U.S. Application Data

(60) Provisional application No. 60/236,470, filed on Sep. 29, 2000.

(51) **Int. Cl.**[7] **G06F 3/00**
(52) **U.S. Cl.** **710/5**; 710/9; 711/147; 711/201
(58) **Field of Search** 710/5–7, 20, 21, 710/28, 33, 36–39, 48, 58, 200, 3–4; 711/100, 147, 150, 200, 201

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,285,528 A	*	2/1994	Hart	710/200
6,167,459 A		12/2000	Beardsley et al.	710/3
6,170,023 B1		1/2001	Beardsley et al.	710/36
6,185,638 B1		2/2001	Beardsley et al.	710/36
6,453,366 B1	*	9/2002	Broberg et al.	710/28

* cited by examiner

Primary Examiner—Kim Huynh
Assistant Examiner—Joshua D Schneider
(74) *Attorney, Agent, or Firm*—George A. Herbster

(57) **ABSTRACT**

A method and system for optimizing input-output requests. Input-output requests are intercepted before any information transfers to a data storage facility. Commands in the input-output request are scanned to determine their starting and ending addresses. The scanned addresses are then converted into an address extent that is coextensive with all the starting and ending addresses associated with commands for transferring data to the data storage facility. That address extent is transferred to the data storage facility and may incorporate fewer addresses than are present in an initial address extent.

36 Claims, 24 Drawing Sheets

US006712668B2

(12) **United States Patent**
Lorincz

(10) **Patent No.: US 6,712,668 B2**
(45) **Date of Patent: Mar. 30, 2004**

(54) **SYSTEM AND METHOD FOR ELECTROPOLISHING NONUNIFORM PIPES**

(75) Inventor: **Thomas A. Lorincz**, Hollister, CA (US)

(73) Assignee: **Therma Corporation, Inc.**, San Jose, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 50 days.

(21) Appl. No.: **09/732,327**

(22) Filed: **Dec. 6, 2000**

(65) **Prior Publication Data**

US 2002/0068507 A1 Jun. 6, 2002

(51) **Int. Cl.**[7] **B24B 49/00**
(52) **U.S. Cl.** .... **451/1**; 451/6; 451/9; 451/11; 451/908; 204/224 M; 204/225; 204/272; 204/279; 204/208
(58) **Field of Search** .... 451/8, 908, 6, 451/9, 11; 204/225, 272, 224 M, 279, 208, 308

(56) **References Cited**

U.S. PATENT DOCUMENTS

2,764,540 A		9/1956	Farin et al.	204/140.5
3,055,812 A		9/1962	Andersson	204/140.5
3,202,598 A		8/1965	Covington et al.	204/224
3,329,596 A		7/1967	Abt et al.	204/143
4,369,101 A		1/1983	Wolff et al.	204/224 R
4,645,581 A		2/1987	Voggenthaler et al.	204/275
4,705,611 A		11/1987	Grimes et al.	204/129.1
4,772,367 A		9/1988	Menzel et al.	204/129.5
4,826,582 A		5/1989	Lavalerie et al.	204/196
5,021,660 A	*	6/1991	Tomita et al.	250/338.3
5,135,625 A		8/1992	Lewy	204/129.6
5,507,923 A		4/1996	Stouse et al.	205/679
5,958,195 A		9/1999	Lorincz et al.	204/224 M
6,217,726 B1		4/2001	Lorincz et al.	204/224 M
6,277,264 B1		8/2001	Lorincz	205/679
6,428,681 B1	*	8/2002	Lorincz	205/645

* cited by examiner

Primary Examiner—Joseph J. Hail, III
Assistant Examiner—Shantese McDonald
(74) *Attorney, Agent, or Firm*—Henneman & Saunders; Larry E. Henneman, Jr.

(57) **ABSTRACT**

A pipe electropolishing system (**10**, **10***a*) for in place polishing of a pipe (**28**) has provision for detecting the instant position of a cathode (**14**) within the pipe (**28**) such as cable marks (**52**) and cable mark sensor (**50**), an infrared camera (**60**), heat sensing crayon marks (**64**), thermisters (**66**), and capacitance sensors (**68**), used individually or in combination. According to the inventive in place electropolishing method (**80**) when it is determined that the cathode is in a nonuniform portion (**70**) of the pipe (**28**), then increased polishing action is provided as by increasing the voltage using a variable power supply (**30***a*) and/or by slowing down the progress of the cathode (**14**) using a variable speed cable puller (**18**).

12 Claims, 3 Drawing Sheets

Note: Green circles are added by the authors. U.S. Patent No. 6678748 contains citations within the primary class only and is valued at 34.88 million dollars, while U.S. Patent No. 6712668 contains citations outside the primary class only and is valued at 0.41 million dollars. Median patent valuation for 2004 is 3.18 million dollars. Patent documents are downloaded from USPTO; patent valuations are based on the estimates provided by Kogan et al. (2017).

Table A1. Database construction and change in sample size

Data Source	Data	Sampling Conditions	Time Period	Number of Observations
PatentsView	Citations	Utility patents from all years, observations with missing information and duplicates dropped	1976-2020	103,790,190
	Patents	Utility patents from all years, observations with missing information dropped	1976-2020	6,901,418
		Data on USPC classes available till 2015 only	1976-2015	5,265,341
		Data from at least previous 10 years required for recombination variables	1986-2015	4,633,903
		Patents without class information dropped	1986-2015	4,453,499
KPSS	Market value	Matched with patent data	1986-2015	1,639,095
Compustat	Firm-level data	Matched with Compustat data	1986-2015	1,569,062
Current Sample		Dropped 955 singleton observations	1986-2015	1,568,107

Table A2. Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Market value	1													
2	Distant recombination	0.107	1												
3	Distant recombination of knowledge from specific domains	0.084	0.621	1											
4	Distant recombination of knowledge from general domains	0.131	0.629	0.190	1										
5	Distant recombination of knowledge from weakly connected domains	0.093	0.589	0.474	0.550	1									
6	Distant recombination of knowledge from strongly connected domains	0.108	0.609	0.452	0.439	0.117	1								
7	Local recombination	-0.075	-0.202	-0.140	-0.222	-0.125	-0.233	1							
8	Domain insularity	0.040	0.040	0.052	0.050	0.026	0.081	0.183	1						
9	Citation lag	0.104	0.170	0.118	0.172	0.228	0.032	-0.012	0.107	1					
10	Number of claims	0.145	0.085	0.087	0.101	0.078	0.101	-0.061	0.054	-0.004	1				
11	Use of science	0.144	0.076	0.102	0.108	0.069	0.137	-0.133	0.015	0.034	0.124	1			
12	Family size	0.035	0.030	0.042	0.047	0.032	0.052	-0.028	0.017	0.091	0.032	0.192	1		
13	Government sponsorship	0.044	0.011	0.010	0.012	0.020	-0.001	-0.011	-0.004	0.023	0.008	0.065	-0.013	1	
14	Percentage missing class	0.068	-0.102	-0.089	-0.083	-0.018	-0.151	0.172	-0.081	0.453	-0.087	-0.104	-0.012	0.011	1
15	Inventor team size	0.018	0.033	0.031	0.051	0.035	0.043	-0.068	-0.001	0.034	0.065	0.123	0.124	0.014	-0.074
16	Percentage female	0.045	-0.002	-0.004	0.005	-0.014	0.017	-0.050	-0.024	-0.017	0.012	0.084	0.036	0.003	-0.047
17	Star inventor	0.008	0.040	0.064	0.040	0.026	0.078	0.006	0.049	-0.032	0.083	0.115	0.099	-0.005	-0.074
18	Reliance of self-citation	-0.001	-0.006	-0.022	-0.019	-0.028	-0.004	0.050	0.086	-0.109	-0.005	-0.014	0.016	-0.001	-0.117
19	Patent count	-0.251	-0.014	-0.029	-0.018	-0.047	0.006	-0.107	-0.030	-0.130	-0.074	-0.090	-0.090	-0.008	-0.206
20	Sales growth	0.000	0.000	0.000	0.000	-0.002	0.002	0.005	0.001	0.000	0.003	0.007	0.004	0.000	0.000
21	R&D intensity	0.017	0.006	0.019	0.001	-0.035	0.053	-0.042	-0.013	-0.102	0.058	0.186	0.054	-0.007	-0.118
22	Financial slack	-0.002	0.014	0.033	0.013	0.003	0.039	-0.016	0.009	-0.054	0.078	0.102	0.030	-0.023	-0.057
23	Number of business segments	-0.095	0.003	-0.032	0.016	0.021	-0.038	-0.062	-0.023	0.085	-0.077	-0.086	-0.017	0.047	0.009
24	Firm age	-0.282	-0.020	-0.017	-0.017	-0.002	-0.040	-0.034	-0.001	0.061	-0.061	-0.070	-0.048	-0.013	-0.148
25	Inventor mobility	-0.317	-0.026	-0.018	-0.028	-0.028	-0.015	-0.028	-0.011	-0.039	-0.034	0.019	0.035	-0.023	-0.163
26	Degree of integration	0.161	0.008	0.025	0.006	0.019	0.012	0.077	0.025	0.047	0.071	0.078	0.059	-0.013	0.124
27	Geographical dispersion	0.127	0.005	-0.015	0.017	-0.015	0.015	-0.137	-0.037	0.067	0.005	0.059	0.072	-0.031	-0.108
28	Experience in the focal domain	0.113	-0.071	-0.038	-0.052	-0.085	0.011	0.200	0.091	0.003	0.102	0.136	0.098	-0.027	-0.053
29	Experience in the cited domains	0.135	0.246	0.312	0.231	0.235	0.279	-0.176	0.047	0.091	0.120	0.262	0.114	-0.010	-0.088
30	Domain growth	-0.007	-0.003	-0.001	-0.001	-0.065	0.079	-0.094	-0.020	-0.148	0.032	0.031	-0.008	-0.016	-0.125
31	Domain uncertainty	-0.004	-0.009	-0.008	-0.016	0.051	-0.084	0.125	0.023	0.099	-0.045	-0.051	-0.008	0.011	0.146
32	Domain size	-0.041	-0.001	0.008	0.007	-0.106	0.152	-0.124	-0.008	-0.167	0.086	0.136	0.045	-0.015	-0.340

