Online appendix for "Risky recombinations: Institutional gatekeeping in the innovation process"

John-Paul Ferguson
Stanford University
jpferg@stanford.edu

Gianluca Carnabuci
European School of Management and Technology
gianluca.carnabuci@esmt.org

A Appendix: Exclusion using art-unit workload

We fit two-stage models that first estimate the effects of recombinant breadth on a patent application's being granted and then the effects of breadth on citations that the patent receives. The difficulty with such models is that many features of an invention that predict its being granted patent (most notably, its quality) should also predict its being cited by future innovators. Including only such variables means relying on assumptions about the models' functional form to identify selection effects. This is an unacceptably strong assumption in the best of circumstances. We would prefer to include a variable or variables in the first stage whose effect on future citations happens only through the indirect channel of selection itself. That is, we need a variable that is correlated with application granting but is uncorrelated with forward citations.

In recent work with USPTO data, researchers have used fixed effects for patent examiners to provide exogenous variation. Patent examiners vary in stringency when evaluating patents, and the assignment of a particular application to a particular examiner is uncorrelated with observable measures of invention quality (Lemley & Sampat 2012). As of this writing, though, the EPO refuses to release detailed information on patent examiners.

Rather than exploit idiosyncratic variation in approval likelihood by examiner, we exploit such variation by time. We leverage the fact that examiners' workloads, and the resulting amounts of time that they have to devote to each application, vary. This variation has nothing to do with the focal patent application. We know that evaluator behavior often shifts under the pressures of increased workloads (Simcoe & Waguespack 2011, Boudreau, Guinan, Lakhani & Riedl 2016), producing greater or lesser stringency depending on the context. And there is a first-order negative relationship in our data between the workloads of EPO art units receive and their approval rates. Thus art-unit workload has the potential to meet the exclusion restriction in our models.

However, several non-idiosyncratic trends must be removed from examiner workload for it to work as a plausible instrument. First, the number of patent applications received by most EPO art units has increased over time. Second, the workload varies systematically across art units, which reflects the uneven nature of advancement along the knowledge frontier as inventors exploit new technological opportunities. It is reasonable to presume that the EPO assigns staff to art units in proportion to their expected workloads.

As described in the main text, we therefore develop a measure of idiosyncratic fluctuations in art-unit workloads by first constructing a panel of workloads by art unit-month and then purging it of systematic variation by time and art unit. Let w_{jt} be the workload in art unit j in month t. Here t is a specific month, such as January 2000. Workload is highly right skew, so we take its natural log; figure A1 shows a histogram of the logged workload. We want to regress $\ln(w_{jt})$ on a full suite of month, art-unit, and art-unit/year indicator variables. Because interacting art units

