Patent Examiner Specialization*

Cesare Righi[1] and Timothy Simcoe[2]

[1]Boston University
[2]Boston University and NBER

August 4, 2018

Abstract

We study the matching of patent applications to examiners at the U.S. Patent and Trademark Office. The distribution of technology classes is more concentrated than would occur under random matching and F-tests reject the hypothesis that family size and claim scope are randomly distributed across examiners. Using the application text, we show that examiner specialization persists even after conditioning on technology sub-classes. Specialization is less pronounced in computers and software than other technology fields. More specialized examiners have a lower grant rate. These findings undermine the idea that random matching justifies instrumental variables based on examiner behaviors or characteristics.

*This is a revised version of the second chapter of Cesare Righi's doctoral dissertation. We thank Sam Arts, Jim Bessen, Bruno Cassiman, Christian Catalini, Iain Cockburn, Dennis Crouch, Stuart Graham, Deepak Hegde, Mark Lemley, Alan Marco, Michael Meurer, Arti Rai, Bhaven Sampat, Mark Schankerman, Andrew Toole, Joel Waldfogel, and Heidi Williams, along with seminar participants at the BU Technology & Policy Research Initiative, BU Strategy & Innovation department, Université de Bordeaux, Copenhagen Business School, KU Leuven, Max Planck Institute for Innovation and Competition, Munich Summer Institute, NBER productivity lunch and the USPTO for useful comments and suggestions. We thank Evgeny Klochikhin for access to the PatentsView patent applications database. The authors have received funding from Google, Intel and the Boston University Technology & Policy Research Initiative. Contact: crighi@bu.edu; tsimcoe@bu.edu.

1 Introduction

In 2015, the U.S. Patent and Trademark Office (USPTO) received 589,410 utility patent applications. Matching each application to a qualified examiner is a fundamental part of the examination process. This matching proceeds in two steps. First, each application is assigned to an "art unit" comprised of several examiners who specialize in a particular technology. Then the application is assigned to an individual examiner within that art unit. Motivated by the accounts in Cockburn, Kortum & Stern (2002) and Lemley & Sampat (2012), several studies have suggested that the second step in this process is more-or-less random, and then, building on an idea proposed by Sampat & Williams (2017), used examiner characteristics as instrumental variables for examination outcomes.[1]

We re-examine the random matching assumption, and find strong evidence of technological specialization by patent examiners within art units. Our statistical tests also reject the hypothesis that proxies for patent quality (family size), scope (the length of the first claim), and the identity of the applicant (assignee) are randomly distributed across examiners. Examiner specialization is more pronounced in the art units that examine Biotechnology, Chemistry, Mechanical Engineering and Semiconductor applications, and less so in the computer-related art units. Using a measure of textual similarity between pairs of patent applications, we show that examiner specialization persists even within U.S. Patent Classification System (USPC) subclasses. Finally, we find that more specialized examiners have a lower grant rate and produce a larger narrowing of claim-scope during the examination process.

These findings have implications for instrumental variable strategies based on examiner behaviors and characteristics. Random assignment would suffice to make leave-one-out grant rates (or any other examiner characteristic) uncorrelated with both observed and unobserved characteristics of an application. But when examiners specialize, examiner-based instruments may be correlated with unobserved variation in the underlying technology, sug-

[1] Papers adopting variants on this identification strategy include Farre-Mensa, Hegde & Ljungqvist (2017), Feng & Jaravel (2017), Gaulé (2015), Kuhn (2016), Kuhn & Thompson (2017), and Sampat & Williams (2017).

gesting an alternative path through which the instrument could influence outcomes. Thus, although examiner-based instruments may be valid in some applications, the identification strategy rests on a stronger assumption than is typically acknowledged: potential outcomes must be uncorrelated with (unobserved) technological heterogeneity. Although we do not generally take a Manichean view of causal inference, we are forced to conclude that in this case, random matching of examiners to applications does not provide a general-purpose tool for causal inference about the patent system.

This is the first paper to systematically test the random matching hypothesis across all of the technology areas examined by the USPTO. We use three different approaches to provide evidence of specialization. First, we examine a pair of test statistics from the literature on industry agglomeration (Mori, Nishikimi & Smith 2005) that ask whether application characteristics (e.g. technology subclass) are more or less dispersed across examiners than would occur under random assignment.[2] These statistics are computed at the art-unit-filing-year level, and we examine the entire distribution of p-values for various application characteristics, including technology subclass, assignee, and indicators of patent value (family size) and scope (first independent claim length). Second, we use simple F-tests to examine whether examiner fixed-effects are correlated with our measures of patent value and scope. Our third approach exploits a new measure proposed by Arts, Cassiman & Gomez (2018), and uses OLS regression to show that the probability of two applications being assigned to the same examiner increases with the textual similarity of their titles and abstracts.

At a substantive level, our findings illustrate how the USPTO manages a tension between efficiency and fairness (Merges 1999). One way to promote fairness is through uniform application of patentability criteria, but prior research suggests that is difficult. Some examiners are simply tougher than others (Sampat & Williams 2017, Kuhn & Thompson 2017), and experienced examiners are more lenient on average, partly because of time constraints

[2]These methods focus specifically on the null hypothesis of random assignment, unlike instrumental variable falsification tests that ask the slightly different question of whether examiner and application characteristics are correlated.