		15	16	17	18	19	20	21	22	23	24	25	26	27	28
15	Inventor team size	1													
16	Percentage female	0.087	1												
17	Star inventor	0.186	0.003	1											
18	Reliance of self-citation	0.051	0.012	0.139	1										
19	Patent count	0.048	0.001	0.067	0.154	1									
20	Sales growth	0.002	0.000	0.002	-0.004	-0.023	1								
21	R&D intensity	0.014	0.024	0.031	-0.036	-0.134	0.021	1							
22	Financial slack	-0.009	-0.012	0.054	-0.019	-0.229	0.006	-0.109	1						
23	Number of business segments	0.034	-0.037	-0.065	0.050	0.435	-0.015	-0.262	-0.310	1					
24	Firm age	0.051	-0.042	-0.002	0.098	0.405	-0.014	-0.133	-0.113	0.322	1				
25	Inventor mobility	0.083	-0.017	-0.006	-0.065	0.062	0.001	0.065	0.025	0.091	0.276	1			
26	Degree of integration	-0.044	0.001	-0.037	-0.128	-0.796	0.014	0.162	0.227	-0.394	-0.336	-0.072	1		
27	Geographical dispersion	0.038	0.026	-0.052	-0.025	0.159	-0.009	0.046	-0.049	0.132	-0.033	-0.064	-0.130	1	
28	Experience in the focal domain	0.052	0.048	0.124	0.051	-0.338	0.021	0.140	0.178	-0.296	-0.137	-0.008	0.291	0.006	1
29	Experience in the cited domains	0.068	0.024	0.150	0.030	-0.116	0.003	0.072	0.105	-0.157	-0.093	-0.009	0.107	0.005	0.012
30	Domain growth	-0.011	0.008	0.021	-0.019	0.085	0.001	0.061	0.035	-0.042	-0.003	-0.023	-0.033	0.026	0.058
31	Domain uncertainty	-0.019	-0.021	-0.031	0.020	-0.083	0.002	-0.052	-0.028	0.036	-0.026	-0.007	0.039	-0.042	-0.068
32	Domain size	0.071	0.072	0.090	-0.016	0.192	-0.002	0.136	0.069	-0.031	0.072	0.115	-0.103	0.158	0.222

		29	30	31	32
29	Experience in the cited domains	1			
30	Domain growth	0.017	1		
31	Domain uncertainty	-0.042	-0.480	1	
32	Domain size	0.077	0.330	-0.314	1

Note: The grayed-out cells indicate that the variables in the corresponding rows and columns are disaggregated measures of the same construct that appear in separate models.

Table A3a. Analysis of short-run and long-run technological impact in all classes

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**	**Model 12**
	All years	*All years*	*All years*	*All years*	*Years 1 to 3*	*Years 1 to 3*	*Years 1 to 3*	*Years 1 to 3*	*Years 4 to 10*	*Years 4 to 10*	*Years 4 to 10*	*Years 4 to 10*
	All classes	*All classes*	*All classes*	*All classes*	*All classes*	*All classes*	*All classes*	*All classes*	*All classes*	*All classes*	*All classes*	*All classes*
Distant recombination	0.1205	0.1258			0.0489	0.0759			0.0969	0.0928		
	(0.0000)	(0.0000)			(0.0000)	(0.0000)			(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0104				-0.0509				0.0083		
		(0.4658)				(0.0000)				(0.5336)		
Distant recombination of knowledge from specific domains			0.0751				0.0345				0.0605	
			(0.0000)				(0.0000)				(0.0000)	
Distant recombination of knowledge from general domains			0.0966				0.0383				0.0822	
			(0.0000)				(0.0000)				(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.0889				0.0309				0.0692
				(0.0000)				(0.0000)				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0883				0.0440				0.0780
				(0.0000)				(0.0000)				(0.0000)
Local recombination	0.1421	0.1425	0.1394	0.1382	0.0841	0.0833	0.0824	0.0811	0.1189	0.1185	0.1158	0.1143
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		-0.1077	-0.1140	-0.1040		0.1508	0.1128	0.1192		0.1224	0.1307	0.1409
		(0.0000)	(0.0000)	(0.0000)		(0.0000)	(0.0000)	(0.0000)		(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107	1568107	1568107	1568107	1568107	1568107	1568107	1568107	1568107
R-squared	0.4065	0.4065	0.4072	0.4072	0.1807	0.1807	0.1811	0.1812	0.2882	0.2882	0.2888	0.2888

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class and grant year. In columns labeled as "All years," we measured the dependent variable, total technological impact, as the logged number of citations received during all years (till 2020) after a patent has been granted plus one; in columns labeled as "Years 1 to 3," the dependent variable, short-run impact, is measured as the logged number of citations received during the first three years after a patent has been granted plus one; in columns labeled as "Years 4 to 10," the dependent variable is long-run impact, which is the logged number of forward citations received between the fourth year and the tenth year plus one.

Table A3b. Analysis of short-run distant technological impact and long-run local technological impact

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
	Years 1 to 3	*Years 1 to 3*	*Years 1 to 3*	*Years 1 to 3*	*Years 4 to 10*	*Years 4 to 10*	*Years 4 to 10*	*Years 4 to 10*
	Outside class	*Outside class*	*Outside class*	*Outside class*	*Within class*	*Within class*	*Within class*	*Within class*
Distant recombination	0.1021	0.0335			-0.0686	0.0160		
	(0.0000)	(0.0000)			(0.0000)	(0.0014)		
Domain insularity × Distant recombination		0.1302				-0.1619		
		(0.0000)				(0.0000)		
Distant recombination of knowledge from specific domains			0.0692				-0.0535	
			(0.0000)				(0.0000)	
Distant recombination of knowledge from general domains			0.0671				-0.0341	
			(0.0000)				(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.0725				-0.0639
				(0.0000)				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0641				-0.0185
				(0.0000)				(0.0000)
Local recombination	-0.0364	-0.0351	-0.0385	-0.0392	0.2200	0.2199	0.2225	0.2209
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		-0.2060	-0.1040	-0.0997		-0.0768	-0.2033	-0.1963
		(0.0000)	(0.0000)	(0.0000)		(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107	1568107	1568107	1568107	1568107
R-squared	0.1648	0.1651	0.1667	0.1665	0.3876	0.3877	0.3879	0.3879

Note: In columns labeled as “Years 1 to 3,” we measured the dependent variable, short-run distant impact, as the logged number of citations received from outside the focal class during the first three years after a patent has been granted plus one; in columns labeled as “Years 4 to 10,” the dependent variable is long-run local impact, which is the logged number of forward citations received from within the focal class between the fourth year and the tenth year plus one.

Table A4. Robustness test of the market valuations of firm patents excluding volatile market conditions

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0098	0.0465		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0687		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0113	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0049	
			(0.0012)	
Distant recombination of knowledge from weakly connected domains				0.0062
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0094
				(0.0000)
Local recombination	0.0164	0.0141	0.0143	0.0140
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4947	0.4426	0.4433
		(0.0000)	(0.0000)	(0.0000)
Observations	1206442	1206442	1206442	1206442
R-squared	0.9010	0.9011	0.9011	0.9011

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and grant year. The dependent variable is the estimated market valuation of the patent.