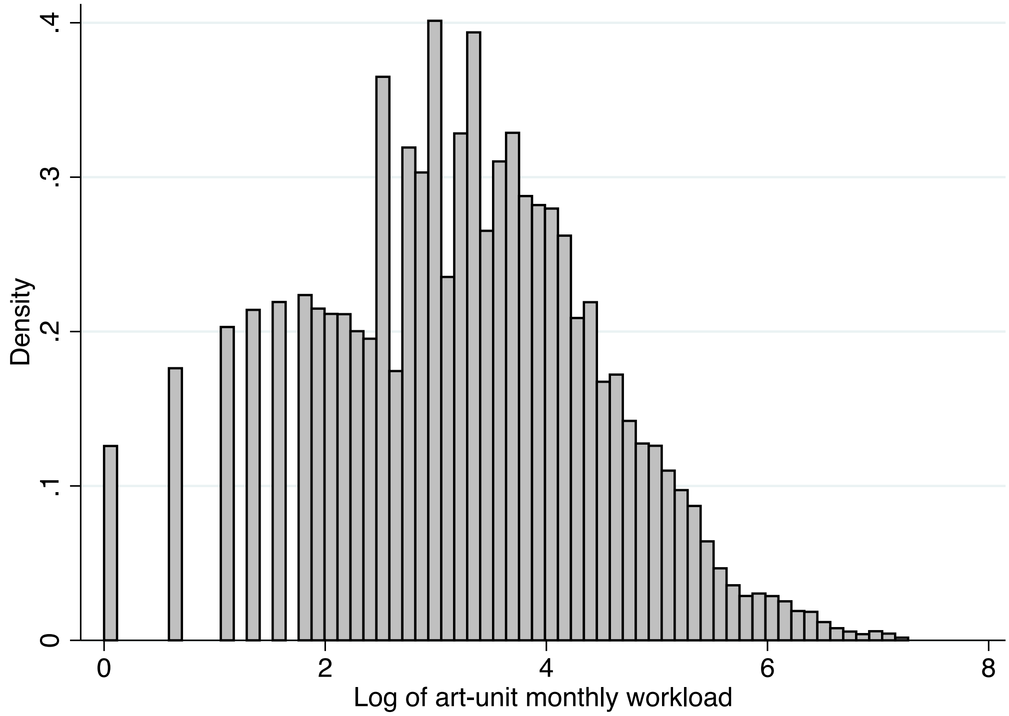

Figure A1: Distribution of the logged values of monthly art-unit workloads. See table 4 for details about the sample.

with years would produce roughly 20,000 variables and because we are interested in removing the systematic variation and not in recovering the fixed effects themselves, we demean the observations by the group mean for their corresponding art-unit/year. We regress the de-meaned variable on the full suite of 615 t month dummies and 631 j art-unit dummies. We use robust standard errors in this regression to account for how the absolute variance in workload grows with the mean. This regression removes any time-series variation as well as any systematic differences in workload across art units that may change over time. We calculate $\ln(\hat{w}_{jt})$ and take the residual. We interpret this residual, $\ln(\dot{w}_{jt})$, as the *idiosyncratic* art-unit workload in a given month, above or below what was suspected given the secular trend of applications arriving in that unit.

Figure A2 plots the residuals against the observed values for the monthly workload. The correlation between the two is less than .001. We are confident that the regression purges any systematic (and thus potentially endogenous) variation from this workload measure.

We next checked the correlation of this workload measure with the other model variables. Table A1 presents a balance test. To produce the coefficients in this table, we regressed each of the listed variables on art-unit workload, with the same sample of observations used in tables 4 and 5. Workload is significantly correlated with the probability of an application's being granted but not significantly correlated with other variables in the model. We also checked its correlation with variables in the citation stage. The closest to significance we see is with technical fertility, where $p < .1$. The idiosyncratic variation in art-unit workload among granted patents is not directly correlated with those patents' forward citation counts.

Finally, we ran a placebo test. Art-unit workload in the month an application is filed is correlated with the likelihood of that application's being granted patent, but one might reasonably argue