(Lemley & Sampat 2012, Frakes & Wasserman 2017). Random matching provides another path to fairness, but forgoes the potential efficiency benefits of specialization. Our analysis shows that the amount of specialization varies across art units, leading some applicants to get tougher examiners on average. Although the important applications (with large families) and broad applications (with short first independent claims) are not randomly distributed, technological specialization explains most of this correlation. There is no evidence that the broadest or most important applications are assigned to specific examiners.

We consider two possible mechanisms that would generate more observed specialization in Chemistry, Biotechnology, Mechanical Engineering and Semiconductors than in the computer-related art units. One possibility is that "generalist examiners" are able to evaluate computing inventions, while more specialized skills and knowledge are required in other areas. Another possibility is that the USPC technology classification system works better outside computers and software, so we simply fail to observe much of the specialization that takes place within computer-related art units. The data provide support for both of these explanations. Application-pairs sharing the same examiner have less textual overlap in Computers and Communications, suggesting that examiners in the computer-related tech centers are more generalist. At the same time, application-pairs in Computers and Communications that share the same primary USPC subclass are less textually similar than in other fields, suggesting differences in the quality of the classification system.

Finally, we find a positive correlation between specialization and a more stringent examination process, suggesting that it is easier for examiners who specialize to find relevant prior art. Under random matching, these estimates have a causal interpretation. Alternatively, they remain important for showing how non-random matching is related to examination outcomes.

2 Patent Examiner Assignment at the USPTO

When a patent application is filed, the Office of Patent Application Processing reviews the formality requirements of the application and assigns it a serial number. A contractor then defines its technology classification by extracting keywords form the text of the application and comparing them to a set of keywords associated with USPC class and subclass codes.[3] Each application has at least one mandatory classification, which is defined as a unique combination of class and subclass identifiers. The current version of the USPC has roughly 450 classes and more than 150,000 subclasses.

The USPTO has eight Technology Centers (TCs) responsible for examination of utility patent applications in broad technological areas. Each TC is comprised of several art units, or teams of patent examiners who specialize in a particular technology. Technological classifications are used to assign each new patent application to a specific art unit.[4] Within each art unit the initial assignment of a new application is handled by a Supervisory Patent Examiner (SPE). The SPE can refine the technological classification of a new application if it is incorrect, or request that an application be transferred to another art unit. But in most cases, the SPE will assign the application to an examiner within her art unit.

Previous research documents that SPEs have substantial discretion in examiner assignment. Some SPEs interviewed by Lemley & Sampat (2012) mention assigning applications to examiners essentially randomly within subclasses. Other SPEs give the oldest unassigned application to an examiner when she finishes the examination of another application. Although these practices suggest random matching, some SPEs may encourage technological specialization of examiners within their art unit. Cockburn et al. (2002) suggest that the degree of technological specialization varies across art units – in some art units an individual examiner is responsible for almost all applications in a specific technology class, and in others the examiners are less specialized. In addition, our conversation with SPEs and examiners confirmed that skills and knowledge of a technical area are often among the main

[3] For details, see `http://www.uspto.gov/sites/default/files/patents/resources/classification/overview.pdf`. Although it was replaced by the Cooperative Patent Classification (CPC) on January 1, 2013, the USPC is the relevant classification for the entire period of our study.

[4] For the current list of classes and subclasses examined by each art unit, see `http://www.uspto.gov/patents-application-process/patent-search/understanding-patent-classifications/patent-classification`.

criteria that drive application assignment, along with fairness to applicants, examiners' workload and learning objectives.

Although the USPTO constantly monitors the performance of art units and examiners to ensure a certain level of quality of the examination process, the assignment to a particular art unit and to a specific examiner can have important consequences for an application. Different practices across art units and the personal approach of each examiner can affect whether an application is eventually granted (Sampat & Williams 2017), how quickly a decision is reached (Farre-Mensa et al. 2017), and the scope and strength of an issued patent (Kuhn & Thompson 2017). This variation in standards led Cockburn et al. (2002) to conclude that "there may be as many patent offices as patent examiners."

3 Methods and Data

We use two approaches to measure specialization. The first approach compares the actual distribution of application characteristics to the predicted distribution under random assignment. For continuous characteristics, we use a simple F-test from an examiner fixed-effects model. For discrete characteristics such as technology classes and assignee, however, we adopt a pair of tests from the literature on industry agglomeration. The second approach uses regression at the application-pair level to ask whether textual similarity predicts having the same examiner.

3.1 Agglomeration Test Statistics

Our main approach to measuring technological specialization uses two statistical tests originally developed to study industry agglomeration: the Divergence Index (D-index) and the Multinomial Test of Agglomeration and Dispersion (MTAD). In our setting, patent examiners are analogous to cities, and technology subclasses (or other application characteristics) are analogous to industries. We briefly describe the two tests here, and provide details in Appendix A.