Table A5. Robustness test of the market valuations of firm patents using grant date fixed effects

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0102	0.0445		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0636		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0124	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0046	
			(0.0007)	
Distant recombination of knowledge from weakly connected domains				0.0062
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0101
				(0.0000)
Local recombination	0.0140	0.0119	0.0120	0.0116
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4782	0.4299	0.4307
		(0.0000)	(0.0000)	(0.0000)
Observations	1568102	1568102	1568102	1568102
R-squared	0.8999	0.8999	0.8999	0.8999

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and patent grant date. The dependent variable is the estimated market valuation of the patent.

Table A6. Logit regression analysis of patent renewal decisions in the fourth year after patent grant

	Model 1	**Model 2**	**Model 3**	**Model 4**
Distant recombination	0.0393	0.1866		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.2239		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0868	
			(0.0000)	
Distant recombination of knowledge from general domains			-0.0109	
			(0.1096)	
Distant recombination of knowledge from weakly connected domains				0.0320
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0604
				(0.0000)
Local recombination	0.1826	0.1510	0.1501	0.1470
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.6964	0.5034	0.5416
		(0.0000)	(0.0000)	(0.0000)
Patent market valuation (logged)	0.0197	0.0213	0.0214	0.0209
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	1234272	1234272	1234272	1234272
chi2	27566.99	27944.37	27992.86	27980.31

Note: p-values in parentheses; all models include controls and fixed effects for grant year. The dependent variable is a dummy variable that switches to one for patents that are renewed in the fourth year after grant.

Table A7. Robustness test excluding patent classes associated with miscellaneous subcategories

	Model 1	**Model 2**	**Model 3**	**Model 4**
Distant recombination	0.0123	0.0259		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0250		
		(0.0393)		
Distant recombination of knowledge from specific domains			0.0142	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0048	
			(0.0014)	
Distant recombination of knowledge from weakly connected domains				0.0061
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0105
				(0.0000)
Local recombination	0.0158	0.0139	0.0139	0.0134
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4402	0.4212	0.4217
		(0.0000)	(0.0000)	(0.0000)
Observations	1322690	1322690	1322690	1322690
R-squared	0.8942	0.8942	0.8942	0.8942

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and grant year. The dependent variable is the estimated market valuation of the patent.

Table A8. Robustness test excluding patents whose technology classification has changed after grant

	Model 1	**Model 2**	**Model 3**	**Model 4**
Distant recombination	0.0112	0.0632		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0964		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0123	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0061	
			(0.0001)	
Distant recombination of knowledge from weakly connected domains				0.0058
				(0.0002)
Distant recombination of knowledge from strongly connected domains				0.0121
				(0.0000)
Local recombination	0.0179	0.0157	0.0161	0.0156
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.5393	0.4656	0.4670
		(0.0000)	(0.0000)	(0.0000)
Observations	1192528	1192528	1192528	1192528
R-squared	0.8963	0.8964	0.8964	0.8964

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and grant year. The dependent variable is the estimated market valuation of the patent.

Table A9. Robustness test excluding patents that have been assigned to more than one technology class

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0081	0.0488		
	(0.0001)	(0.0000)		
Domain insularity × Distant recombination		-0.0751		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0120	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0012	
			(0.5145)	
Distant recombination of knowledge from weakly connected domains				0.0030
				(0.0976)
Distant recombination of knowledge from strongly connected domains				0.0103
				(0.0000)
Local recombination	0.0211	0.0188	0.0191	0.0186
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4835	0.4317	0.4327
		(0.0000)	(0.0000)	(0.0000)
Observations	829589	829589	829589	829589
R-squared	0.9001	0.9001	0.9001	0.9001

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and grant year. The dependent variable is the estimated market valuation of the patent.

Table A10. Robustness test with the number of classes assigned to a patent as an additional control

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0106	0.0461		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0663		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0121	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0052	
			(0.0001)	
Distant recombination of knowledge from weakly connected domains				0.0062
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0105
				(0.0000)
Local recombination	0.0152	0.0132	0.0133	0.0129
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4823	0.4316	0.4327
		(0.0000)	(0.0000)	(0.0000)
Number of classes	-0.0009	-0.0004	-0.0008	-0.0007
	(0.2020)	(0.5561)	(0.2558)	(0.3380)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8955	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and grant year. The dependent variable is the estimated market valuation of the patent.

Table A11. Exploring the implications of the temporal change in a focal domain's average connectivity

	Model 1	Model 2	Model 3
Distant recombination of knowledge from strongly connected domains	0.0053	0.0053	0.0097
	(0.0001)	(0.0001)	(0.0000)
Distant recombination of knowledge from weakly connected domains	0.0107	0.0097	0.0093
	(0.0000)	(0.0000)	(0.0000)
Change in focal domain connectivity	0.9967	0.9150	1.1893
	(0.0000)	(0.0000)	(0.0000)
Distant recombination of knowledge from strongly connected domains × Change in focal domain connectivity		0.1285	0.1977
		(0.0203)	(0.0004)
Distant recombination of knowledge from weakly connected domains × Change in focal domain connectivity			-0.5285
			(0.0000)
Local recombination	0.0144	0.0144	0.0145
	(0.0000)	(0.0000)	(0.0000)
Observations	1545339	1545339	1545339
R-squared	0.8979	0.8979	0.8979

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. *Change in focal domain connectivity* is the change in the focal patent class's average connectivity strength to all other classes relative to the previous year and is multiplied by 1000 for scale.

Table A12. Robustness test including search scope and search depth as additional controls

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0099	0.0460		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0671		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0115	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0047	
			(0.0007)	
Distant recombination of knowledge from weakly connected domains				0.0058
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0098
				(0.0000)
Local recombination	0.0153	0.0131	0.0133	0.0129
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4838	0.4327	0.4336
		(0.0000)	(0.0000)	(0.0000)
Search scope	-0.4081	-0.4087	-0.4081	-0.4081
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Search depth	-0.0143	-0.0146	-0.0146	-0.0146
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8957	0.8957	0.8957	0.8957

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and grant year. The dependent variable is the estimated market valuation of the patent. Search depth and search scope are defined in accordance with Katila and Ahuja (2002).

Table A13. Supplemental analysis of technological distance of prior art

	Model 1	Model 2	Model 3
Technological distance of prior art	-0.0073	0.0402	
	(0.0011)	(0.0000)	
Domain insularity × Technological distance of prior art		-0.0915	
		(0.0000)	
Technological distance of prior art in specific domains			-0.0068
			(0.0111)
Technological distance of prior art in general domains			-0.0009
			(0.7144)
Number of backward citations (logged)	0.0148	0.0144	0.0140
	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4576	0.4232
		(0.0000)	(0.0000)
Observations	1568107	1568107	1568107
R-squared	0.8955	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. *Technological distance of prior art* is defined using the approach by Trajtenberg, Henderson, and Jaffe (1997) and Kaplan and Vakili (2015). The measure assumes the value of one if a patent has no backward citations.

Table A14. Supplemental analysis using count measures of distant recombination

	Model 1	**Model 2**	**Model 3**	**Model 4**
Distant recombination (logged count)	0.0051	0.0120		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination (logged count)		-0.0130		
		(0.0013)		
Distant recombination of knowledge from specific domains (logged count)			0.0049	
			(0.0000)	
Distant recombination of knowledge from general domains (logged count)			0.0004	
			(0.6135)	
Distant recombination of knowledge from weakly connected domains (logged count)				0.0016
				(0.0384)
Distant recombination of knowledge from strongly connected domains (logged count)				0.0040
				(0.0000)
Local recombination (logged count)	0.0105	0.0091	0.0095	0.0091
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4407	0.4221	0.4233
		(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8955	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. Distant recombination and local recombination are measured as the logged value of the count of citations in the corresponding category plus one.

Table A15. Supplemental analysis using the percentage measures of distant recombination

	Model 1	Model 2	Model 3	Model 4
Distant recombination	-0.0080	0.0152		
	(0.0001)	(0.0285)		
Domain insularity × Distant recombination (percentage)		-0.0417		
		(0.0032)		
Distant recombination of knowledge from specific domains (percentage)			-0.0025	
			(0.3282)	
Distant recombination of knowledge from general domains (percentage)			-0.0076	
			(0.0023)	
Distant recombination of knowledge from weakly connected domains (percentage)				-0.0048
				(0.0449)
Distant recombination of knowledge from strongly connected domains (percentage)				-0.0054
				(0.0400)
Number of backward citations (logged)	0.0154	0.0148	0.0148	0.0147
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4250	0.4047	0.4044
		(0.0000)	(0.0000)	(0.0000)
Observations	1501682	1501682	1501682	1501682
R-squared	0.8947	0.8948	0.8948	0.8948

Note: p-values based on robust errors in parentheses; all the models include firm, class and grant year fixed effects and the controls. The percentage measure of distant recombination is calculated as the number of backward citations outside the focal class divided by the total number of backward citations. The dependent variable is the market valuation of a patent.