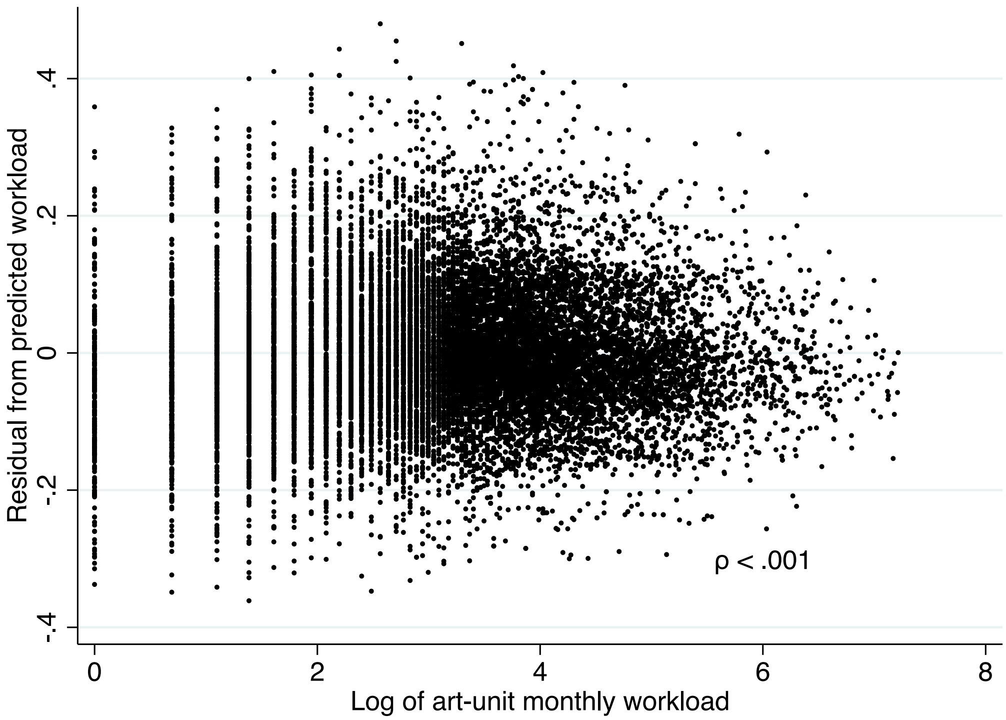

Figure A2: Scatterplot of observed (logged) monthly art-unit workloads against regression residuals. A random 1% sample of points are plotted here for clarity. We regressed $\ln(w_{jt}) = \alpha_0 + \beta' t + \gamma' j + \delta' jy$ where t are unique month indicators (i.e., January 2000), j are four-digit IPC art units, and y are unique years. robust standard errors were calculated to account for the absolute variance of workloads growing with the mean. See table 4 for details about the sample. Gaps in the scatter plot result from logging the monthly workload counts; compare with figure A1.

Table A1: Balance test of Art-unit workload against other variables used in models

Patent-granting stage	
Application granted	-0.031**
	(0.008)
Patent breadth	-0.031
	(0.029)
ln(Backward citation count)	0.023
	(0.023)
ln(Inventor count)	-0.034
	(0.021)
ln(Technical fertility)	-0.017
	(0.018)
Patent-citation stage	
Forward citation count	1.691
	0.921
Patent breadth	-0.025
	(0.018)
ln(Backward citation count)	-0.007
	(0.013)
ln(Avg. citation age)	-0.005
	(0.004)
ln(Inventor count)	-0.039
	(0.025)
ln(Technical fertility)	-0.141^{+}
	(0.082)

See tables 4 and 5 for details about the samples used in the two stages. Each coefficient represents the estimated effect of our art-unit workload measure on the listed variable in a bivariate regression with standard errors clustered by art units defined at the four-digit IPC level.
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

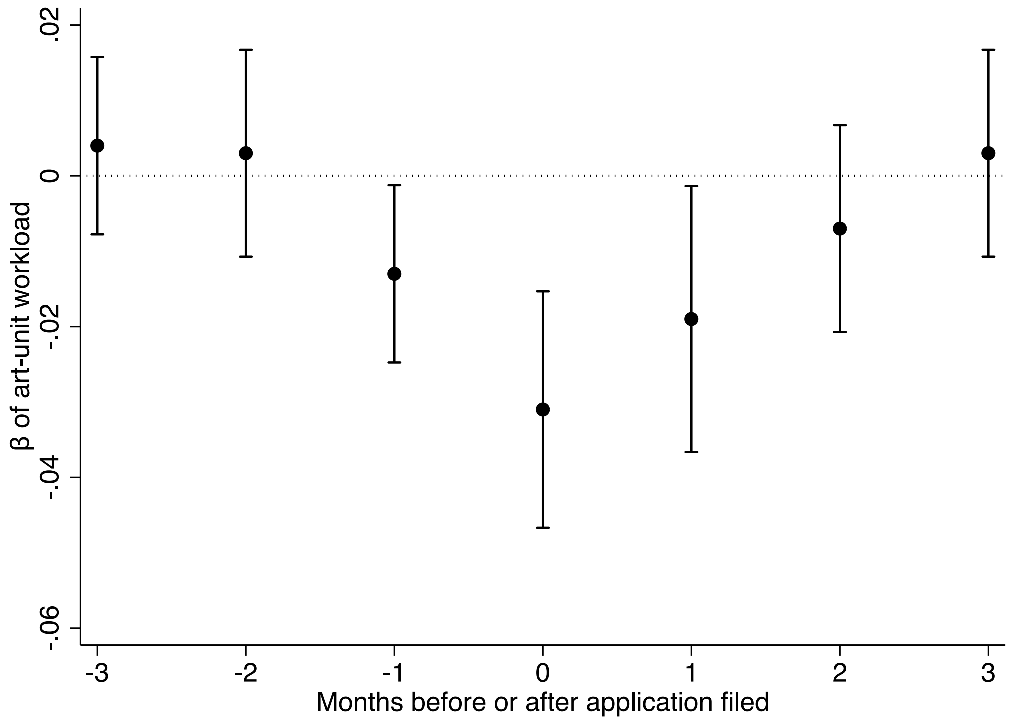

Figure A3: Robustness check of idiosyncratic art-unit workload's impact on patent-granting rates. See table 4 for details about the sample. Each point on the x-axis represents a bivariate regress of the probability of granting on workload, in which each observation has been assigned the value for the art unit's workload x months before or after the application was actually filed.