Table A16. Robustness test of domain-related measures based on one-year window

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0103	0.0426		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0608		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0117	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0061	
			(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.0054
				(0.0001)
Distant recombination of knowledge from strongly connected domains				0.0114
				(0.0000)
Local recombination	0.0154	0.0132	0.0133	0.0128
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4241	0.3784	0.3801
		(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8955	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all models include controls and fixed effects for firm, class, and grant year. The dependent variable is the market valuation of a patent.

Table A17. The effects of distant recombination split by quartiles of domain generality and inter-domain connectivity.

	Model 1	**Model 2**
Distant Recombination: Generality Q1	0.0018	
	(0.2049)	
Distant Recombination: Generality Q2	0.0145	
	(0.0000)	
Distant Recombination: Generality Q3	0.0102	
	(0.0000)	
Distant Recombination: Generality Q4	-0.0059	
	(0.0000)	
Distant Recombination: Connectivity Q1		0.0067
		(0.0000)
Distant Recombination: Connectivity Q2		0.0028
		(0.0348)
Distant Recombination: Connectivity Q3		0.0041
		(0.0026)
Distant Recombination: Connectivity Q4		0.0135
		(0.0000)
Local Recombination	0.0132	0.0125
	(0.0000)	(0.0000)
Observations	1568107	1568107
R-squared	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. Q1-Q4 represents the first (bottom 25%), second (25th percentile to median), third (median to 75th percentile), and fourth (top 25%) quartiles in terms of domain generality or inter-domain connectivity. In accordance with the main analysis, distant recombination measures are dummies that switch to one if a focal patent has at least one backward citation in the corresponding quartile.

Table A18. Biprobit models of the effects of distant recombination on the likelihood of technological breakthroughs and market hits

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
	Market Hits				*Technological Breakthroughs*			
Distant recombination	0.0540	0.0264			0.1626	0.1412		
	(0.0000)	(0.3831)			(0.0000)	(0.0000)		
Domain insularity × Distant recombination		0.0801				0.0730		
		(0.1403)				(0.2199)		
Distant recombination of knowledge from specific domains			0.0636				0.1380	
			(0.0000)				(0.0000)	
Distant recombination of knowledge from general domains			0.0495				0.1224	
			(0.0000)				(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.0632				0.2043
				(0.0000)				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0180				0.0209
				(0.0040)				(0.0011)
Local recombination	-0.0012	-0.0282	-0.0323	-0.0297	0.0628	0.0335	0.0243	0.0337
	(0.8795)	(0.0004)	(0.0000)	(0.0002)	(0.0000)	(0.0001)	(0.0037)	(0.0001)
Domain insularity		0.2986	0.3812	0.3634		0.3512	0.4528	0.4037
		(0.0000)	(0.0000)	(0.0000)		(0.0000)	(0.0000)	(0.0000)
Observations	1528613	1528613	1528613	1528613	1528613	1528613	1528613	1528613
Chi-squared	33103.87	34133.70	34191.42	34375.14	33103.87	34133.70	34191.42	34375.14

Note: p-values in parentheses; all models include controls. *Market Hits* are the top 1% of the sample patents in the same class and grant year in terms of market valuation. *Technological Breakthroughs* are the top 1% most cited patents in the same cohort.

Table A19. Robustness test accounting for selection bias of distant recombination

First Stage	**Model 1-1**	**Model 2-1**	**Model 3-1a**	**Model 3-1b**	**Model 4-1a**	**Model 4-1b**
DV:	*Dist. Recomb. = 1*	*Dist. Recomb. = 1*	*Dist. Recomb. Specific = 1*	*Dist. Recomb. General. = 1*	*Dist. Recomb. Weakly Connected = 1*	*Dist. Recomb. Strongly Connected = 1*
Searchable distant front-page records: All	0.0035	0.0035				
	(0.0340)	(0.0340)				
Searchable distant full-text records: All	0.0065	0.0065				
	(0.0022)	(0.0022)				
Searchable distant front-page records: Specific			0.0034			
			(0.0254)			
Searchable distant full-text records: Specific			0.0049			
			(0.0113)			
Searchable distant front-page records: General				0.0007		
				(0.6502)		
Searchable distant full-text records: General				0.0022		
				(0.2567)		
Searchable distant front-page records: Weakly connected					0.0119	
					(0.0000)	
Searchable distant full-text records: Weakly connected					0.0192	
					(0.0000)	
Searchable distant front-page records: Strongly connected						-0.0023
						(0.1334)
Searchable distant full-text records: Strongly connected						-0.0026
						(0.1790)
Main Stage	**Model 1-2**	**Model 2-2**	**Model 3-2**		**Model 4-2**	
DV:	*Patent Valuation*					
Distant recombination	0.0115	0.0453				

	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0647		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0120	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0048	
			(0.0004)	
Distant recombination of knowledge from weakly connected domains				0.0062
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0100
				(0.0000)
Local recombination	0.0157	0.0133	0.0137	0.0128
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4844	0.4378	0.4414
		(0.0000)	(0.0000)	(0.0000)
IMR: All	0.0608	0.0855		
	(0.0086)	(0.0002)		
IMR: Specific			0.7205	
			(0.0000)	
IMR: General			-0.5380	
			(0.0000)	
IMR: Weakly connected				-0.4712
				(0.0000)
IMR: Strongly connected				0.6182
				(0.0000)
Post-digitalization grant	0.0623	0.0626	0.0621	0.0597
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	1568013	1568013	1568007	1566713
R-squared	0.8955	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; Models 1-1, 2-1, 3-1a, 3-1b, 4-1a, and 4-1b in the upper panel are Probit models with class and application year fixed effects. Models 1-2, 2-2, 3-2, and 4-2 in the lower panel are high-dimension fixed-effects models that follow the specifications of the main analysis in Table 3.

Searchable distant front-page records is calculated as the logged count of patent records that were granted outside the focal HJT subcategory in the twenty years preceding the focal patent application and were searchable solely through bibliographic data on the USPTO's website in the month prior to the focal patent application. This variable takes on non-negative values from the month following the 1995 rollout of the patent office's first online patent search tool until full-text data became accessible online in November 1998. *Searchable distant full-text records* is calculated as the logged count of the records of patents that were granted outside the focal HJT subcategory over the same twenty-year period and were searchable online through full-text data in the month prior to the focal application. This measure assumes a non-negative value starting in the month after full-text data became searchable on the USPTO website in 1998, and it counts records in both USPTO and Google Patents after the latter's launch in December 2006. These instruments capture all distant subcategories (i.e., all subcategories excluding the focal subcategory) subcategories in Models 1-1 and 2-1 (*All*), distant subcategories with below-median generality (*Specific*) in Model 3-1a, distant subcategories with above-median generality (*General*) in Model 3-1b, distant subcategories with below-median connectivity to all other subcategories (*Weakly connected*) in Model 4-1a, and distant subcategories with above-median connectivity (*Strongly connected*) in Model 4-1b. We added one to all count measures before taking the logarithm to accommodate zeroes. Correspondingly, the distant recombination dummy (*Dist. Recomb.*) is identified through all distant classes (i.e., all classes excluding the focal class) in Models 1-1 and 2-1 and restricted to specific classes in Model 3-1a, general classes in Model 3-1b, weakly connected classes in Model 4-1a, and strongly connected classes in Model 4-1b. Inverse Mills ratios (*IMR*) are then calculated from the corresponding first-stage selection models and included in the main-stage models. *Post-digitization grant* is a dummy that switches to one if a patent is granted after the USPTO first introduced the online search tool for patent grants in November 1995. Sample size differences are due to some cases being perfectly predicted by fixed effects in the first stage.

Table A20. Falsification Test of Instruments

	Model 1	Model 2
	Distant Recombination: Examiner Citation	*Distant Recombination: Non-examiner Citation*
Searchable distant front-page records	-0.0449	0.0593
	(0.0701)	(0.0000)
Searchable distant full-text records	-0.0448	0.0645
	(0.0722)	(0.0000)
Observations	981835	1008425

Note: p-values based on robust errors in parentheses; both models are Probit models with class and application year fixed effects and controls. Instruments are defined as in Model 1-1 in Table A19 above. In Model 1, the dependent variable is a dummy for distant recombination in a focal patent, which equals one if the backward citations outside the focal patent class are added by examiners only; in Model 2, the dependent variable is a dummy for distant recombination in a focal patent, which equals one if the backward citations outside the focal patent class are added by non-examiners only. Sample size differences are due to citation type data not being available for all years and exclusion of cases perfectly predicted by fixed effects.

Table A21. The effects of distant recombination on patent grant delay

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0532	0.0694		
	(0.0000)	(0.0000)		
Domain insularity × Distant recombination		-0.0306		
		(0.0125)		
Distant recombination of knowledge from specific domains			0.0444	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0616	
			(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.0519
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0403
				(0.0000)
Local recombination	0.0118	0.0112	0.0100	0.0100
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.1147	0.0963	0.0970
		(0.0000)	(0.0000)	(0.0000)
Observations	1568059	1568059	1568059	1568059
R-squared	0.2126	0.2126	0.2134	0.2131

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and all the controls. We first calculated grant lag as the number of days between the filing and grant dates of the focal patent. The dependent variable, patent grant delay, is the difference between the focal patent's grant lag and the average grant lag in the same patent class and 2-digit SIC industry, normalized by the standard deviation of the same class-industry pair.

Table A22. Mechanism test: The moderating effects of firm-level intentionality toward distant recombination on the relationship between distant recombination and patent valuation.

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0057	0.0458		
	(0.0008)	(0.0000)		
Distant recombination × Firm tendency	-0.1915	-0.2300		
	(0.0000)	(0.0000)		
Distant recombination × Domain insularity		-0.0752		
		(0.0000)		
Distant recombination × Domain insularity × Firm tendency		0.0559		
		(0.4511)		
Distant recombination of knowledge from specific domains			0.0104	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0031	
			(0.0224)	
Distant recombination of knowledge from specific domains × Firm tendency			-0.0972	
			(0.0000)	
Distant recombination of knowledge from general domains × Firm tendency			-0.1773	
			(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.0043
				(0.0014)
Distant recombination of knowledge from strongly connected domains				0.0088
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Firm tendency				-0.2643
				(0.0000)
Distant recombination of knowledge from strongly connected domains × Firm tendency				-0.0291
				(0.1245)
Local recombination	0.0157	0.0135	0.0140	0.0136
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Firm tendency	0.2724	0.2837	0.2848	0.2921
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4899	0.4330	0.4334
		(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8955	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. *Firm tendency* measures the patenting firm's inclination toward distant recombination relative to its peers in the same 3-digit SIC industry. It is calculated as the difference between the firm-level average and industry-level average of the proportion of distant backward citations in patent grants in the same year. The dependent variable is the market valuation of a patent.

Table A23. Disentangling substance effects from the presence effects of distant recombination on patent valuation.

	Model 1	Model 2	Model 3	Model 4
Distant recombination – low similarity	-0.0120	-0.0277		
	(0.0000)	(0.0000)		
Distant recombination – high similarity	0.0000	0.0128		
	(0.9748)	(0.0113)		
Distant recombination – low similarity × Domain insularity		0.0319		
		(0.0010)		
Distant recombination – high similarity × Domain insularity		-0.0256		
		(0.0100)		
Distant recombination of knowledge from specific domains – low similarity			-0.0077	
			(0.0000)	
Distant recombination of knowledge from general domains – low similarity			-0.0129	
			(0.0000)	
Distant recombination of knowledge from specific domains – high similarity			0.0031	
			(0.0317)	
Distant recombination of knowledge from general domains – high similarity			-0.0041	
			(0.0049)	
Distant recombination of knowledge from weakly connected domains – low similarity				-0.0143
				(0.0000)
Distant recombination of knowledge from strongly connected domains – low similarity				-0.0053
				(0.0005)
Distant recombination of knowledge from weakly connected domains – high similarity				-0.0052
				(0.0004)
Distant recombination of knowledge from strongly connected domains – high similarity				0.0034
				(0.0233)
Local recombination	0.0132	0.0133	0.0135	0.0127
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity	0.4265	0.4251	0.4262	0.4279
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8956	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. The similarity score for each type of backward citations is constructed using the measure developed by Arts, Cassiman, and Gomez (2018)[i] and represents the average textual similarity between a focal patent's title and abstract and those of its cited patents. For each type of distant recombination, we use the median similarity score calculated within the focal patent's class and grant year as the cutoff to create two dummy variables representing high- and low-similarity groups.

[i] Arts, S., Cassiman, B., & Gomez, J. C. (2018). Text matching to measure patent similarity. *Strategic Management Journal*, *39*(1), 62-84.

Table A24. Mechanism test: The accentuating effects of information crowding on the relationship between distant recombination and patent valuation

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0071	0.0381		
	(0.0000)	(0.0000)		
Distant recombination × Crowding on grant date	0.0020	0.0084		
	(0.0000)	(0.0000)		
Distant recombination × Domain insularity		-0.0554		
		(0.0000)		
Distant recombination × Domain insularity × Crowding on grant date		-0.0143		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0087	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0043	
			(0.0026)	
Distant recombination of knowledge from specific domains × Crowding on grant date			0.0026	
			(0.0000)	
Distant recombination of knowledge from general domains × Crowding on grant date			-0.0000	
			(0.9165)	
Distant recombination of knowledge from weakly connected domains				0.0056
				(0.0001)
Distant recombination of knowledge from strongly connected domains				0.0073
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Crowding on grant date				-0.0009
				(0.0060)
Distant recombination of knowledge from strongly connected domains × Crowding on grant date				0.0034
				(0.0000)
Local recombination	0.0167	0.0145	0.0148	0.0144
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Crowding on grant date	-0.0197	-0.0197	-0.0196	-0.0200
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4458	0.4261	0.4270
		(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8963	0.8963	0.8963	0.8963

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. *Crowding on grant date* is defined as the sum of the number of grants in the same patent class and 3-digit SIC industry, plus the number of public firms receiving patents in the same 3-digit SIC industry on the focal patent's grant date, net of the median value for all grant dates in the same year; results are similar when the measure is based on patent grant count only.

Table A25. The implications of AIPA-enabled early disclosure for the relationship between distant recombination and patent valuation

	Model 1	**Model 2**	**Model 3**	**Model 4**
Distant recombination	0.0333	0.0734		
	(0.0000)	(0.0000)		
Distant recombination × Early disclosure	-0.0150	-0.0364		
	(0.0042)	(0.0001)		
Distant recombination × Domain insularity		-0.0817		
		(0.0000)		
Distant recombination × Domain insularity × Early disclosure		0.0425		
		(0.0167)		
Distant recombination of knowledge from specific domains			0.0240	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0071	
			(0.0987)	
Distant recombination of knowledge from specific domains × Early disclosure			-0.0159	
			(0.0004)	
Distant recombination of knowledge from general domains × Early disclosure			0.0013	
			(0.7761)	
Distant recombination of knowledge from weakly connected domains				0.0099
				(0.0187)
Distant recombination of knowledge from strongly connected domains				0.0268
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Early disclosure				-0.0026
				(0.5468)
Distant recombination of knowledge from strongly connected domains × Early disclosure				-0.0197
				(0.0000)
Local recombination	0.0238	0.0248	0.0227	0.0210
	(0.0000)	(0.0000)	(0.0001)	(0.0003)
Local recombination × Early disclosure	-0.0082	-0.0101	-0.0078	-0.0062
	(0.1742)	(0.0986)	(0.1951)	(0.3041)
Early disclosure	0.0112	0.0144	0.0080	0.0107
	(0.0970)	(0.0340)	(0.2032)	(0.0873)
Domain insularity		0.2729	0.2362	0.2391
		(0.0000)	(0.0000)	(0.0000)
Observations	900340	900340	900340	900340
R-squared	0.9201	0.9201	0.9201	0.9201

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and all the controls. The dependent variable is the estimated market valuation of the patent. *Early disclosure* is a dummy that switches to one if a focal patent is subject to the AIPA doctrine and is granted after the 18-month publication deadline. The sample includes post-AIPA patents only.

Table A26. Mechanism test: The attenuating effects of patent keyword similarity on the relationship between distant recombination and patent valuation

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0102	0.0507		
	(0.0000)	(0.0000)		
Distant recombination × Patent keyword similarity	0.5423	-0.0402		
	(0.0000)	(0.8649)		
Distant recombination × Domain insularity		-0.0765		
		(0.0000)		
Distant recombination × Domain insularity × Patent keyword similarity		1.4224		
		(0.0008)		
Distant recombination of knowledge from specific domains			0.0121	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0049	
			(0.0004)	
Distant recombination of knowledge from specific domains × Patent keyword similarity			0.2618	
			(0.0128)	
Distant recombination of knowledge from general domains × Patent keyword similarity			0.3881	
			(0.0003)	
Distant recombination of knowledge from weakly connected domains				0.0061
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0104
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Patent keyword similarity				0.4689
				(0.0000)
Distant recombination of knowledge from strongly connected domains × Patent keyword similarity				0.2138
				(0.0396)
Local recombination	0.0157	0.0135	0.0137	0.0133
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Patent keyword similarity	-0.9042	-0.9614	-0.8580	-0.8574
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4895	0.4296	0.4304
		(0.0000)	(0.0000)	(0.0000)
Observations	1567827	1567827	1567827	1567827
R-squared	0.8955	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. *Patent keyword similarity* is the backward cosine similarity measure estimated by Arts, Hou, and Gomez[ii] (2021). The measure captures the average degree of similarity between the keywords in a focal patent's text and those in patents filed over the previous five years.