that this measure is picking up something else if the workload in *other* months is also correlated with granting rates. We therefore estimated correlations between granting and $\ln(\dot{w}_{j,t-3})$ through $\ln(\dot{w}_{j,t+3})$. We present these in figure A3. There are significant correlations between workloads in the months immediately before and after an application is filed and granting rates. This surely reflects how work on specific applications can spill over from month to month, and how the art units can adjust to increased workloads with some lag. But the coefficients in $t-1$ and $t+1$ are much smaller than in t, and there is no significant correlation for the more temporally distant variables.

As with any instrumental variable, we cannot rule out the possibility that idiosyncratic changes in art-unit workloads are correlated with some unobserved variable. But its non-correlation with the observed variables puts some bounds on that concern. This, combined with its non-correlation with predicted workloads and the placebo test, gives us some confidence in using this measure to satisfy the exclusion restriction in the first stage of our models.

B Appendix: Citations to failed patent applications

Within the European Patent Office, the examiners' mandate is to establish the novelty of a patent application, which requires identifying and citing all relevant public knowledge that constitutes prior art. Specifically, Article 54 of the European Patent Convention states the following:

1. An invention shall be considered to be new if it does not form part of the state of the art.

2. The state of the art shall be held to comprise everything made available to the public by means of a written or oral description, by use, or in any other way, before the date of filing of the European patent application.

3. *Additionally, the content of European patent applications as filed*, the dates of filing of which are prior to the date referred to in paragraph 2 and which were published on or after that date, shall be considered as comprised in the state of the art. [Emphasis added]

Because the EPO publishes patent applications before a granting decision is made (I.e., while the application is pending), the details of applications enter the stock of public knowledge before the application is granted. This can produce situations where an application is denied (perhaps the application is judged to have insufficient industrial applicability) yet where part of the invention is novel or otherwise relevant to subsequent work. When EPO examiners compile applications' search reports, they are required to acknowledge such relevant prior work, even if the application in which the work appeared was ultimately denied patent.

The European Patent Convention offers a legal rationale for this routine: If inventor A applies for patent with an invention but is ultimately denied, and then inventor B learns of advances in the state of knowledge from the published record of A's application and produces a new invention that receives patent, then inventor A should still be able to develop inventions that build upon their own original ideas. If no citations to A's failed application were made in inventor B's patent, the evidence of A's inventive step would be separated from the patent and only visible in the application file. Hence the inclusion of references to A's failed application in B's patent's prior-art references.

More information can be found in reference to the European Patent Convention itself: `http://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ar54.html`.

C Appendix: Comparing recombinant breadth to similar measures

In their presentation of the measure of technological breadth, Gruber, Harhoff & Hoisl (2013) argued and showed that it performs similarly to a simple Herfindahl index of the classes of a patent's backward citations. We constructed such a Herfindahl index and, as with the recombinant-breadth models, estimated its effects on the likelihood of application approval, built an inverse Mills ratio, and then estimated its effects on impact while controlling for selection. Such a Herfindahl is what Hall, Jaffe & Trajtenberg (2001) referred to as the "originality" of a patent, and indeed our interpretation would be the same: original inventions are somewhat less likely to be approved but, once approved, have a greater impact. Tables C1 and C2 show similar patterns to what we find with our recombinant breadth measure. Finding comparable effects of institutional gatekeeping for this measure gives us some confidence that our main results are not simply an artifact of which measure of breadth we chose to use.