[ii] Arts, S., Hou, J., & Gomez, J. C. (2021). Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures. *Research Policy*, *50*(2), 104144.

Table A27. Mechanism test: The attenuating effects of product market proximity on the relationship between distant recombination and patent valuation

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0056	-0.0063		
	(0.0086)	(0.3820)		
Distant recombination × Product market proximity	-0.0481	-0.4765		
	(0.4841)	(0.0007)		
Distant recombination × Domain insularity		0.0239		
		(0.0894)		
Distant recombination × Domain insularity × Product market proximity		0.8587		
		(0.0006)		
Distant recombination of knowledge from specific domains			0.0076	
			(0.0000)	
Distant recombination of knowledge from general domains			-0.0006	
			(0.7066)	
Distant recombination of knowledge from specific domains × Product market proximity			-0.0917	
			(0.1534)	
Distant recombination of knowledge from general domains × Product market proximity			0.1389	
			(0.0292)	
Distant recombination of knowledge from weakly connected domains				-0.0022
				(0.1812)
Distant recombination of knowledge from strongly connected domains				0.0081
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Product market proximity				0.0999
				(0.1274)
Distant recombination of knowledge from strongly connected domains × Product market proximity				-0.0693
				(0.2676)
Local recombination	0.0096	0.0096	0.0093	0.0087
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Product market proximity	-0.3618	-0.3498	-0.4233	-0.3943
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.0275	0.0466	0.0471
		(0.3367)	(0.0725)	(0.0699)
Observations	867084	867084	867084	867084
R-squared	0.7562	0.7562	0.7562	0.7562

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. *Product market proximity* is the average product similarity score between the patenting firm and other competitors in the industry in the grant year, net of the sample median. The similarity scores are estimated by Hoberg and Phillips (2016)[iii] and are available for a subset of the sample patents.

[iii] Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), 1423-1465.

Table A28. Testing the interaction effects between local recombination and distant recombination

	Model 1	**Model 2**	**Model 3**
Distant recombination	0.0193		
	(0.0000)		
Distant recombination × Local recombination	-0.0107		
	(0.0081)		
Distant recombination of knowledge from specific domains		0.0147	
		(0.0000)	
Distant recombination of knowledge from general domains		0.0151	
		(0.0000)	
Distant recombination of knowledge from specific domains × Local recombination		-0.0031	
		(0.3488)	
Distant recombination of knowledge from general domains × Local recombination		-0.0120	
		(0.0003)	
Distant recombination of knowledge from weakly connected domains			0.0201
			(0.0000)
Distant recombination of knowledge from strongly connected domains			0.0096
			(0.0015)
Distant recombination of knowledge from weakly connected domains × Local recombination			-0.0167
			(0.0000)
Distant recombination of knowledge from strongly connected domains × Local recombination			0.0015
			(0.6484)
Local recombination	0.0240	0.0222	0.0218
	(0.0000)	(0.0000)	(0.0000)
Domain insularity		0.4317	0.4329
		(0.0000)	(0.0000)
Observations	1568107	1568107	1568107
R-squared	0.8955	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent.

Table A29. Mechanism test: The effects of distant recombination on patent valuation under different appropriability conditions

	Model 1	**Model 2**	**Model 3**	**Model 4**
Distant recombination	0.0287	0.1124		
	(0.0000)	(0.0000)		
Distant recombination × Low-appropriability industry	-0.0098	-0.1197		
	(0.0103)	(0.0000)		
Distant recombination × Domain insularity		-0.1592		
		(0.0000)		
Distant recombination × Domain insularity × Low-appropriability industry		0.2216		
		(0.0000)		
Distant recombination of knowledge from specific domains			0.0158	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0245	
			(0.0000)	
Distant recombination of knowledge from specific domains × Low-appropriability industry			0.0033	
			(0.3047)	
Distant recombination of knowledge from general domains × Low-appropriability industry			-0.0274	
			(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.0249
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0170
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Low-appropriability industry				-0.0308
				(0.0000)
Distant recombination of knowledge from strongly connected domains × Low-appropriability industry				0.0015
				(0.6432)
Local recombination	0.0281	0.0242	0.0239	0.0232
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Local recombination × Low-appropriability industry	-0.0028			
	(0.4868)			
Domain insularity		0.5439	0.5020	0.5043
		(0.0000)	(0.0000)	(0.0000)
Observations	1052622	1052622	1052622	1052622
R-squared	0.8850	0.8851	0.8851	0.8851

Note: p-values based on robust errors in parentheses. The dependent variable is the estimated market valuation of the patent. *Low-appropriability industry* is a dummy that switches to one if the patenting firm is in an industry with below-median effective scores for patenting as an appropriability mechanism for product and process innovations (Cohen, Nelson & Walsh, 2000). All the models include firm, class, and grant year fixed effects and the controls.

Table A30. Falsification test of the technological impact of firm patents under different appropriability conditions

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.1406	0.1342		
	(0.0000)	(0.0000)		
Distant recombination × Low-appropriability industry	-0.0352	-0.0230		
	(0.0000)	(0.0239)		
Distant recombination × Domain insularity		0.0138		
		(0.4744)		
Distant recombination × Domain insularity × Low-appropriability industry		-0.0271		
		(0.1374)		
Distant recombination of knowledge from specific domains			0.0883	
			(0.0000)	
Distant recombination of knowledge from general domains			0.1045	
			(0.0000)	
Distant recombination of knowledge from specific domains × Low-appropriability industry			-0.0237	
			(0.0000)	
Distant recombination of knowledge from general domains × Low-appropriability industry			-0.0188	
			(0.0000)	
Distant recombination of knowledge from weakly connected domains				0.1040
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.1012
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Low-appropriability industry				-0.0253
				(0.0000)
Distant recombination of knowledge from strongly connected domains × Low-appropriability industry				-0.0338
				(0.0000)
Local recombination	0.1620	0.1446	0.1413	0.1402
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Local recombination × Low-appropriability industry	-0.0352			
	(0.0000)			
Domain insularity		-0.0163	-0.0178	-0.0101
		(0.5989)	(0.5278)	(0.7192)
Observations	1052622	1052622	1052622	1052622
R-squared	0.4113	0.4113	0.4119	0.4119

Note: p-values based on robust errors in parentheses. The dependent variable is the logged count of total forward citations received by the focal patent plus one. *Low-appropriability industry* is a dummy that switches to one if the patenting firm is in an industry with below-median effective scores for patenting as an appropriability mechanism for product and process innovations (Cohen, Nelson & Walsh, 2000). All the models include firm, class, and grant year fixed effects and the controls.

Table A31. Mechanism test: The accentuating effects of short-term ownership on the relationship between distant recombination and patent valuation

	Model 1	Model 2	Model 3	Model 4
Distant recombination	0.0101	0.0445		
	(0.0000)	(0.0000)		
Distant recombination × Short-term ownership	0.0055	0.0330		
	(0.7319)	(0.2585)		
Distant recombination × Domain insularity		-0.0633		
		(0.0000)		
Distant recombination × Domain insularity × Short-term ownership		-0.0733		
		(0.1574)		
Distant recombination of knowledge from specific domains			0.0113	
			(0.0000)	
Distant recombination of knowledge from general domains			0.0068	
			(0.0000)	
Distant recombination of knowledge from specific domains × Short-term ownership			0.0154	
			(0.2396)	
Distant recombination of knowledge from general domains × Short-term ownership			-0.0415	
			(0.0018)	
Distant recombination of knowledge from weakly connected domains				0.0086
				(0.0000)
Distant recombination of knowledge from strongly connected domains				0.0091
				(0.0000)
Distant recombination of knowledge from weakly connected domains × Short-term ownership				-0.0624
				(0.0000)
Distant recombination of knowledge from strongly connected domains × Short-term ownership				0.0315
				(0.0152)
Local recombination	0.0154	0.0133	0.0135	0.0131
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Short-term ownership	-0.0307	-0.0232	-0.0082	-0.0060
	(0.0607)	(0.1580)	(0.5672)	(0.6742)
Domain insularity		0.4826	0.4321	0.4333
		(0.0000)	(0.0000)	(0.0000)
Observations	1568107	1568107	1568107	1568107
R-squared	0.8955	0.8956	0.8956	0.8956

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the estimated market valuation of the patent. *Short-term ownership* is defined as the percentage of shares owned by transient institutional investors, who are characterized by high portfolio turnover and short horizons, net of the shares owned by dedicated institutions that are long-term oriented and have low portfolio turnover, per the classification method and data by Bushee (2001).

Table A32. The effects of distant recombination on long-term patent enforcement

	Panel A: Dummy Measure				Panel B: Count Measure			
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Distant recombination	-0.0002	-0.0016			0.0009	-0.0004		
	(0.4855)	(0.0909)			(0.0000)	(0.4983)		
Domain insularity × Distant recombination		0.0027				0.0027		
		(0.1345)				(0.0154)		
Distant recombination of knowledge from specific domains			0.0003				0.0011	
			(0.2662)				(0.0000)	
Distant recombination of knowledge from general domains			0.0000				0.0006	
			(0.9812)				(0.0005)	
Distant recombination of knowledge from weakly connected domains				0.0005				0.0011
				(0.0235)				(0.0000)
Distant recombination of knowledge from strongly connected domains				-0.0001				0.0005
				(0.7701)				(0.0019)
Local recombination	0.0001	0.0001	0.0001	0.0002	0.0008	0.0008	0.0007	0.0008
	(0.6410)	(0.6231)	(0.6268)	(0.5621)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Market value	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	(0.0553)	(0.0553)	(0.0559)	(0.0553)	(0.0571)	(0.0564)	(0.0563)	(0.0559)
Domain insularity		0.0001	0.0024	0.0022		-0.0028	0.0014	0.0012
		(0.9866)	(0.6994)	(0.7181)		(0.6499)	(0.8166)	(0.8415)
Observations	541515	541515	541515	541515	541515	541515	541515	541515
R-squared	0.0883	0.0883	0.0883	0.0883	0.0884	0.0884	0.0885	0.0885

Note: p-values based on robust errors in parentheses; all the models include firm, class, and grant year fixed effects and the controls. The dependent variable is the natural logarithm of the count of infringement lawsuits filed by the firm on the basis of the focal patent against other parties within the 10-year period following the grant, plus one. The information on infringement suits is collected from the USPTO's Patent Litigation Docket Reports Data (Toole, Miller & Sichelman, 2024).[iv] Panel A reports the results using the dummy measures for knowledge recombination in the main analysis. Panel B reports the results based on count measures for knowledge recombination, which are calculated as the logged value of the number of citations in a corresponding recombination category plus one (see Table A14). The sample period is set to 2003-2010 to minimize left and right truncation issues and to ensure comprehensive coverage of litigation records in the time frame.

[iv] Toole, A., Miller, R., & Sichelman, T. M. (2024). Technical Documentation for Patent Litigation Docket Reports Data, 1963-2020. USPTO Economic Working Paper No. 2024-1.

Appendix B. Summary of Qualitative Evidence Corroborating Investor Evaluation of Patents

Our qualitative work spans multiple sources, including media archives, investor community discussions, analyst reports, and firm communications. Through these comprehensive efforts, we have compiled a list of relevant and representative quotes from industry experts and presented them in Table B1 below. Many important findings emerge from these inquiries. First, patent evaluation is generally considered important but challenging due to a combination of factors. Industry insiders generally view patents as intangible assets whose value is uncertain, speculative, but increasingly crucial (Quotes A2 and A4). Meanwhile, there is no consensus on valuation methodologies, which come with significant constraints (Quote A26). Furthermore, additional disclosure on specific patents is rare (Quote A1).

These insights are supported by our own survey of media coverage for 170 sample patents. We have tabulated the counts of news articles and press releases for these patents in Factiva and reported the statistics in Table B2. We find no substantial difference in the scant news coverage received by the most and least valued patents or between patents built upon only local and only distant recombination. As part of this analysis, we manually reviewed the news articles and found that in almost all cases they reiterate the information that is already disclosed in the patent. Thus, it seems unlikely that investors would bypass the original grant document, which remains the primary source of information for assessing new patents and one of the few external channels for firm R&D disclosure.

Furthermore, patent citations and class information are highly visible and considered valuable signals for stock prices by market participants. As noted by patent experts, citations are displayed prominently in the reference section on the front page of each patent document (Quote B1; also see Figure A5 for illustrative examples) and thus are evident to investors who use patent references to judge the value of the patented idea (Quotes B4 and B6). We have identified multiple anecdotal cases in which citation and class information appears to have been utilized to inform patent valuation and related investment decisions (Quotes C3, C7, and C8). These instances include the chief executive of a patent financing firm calling patent references "a good look at how good the patent might be" (Quote B1) and Boston Consulting Group leveraging patent citation data to identify the domain expertise and value for a corporate client (Quote C3). In line with this, we have further uncovered first-hand accounts suggesting that professional market participants have used patent citations to inform stock trading decisions. MDB Capital Group, an investment banking firm, disclosed in a research report that the firm analyzed patent data on a weekly basis and developed valuation metrics based on firm's patent citations and technological areas (Quote B2). Similarly, CHI Research Inc., a consulting firm, stated that some of its investment decisions were based upon firm patent citation information and updated monthly (Quote B3).

Table B1. Summary of key quotes from industry experts

Quote No.	Source	Quotes
	Theme A: Practical Challenges of Patent Evaluation	
A1	MDB Capital Group (2011)[v], an investment banking firm	• "Despite the potential significance of intellectual property to an individual public company's market value, **huge inefficiencies remain in the valuation of these assets**. These inefficiencies arise primarily from the lack of adequate IP disclosure. For example, **in regulatory filings the IP discussion typically consists of just a single sentence** listing the number of granted and pending patents and **in most corporate investor presentations the existing IP is detailed in a simple bullet point or two**."
A2	Nasdaq (2021)[vi], a major stock exchange	• "But a problem arises when investors try to concretely account for the innovativeness of a particular company: **innovation is intangible and thus difficult to evaluate**…As a leading indicator of future success, patent value estimates can help investors decide the value of a stock and augment their strategies with more relevant and accurate information."
A3	Krajec (2016)[vii], CEO of patent finance company BlueIron IP	• "**Good patents are clear, straightforward, and easy to read**…Good patents have real, solid commercial value. **The value of a patent only comes when it captures commercial value - not when it captures some cool technology**."
A4	Morah (2023)[viii] for *Investopedia*, a financial media outlet	• "As **intangible assets**, **patents present a challenge in terms of valuation**, but they can be pivotal in determining a company's success – and the success of investors who buy these companies' stocks."
A5	UnitedLex (2024)[ix], a legal consulting firm	• "Patent valuation is an important aspect of the patent portfolio management process. **The details that contribute to the value of a given patent might be complicated and not always obvious**…every instance comes with its unique parameters and challenges. **For every patent, there will be a separate evaluation**."
A6	Eqvista (2024)[x], an equity management software firm	• "**Valuing a patent is a challenging process** involving technology, finance, economics, and law expertise." • "…**Challenging to accurately quantify indirect costs** like opportunity costs and risk premiums [for cost-based valuation of patent]…**Challenging to find truly comparable patents** due to differences in scope, technology, geographic coverage, and other factors [for market-based valuation]…**Obtaining accurate and reliable data for cash flow projections and risk assessments can be challenging**, particularly for new or emerging technologies…Option-based method of patent valuation can be **challenging**..."
	Theme B: Practical Implications of Patent Citations	

[v] MDB Capital Group. 2011. "iGo Inc (Nasdaq: IGOI) – A Four Star PatentVest IP Value Company 2011 "Best And Brightest" Patentvest Reports". S&P Capital IQ.
[vi] https://blog.data.nasdaq.com/investing-in-innovators-patent-valuation
[vii] Krajec, R. 2016. "Investing in Patents—Everything Startup Investors Need to Know about Patents". BlueIron Press.
[viii] https://www.investopedia.com/articles/fundamental-analysis/09/valuing-patent.asp
[ix] https://unitedlex.com/insights/increase-the-value-of-your-patent-portfolio-with-patent-valuation/
[x] https://eqvista.com/intangible-asset-valuation/value-patents/