Table C1: Probit models predicting granting of patent application, using Herfindahl rather than recombinant breadth

	(1)	(2)	(3)
IPC Herfindhal	-0.151***	-0.171***	-0.171***
	(0.028)	(0.027)	(0.027)
ln(Backward citation count)		0.033***	0.033***
		(0.008)	(0.008)
ln(Inventor count)		0.106***	0.106***
		(0.006)	(0.006)
ln(Tech. fertility)		-0.071***	-0.071***
		(0.011)	(0.011)
Art-unit workload			-0.028***
			(0.008)
Filing-year F.E.s	Y	Y	Y
Constant	0.770***	2.239***	2.233***
	(0.067)	(0.257)	(0.257)
Observations	844060	844060	844060
Log-likelihood	-550434	-548297	-548286

See table 4 for details about the sample. Herfindahl index is calculated at the four-digit IPC level. Compare coefficients with table 4 in the main text. Standard errors, clustered by patent art units defined at the four-digit IPC level, are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C2: QML Poisson models predicting forward-citation impact of granted patents, using Herfindahl rather than recombinant breadth

	(1)	(2)	(3)
IPC Herfindhal	0.308***	0.124***	0.083***
	(0.009)	(0.009)	(0.010)
ln(Backward citation count)		0.143***	0.130***
		(0.005)	(0.005)
ln(Avg. citation age)		1.119***	1.122***
		(0.025)	(0.025)
ln(Inventor count)		0.158***	0.172***
		(0.005)	(0.005)
ln(Tech. fertility)		0.142***	0.032
		(0.003)	(0.018)
Inverse Mills ratio			-0.420***
			(0.065)
Filing-year F.E.s	Y	Y	Y
Granting-year F.E.s	Y	Y	Y
Constant	2.554***	-10.743***	-7.124***
	(0.005)	(0.222)	(0.601)
Observations	519475	519475	519475
Log-likelihood	-6872519	-6584935	-6584196

See table 5 for details on the sample. Herfindahl index is calculated at the four-digit IPC level. Compare coefficients with table 5 in the main text.
Standard errors, clustered by patent art units defined at the four-digit IPC level, are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

D Appendix: Investigating application review times

If broader patent applications spend longer under review then broader applications will be over-represented in the pending subset of our data. If many of these broad applications would ultimately be approved, then our estimate of the approval penalty on breadth in table 4 would be overstated.

We check against the possibility in two ways. First, we fit hazard models of approval on the data. These model results are shown in table D1. While broader patent applications *are* over-represented among pending records, the coefficients on recombinant breadth in Cox proportional-hazard models suggest that, if anything, these pending records are *less* likely to be approved at any given time than narrower pending applications are. Accounting for the greater likelihood that a broader application will be pending at any time of observation does not eliminate the negative coefficient on breadth. The approval penalty we find in our main analyses is not an artifact of ignoring pending applications.

Table D1: Cox proportional-hazard models predicting granting of patent application

	(1)	(2)	(3)
Patent breadth	-0.269***	-0.161***	-0.161***
	(0.003)	(0.004)	(0.004)
ln(Backward citation count)		-0.101***	-0.101***
		(0.002)	(0.002)
ln(Inventor count)		-0.117***	-0.116***
		(0.002)	(0.002)
ln(Tech. fertility)		-0.157***	-0.156***
		(0.001)	(0.001)
Art-unit workload			-0.119***
			(0.006)
Filing-year F.E.s	Y	Y	Y
Observations	1240377	1240377	1240377
Log-likelihood	-6796611	-6786600	-6786390

See table 4 for details on the sample.
Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Second, we estimate OLS models of the amount of time applications spend under review. Harhoff & Wagner (2009) for example found that applications that cite more IPC classes spend longer in the examination process. We tested to see whether the same relationship holds when considering recombinant breadth. Table D2 presents OLS regressions of review time on breadth. Models 1 and 2 show that more breadth is associated with longer review times. We also examined the impact of radical innovations (as described in the main text, these are combinations of technology classes that have never been combined by previous patents) on review times. Model 1 shows that radical novelty is associated with shorter time under review, as we might predict. Model 2 tests an interaction between recombinant breadth and radical novelty; it shows that radical novelty can offset much of the increases in review time associated with breadth. This is consistent with our argument that the greatest hurdle in parsing boundary-spanning applications lie in determining the scope of the contribution of an invention that spans many domains. Approving an invention is quicker and easier when any examiner can see that a new bridge is being formed, regardless of the organizational and cognitive demands required to parse the rest of the application.

Table D2: OLS models of review time for patent-granting process

	(1)	(2)
Patent breadth	229.208***	231.485***
	(38.166)	(38.698)
Radical novelty	-79.946*	49.814
	(36.912)	(78.377)
Breadth × Novelty		-174.741*
		(85.907)
ln(Backward citation count)	48.650**	48.631**
	(15.876)	(15.876)
ln(Inventor count)	68.676***	68.636***
	(12.521)	(12.523)
Filing-year F.E.s	Y	Y
Constant	1292.461***	1293.289***
	(112.462)	(112.379)
Observations	546593	546593
R^2	0.13	0.13

Models are estimated on all closed cases of patents with non-missing data filed with the EPO between 1983 and 2007.
Standard errors, clustered by patent art units defined at the 4-digit IPC level, are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We also think that review time is worth considering because it gives indirect evidence of the mechanisms that we have proposed are at work behind approval hurdles. If patent examiners interpret atypical combinations as a mark of novelty and inventiveness, they should approve an application faster if that application comprises unambiguously new combinations, like those we have flagged with our novelty variable. Furthermore, such unambiguous marks of novelty should be especially relevant when examiners face broad, boundary-spanning applications. We say this because, as we have argued, the greatest difficulties in parsing broad applications should lie in determining the scope of the contribution of an invention that bridges many domains. In other words, although examiners are generally less likely to identify the inventive step of broader applications, when such applications comprise an unambiguously novel combination then they can recognize *that* as something novel and inventive.

E Appendix: Additional comparison of recombinant breadth and typicality

When we compare the effects of recombinant breadth and typicality in the same model, we must use a subset of the observations that we use in our main analysis. This is because, while breadth is defined (as zero) for applications that only include citations to a single four-digit IPC class, the Jaccard is only defined over applications that cite multiple classes. Thus we first reproduce (in table E1 and E2 our main analyses on the subset of observations where recombinant breadth is greater than zero.

Table E1: Probit models predicting granting of patent application (breadth > 0)

	(1)	(2)	(3)
Patent breadth	-0.452***	-0.441***	-0.441***
	(0.052)	(0.048)	(0.048)
ln(Backward citation count)		0.011	0.011
		(0.007)	(0.007)
ln(Inventor count)		0.101***	0.101***
		(0.006)	(0.006)
ln(Tech. fertility)		-0.084***	-0.084***
		(0.010)	(0.010)
Art-unit workload			-0.032***
			(0.007)
Filing-year F.E.s	Y	Y	Y
Constant	1.115***	2.946***	2.938***
	(0.085)	(0.254)	(0.254)
Observations	617997	617997	617997
Log-likelihood	-404481	-402831	-402819

Models are estimated on the subset of patent applications with non-missing data filed at the EPO between 1983 and 2007, where recombinant breadth of the application's prior work is greater than zero.
Standard errors, clustered by patent art units defined at the four-digit IPC level, are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

One point that should be noted is that, on this subset of observations, when controlling for selection the effect of recombinant breadth on citation impact is not statistically different from zero. However, this seems to be due to weaker precision rather than a real effect. The absolute difference in the size of the coefficients between models 2 and 3 of table E2 is smaller than in table 5, and the standard error on breadth in model 3 is considerably larger.

There is not no best practice for how to aggregate the dyadic Jaccard indices across all categories of an object into a single measure. In the main text we have used each application's minimum Jaccard. A similar case could be made for using the average Jaccard instead. We present results here with an average Jaccard as well. The substantive pattern of results is not changed.

Table E2: QML Poisson models predicting forward-citation impact of granted patents (breadth > 0)

	(1)	(2)	(3)
Patent breadth	0.189***	0.081**	0.062
	(0.027)	(0.025)	(0.056)
ln(Backward citation count)		0.158***	0.156***
		(0.006)	(0.008)
ln(Avg. citation age)		1.220***	1.220***
		(0.031)	(0.031)
ln(Inventor count)		0.141***	0.141***
		(0.006)	(0.006)
ln(Tech. fertility)		0.132***	0.129***
		(0.003)	(0.011)
Inverse Mills ratio			-0.044
			(0.130)
Constant	2.543***	-11.476***	-11.258***
	(0.023)	(0.277)	(0.689)
Observations	375122	375122	375122
Log-likelihood	-5274047	-5044364	-5044362

Models are estimated on the subset of patents with non-missing data filed with the EPO between 1983 and 2007 and granted by the EPO between 1984 and 2008, where the recombinant breadth of the patent's prior work is greater than zero.