B1	Krajec (2023)[xi], CEO of patent finance company BlueIron	• "**One easy way to spot an above-average patent is the References Cited on the front page. The references…are a good look at how good the patent might be**…There are many services that provide some kind of scoring system…of assessing patent value or quality. Virtually every one of the scoring systems **uses the references as a key indicator of quality**."
B2	Algaba, Hickman & Conley (2010)[xii] and MDB Capital Group (2011)[xiii], an investment banking firm	• "In recent years, **patent data mining and bibliometric analysis have become increasingly important fields of competitive intelligence for investors seeking to generate excess returns**." • "With its proprietary PatentVest (PV) database platform, MDB Capital has developed a set of proprietary tools/metrics to reliably identify and quantify leading innovation…**enabling price-discovery and a more appropriate valuation of patent assets…**The PatentVest analytic engine automatically assigns a number of proprietary IP related metrics to each company, **using patent information updated weekly from the USPTO** and...financial market data..." • "Tech Score…measures the impact of a company's patents. The metric is based upon **an analysis of the citation ratio of a company's patents**…" • "Depth Rating…measures a patent estates' degree of concentration…**within an area of technology**." • "Isolation Rating…measures the degree of isolation…of a company's patent estate… **Isolation is measured by looking at the percentage of self-citations versus citations by non-affiliated entities** and reflects the degree of how strongly a company's technology is related to that of its peers."
B3	Narin, Breitzman & Thomas (2004)[xiv] for CHI Research Inc, a management consulting firm	• "CHI Research, Inc. (CHI) has developed two different, but related, **models which use patent indicators to forecast stock price movements**…The second model…**identifies companies which have both strong patent citation indicators** and a low stock market valuation." • "In order to place a value on companies based on their patent portfolios, these portfolios were evaluated…using a number of quantitative patent indicators. **Science Linkage is**…**calculated on the basis of the average number of references on a company's patents to scientific papers**, as distinct from references to previous patents…**Technology Cycle Time…is a measure of the median age of the US patents cited on the front page of a company's patents**." • "In this paper we have shown **a strong association between the quality of companies' patent portfolios, as measured by patent citation indicators, and their stock market performance**…A particularly interesting aspect of this finding is that the number of patents a company holds is not a significant predictor of its performance. **The important factor is the quality of a company's patents, rather than their number.**"
B4	Caldwell & Troyer (2017)[xv] for *IP Watchdog*, an IP news site	• "**Citation analysis has been routinely used for several decades to help assess patent quality** and determine how patents impact the competition and affect the world at large…**All citations begin life as backward citations**…When a patent grants, **citations are listed on the front page of the published document.**"

[xi] https://blueironip.com/patent-searches/
[xii] MDB Capital Group. 2010. "PatentVest Tech Score Correlates to Business Performance and Alpha in Micro-Cap Asset Class". S&P Capital IQ.
[xiii] MDB Capital Group. 2011. "iGo Inc (Nasdaq: IGOI) – A Four Star PatentVest IP Value Company 2011 "Best And Brightest" Patentvest Reports". S&P Capital IQ.
[xiv] Narin, F., Breitzman, A., & Thomas, P. (2004). Using patent citation indicators to manage a stock portfolio. In *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T systems* (pp. 553-568). Dordrecht: Springer Netherlands.
[xv] https://ipwatchdog.com/2017/01/02/value-hidden-citations-patent-evaluation/id=76138/

		•
B5	Minesoft (2024)[xvi], an IP service firm	• "…there are many benefits to **the use of citations for patent analysis and competitive and business intelligence**…Patent citations can **bring to your attention related patents and ideas that align with your own innovations**…In essence, patent citations serve as a valuable resource for strategic decision-making, innovation tracking, and competitive positioning in the business world."
B6	KIA Investment Research (2015)[xvii] quoting Envision IP, an IP analytics firm	• "A reverse citation is a prior art reference considered by the USPTO…**A high reverse citation count indicates that the USPTO has deemed the patent claims novel over a larger number of prior art references**. This presumably provides difficulty in identifying previously unknown prior art during an invalidity challenge." • "BlackBerry's patents also have a **relatively strong average reverse citation count**, with an average of 32 reverse citations per patent…this data suggests that BlackBerry's patents (in general) may stand a **higher chance of surviving an invalidity challenge** than…some of its peers."
	Theme C: Practical Implications of Patent Classes	
C1	Oldham (2022)[xviii] for World Intellectual Property Organization	• "**Back citations may be from a wide range of technology fields in the prior art**…Attention is required to International Patent Classification/Cooperative Patent Classification codes to determine how close cited inventions are to each other…**citation networks and classification codes can be used for the identification of technology paths or trajectories**…citation analysis combined with classification and citation metrics is an important field of research that increasingly promises to make navigating citation networks significantly easier."
C2	United States Patent and Trademark Office (2021)[xix]	• "Three major types of prior art patent searches…Text searching, **Patent classification searching, and Patent citation searching**. A complete patent search will involve all three of these methods."
C3	Manly et al. (2024) [xx] for Boston Consulting Group, a management consulting firm	• "…in 2011 **BCG's Center for Growth and Innovation Analytics (GIA) helped a leading semiconductor equipment company that was facing slowing market growth identify an unexpected adjacency by analyzing patent-citation networks**. The company's expertise in moving materials in a vacuum at ultra-low temperatures could be of huge value in human tissue handling…**By mining**…**patent data**, GIA…accelerated the company's successful move into life sciences." • "…**Aptiv defined clear innovation domains** for both hardware and software. These included central vehicle controllers, high-voltage power electronics, radars, cameras, and their related software…**an organization needs to focus its innovation investment on a focused list of domains** consistent with its business strategy."
C4	Singh (2020)[xxi] for Copperpod Intellectual Property, an IP service firm	• "While searching for a prior art, a combination of two patents…can be considered…to render a claimed invention obvious… in a very broad sense, **the combination can be said to be obvious if both the references are from same field;** even better, if they both address the same problem.

[xvi] https://minesoft.com/the-power-of-patent-citations
[xvii] https://seekingalpha.com/article/3521586-where-blackberrys-44000-patents-come-from
[xviii] Oldham, P. 2022. The WIPO Patent Analytics Handbook. https://wipo-analytics.github.io/handbook/
[xix] https://www.uspto.gov/sites/default/files/documents/Basics-of-Prior-Art-Searching.pdf
[xx] https://media-publications.bcg.com/innovation-systems-need-a-reboot-layout.pdf
[xxi] https://www.copperpodip.com/post/10-prior-art-search-mistakes-that-can-undermine-your-inter-partes-review

		Combining two or more references that relate to different technology areas would have little chance of success in invalidating a subject patent."
C5	Ahmed (2019)[xxii] for *Forbes*, a financial media outlet	• "…**patent search will likely take advantage of the patent classification search model**. All patent applications are classified by the USPTO and other patent search databases into patent classification schemes. These schemes are generally structured according to the technical content of any patent applications they contain…**This classification model can prove helpful during a patent search in a number of ways, especially the ease with which attorneys can cross-search patents classified as containing any specifically relevant feature**."
C6	Upadhyaya (2024)[xxiii] for Sagacious IP, an IP service firm	• "**Patents,** unlike other repositories of documents, **are classified based on their technology areas**. Such technological classifications make the process of searching relevant prior art documents extremely easy."
C7	Rana (2021)[xxiv] for S&P Global, a financial information provider	• "**Technology overlaps between patents can be identified by the International Patent Classification (IPC) codes which maps each patent to relevant technology classes**…Patent data reveals connectedness between companies not captured by sector classification schemes. This offers investors an alternate approach to group peer companies by shared technology attribute."
C8	Innovalpha (2019)[xxv], an investment research firm	• "**In terms of the various classes under which GE filed patent applications**, the top two classes have remained pretty much constant throughout the 5-year period… a change in strategy has occurred within the company **in relation to R&D fields of activity."**

[xxii] https://www.forbes.com/councils/theyec/2019/09/27/patent-classification-search-what-you-need-to-know/
[xxiii] https://sagaciousresearch.com/blog/why-is-prior-art-search-important-for-innovators/
[xxiv] https://www.spglobal.com/marketintelligence/en/documents/technology-momentum-peer-networks-from-patents.pdf
[xxv] https://seekingalpha.com/article/4308702-ge-is-buy-considering-patent-dynamics

Table B2. Explorative analysis of press coverage of contrasting patent samples

Panel A:	*N*	*Number of Articles*	*Press Release = 1*
Top 1% in Market Value	50	2.26(0.40)	0(0)
Bottom 1% in Market Value	50	1.80 (0.44)	0.04(0.03)
		p-value = 0.44	*p-value = 0.16*
Panel B:	***N***	***Number of Articles***	***Press Release = 1***
Top 5% in Local Recombination, No Distant Recombination, and Top 5% in Market Value	48	0.71(0.16)	0(0)
Top 5% in Distant Recombination, No Local Recombination, and Bottom 5% in Market Value	22	0.86(0.24)	0.05(0.05)
		p-value = 0.59	*p-value = 0.14*

Note: Standard errors in parenthesis; p-values are based on two-tail t-tests. For Panel A, the top 1% and bottom 1% patents in market value were randomly drawn from the full sample. For Panel B, all patents fitting the criteria are included for news archive data collection and analysis. We use count measures to define the top 5% thresholds for local and distant recombinations. Manual searches and reviews of news articles and press releases were conducted in Factiva. The count of articles comes from the search results for i. the keyword “patent” and the firm being an article subject in the three-day window of patent grant, and ii. the patent number. We also reviewed each article to determine if it was a press release issued by the firm or contained information not disclosed in the patent grant.