Standard errors, clustered by patent art units defined at the four-digit IPC level, are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table E3: Probit models predicting granting of patent application, using average rather than minimum Jaccard

	(1)	(2)	(3)	(4)
Average Jaccard	-0.865***	-0.746***	-0.744***	-0.728***
	(0.098)	(0.099)	(0.099)	(0.096)
Patent breadth				-0.072***
				(0.009)
ln(Backward citation count)		0.020***	0.020***	0.027***
		(0.005)	(0.005)	(0.005)
ln(Inventor count)		0.095***	0.095***	0.097***
		(0.005)	(0.005)	(0.005)
ln(Tech. fertility)		-0.097***	-0.097***	-0.096***
		(0.007)	(0.007)	(0.007)
Art-unit workload			-0.024*	-0.024*
			(0.011)	(0.011)
Filing-year F.E.s	Y	Y	Y	Y
Constant	0.839***	3.106***	3.100***	3.104***
	(0.025)	(0.183)	(0.183)	(0.182)
Observations	430375	430375	430375	430375
Log-likelihood	-279316	-278224	-278220	-278155

See table 4 for details on the sample. Models here are estimated on the subset where the application lists two or more IPC classes. The Jaccard is not defined for single-class cases—hence the smaller number of observations compared to table 4.

Standard errors, clustered by patent art units defined at the four-digit IPC level, are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table E4: QML Poisson models predicting forward-citation impact of granted patents, using average rather than minimum Jaccard

	(1)	(2)	(3)	(4)	(5)
Average Jaccard	-0.462***	-1.280***	-1.334***	-1.290***	-1.029***
	(0.034)	(0.039)	(0.044)	(0.039)	(0.077)
Patent breadth				0.212***	0.040***
				(0.043)	(0.011)
ln(Backward citation count)		0.155***	0.149***	0.152***	0.170***
		(0.007)	(0.007)	(0.007)	(0.007)
ln(Avg. citation age)		1.165***	1.162***	1.161***	1.160***
		(0.036)	(0.036)	(0.036)	(0.036)
ln(Inventor count)		0.150***	0.150***	0.149***	0.149***
		(0.007)	(0.007)	(0.007)	(0.007)
ln(Tech. fertility)		0.174***	0.167***	0.174***	0.215***
		(0.005)	(0.005)	(0.005)	(0.011)
Inverse Mills ratio			-0.075**		0.426***
			(0.027)		(0.105)
Filing-year F.E.s	Y	Y	Y	Y	Y
Granting-year F.E.s	Y	Y	Y	Y	Y
Constant	1.647***	-13.802***	-13.368***	-13.781***	-16.143***
	(0.240)	(0.412)	(0.440)	(0.412)	(0.680)
Observations	266175	266175	266175	266175	266175
Log-likelihood	-4094042	-3914575	-3914325	-3914172	-3913830

See table 5 for details on the sample. Models here are estimated on the subset of patents where the patent lists two or more IPC classes. The Jaccard is not defined for single-class cases—hence the smaller number of observations compared to table 5.

Standard errors, clustered by patent art units defined at the four-digit IPC level, are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani & Christoph Riedl. 2016. "Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science." *Management Science* . First published online at http://dx.doi.org/10.1287/mnsc.2015.2285.

Gruber, Marc, Dietmar Harhoff & Karin Hoisl. 2013. "Knowledge Recombination across Technological Boundaries: Scientists vs. Engineers." *Management Science* 59(4):837–851.

Hall, Bronwyn H., Adam B. Jaffe & Manuel Trajtenberg. 2001. The NBER Patent Citation Data File: Lessons, Insights, and Methodological Tools. Technical Report 8498 National Bureau of Economic Research.

Harhoff, Dietmar & Stefan Wagner. 2009. "The Duration of Patent Examination at the European Patent Office." *Management Science* 55(12):1969–1984.

Lemley, Mark A. & Bhaven Sampat. 2012. "Examiner Characteristics and Patent Office Outcomes." *Review of Economic Studies* 94(3):817–827.

Simcoe, Timothy S. & Dave M. Waguespack. 2011. "Status, Quality, and Attention: What's in a (Missing) Name?" *Management Science* 57(2):274–290.