The D-index was proposed by Mori et al. (2005), building on Kullback & Leibler (1951), and is based on the concept of relative entropy.[5] For a given technology subclass (or other discrete characteristic) indexed by i, the D-index is given by

$$D(\hat{p}_i|p_0) = \sum_{r\in\mathbf{R}} \hat{p}_{ir} \ln\left(\frac{\hat{p}_{ir}}{p_{0r}}\right). \tag{1}$$

where r indexes individual examiners in set $\mathbf{R}$; p_{0r} is the share of all applications assigned to examiner r; and $\hat{p}_{ir}$ is the share of all applications in subclass i assigned to examiner r. If there are N_i applications in subclass i, then $2N_iD(\hat{p}_i|p_0)$ has a chi-square distribution with $R-1$ degrees of freedom under the null of random matching. In our analysis, the number of D-index tests will equal the number of categories (e.g. one per technology subclass) and we examine the distribution of p-values from all of these tests conditional on a given sample-size threshold (e.g. $N_i > 20$).

MTAD was developed by Rysman & Greenstein (2005), and computes multinomial likelihood functions for an allocation of agents to a set of discrete locations. In our setting, the agents are patent applications and locations correspond to examiners. In particular, suppose we have R examiners, each receiving n_r applications, and C subclasses, each having unconditional probability p_c. The observed number of applications of type c assigned to examiner r is x_r^c. Under random matching of applications to examiners, the likelihood of observing allocation $\mathbf{x}_r$ for examiner r is the multinomial pdf

$$\mathcal{L}(\mathbf{x}_r, n_r, \mathbf{p}) = \binom{n_r}{x_r^1, ..., x_r^C} p_1^{x_r^1} ... \ p_C^{x_r^C}. \tag{2}$$

The intuition behind MTAD is to replace p_c with the observed share of subclass c, and then compare the sample average log-likelihood to the expected log-likelihood under simulated random assignment. If the likelihood of the observed data is lower (higher) than the likelihood under random choice, MTAD indicates that the technology classes are agglomerated (dispersed).[6]

[5]Statisticians often refer to the D-index as a G test statistic. The main advantage of a G-test relative to a chi-squared test of independence occurs when some cells in a frequency table have very small expected counts, which is the case in our setting.

[6]The difference between the observed and simulated likelihood is distributed asymp-

This approach differs from the D-index because the statistic is computed for an entire art unit, and because it can detect whether deviations from random assignment are due to agglomeration or over-dispersion.

3.2 Data

Our main data source is the USPTO Patent Examination Research Dataset (Graham, Marco & Miller 2015), which is based on information from the Public Patent Application Information Retrieval system (Public PAIR). We also use information from PatentsView (`http://www.patentsview.org`), PATSTAT, the USPTO Patent Assignment Dataset (Marco, Myers, Graham, D'Agostino & Kucab 2015) and the Patent Claims Research Dataset (Marco, Sarnoff & deGrazia 2016).

We restrict our analysis to published utility patent applications filed on or after the enactment of the American Inventor's Protection Act of 1999 (November 29, 2000) and before January 1st 2013, whose examiner is affiliated with one of the eight TCs responsible for the examination of utility patent applications. The USPTO Patent Examination Research Dataset provides information on the examiner of record for each application as of January 24, 2015. This is the examiner as of that date for pending applications and the examiner at the time of disposal for disposed applications. The USPTO Patent Examination Research Dataset also provides the art unit of the examiner of record at the time of the last office action recorded for a given application. Under the AIPA, regular utility patent applications are generally published eighteen months after filing.[7]

The data have several limitations. First, applications will not appear in our data if they are abandoned or granted before publication, or if the applicant files only in the United States and requests that the application not

totically normal and we use simulation to generate its confidence intervals. See Rysman & Greenstein (2005) for details on the test. Timothy Simcoe developed a software module to easily perform this test in Stata, available at the following link: `https://ideas.repec.org/c/boc/bocode/s457205.html`

[7] As in Graham et al. (2015) and in the Public PAIR data, we use the term "regular utility patent application" to distinguish nonprovisional utility patent applications from provisional, PCT, reissue or re-examination applications.

be published. Previous research suggests that these outcomes are relatively rare.[8] A second limitation is that we do not observe whether applications are transferred from one examiner to another.

Our primary analysis sample contains 2,717,032 applications examined by 12,338 unique examiners affiliated with 590 art units. Table 1 shows the number of art units, examiners, classes, subclasses and applications in each Technology Center (TC). We exclude applications filed after 2012 to avoid problems related to publication lags and a change in the USPTO technological classification scheme. We also exclude serialized continuations (continuation applications, continuations in part and divisional applications) because these applications are usually assigned to the same examiner of the original application, and would therefore lead us to overstate the extent of agglomeration.

Table 1 also reports sample sizes for the application-pair-level analyses. The full sample contains roughly 1.1 billion application-pairs filed in the same calendar year and processed by the same art unit. Our regressions utilize a sub-sample of 11.7 million application-pairs that have the same filing year, art unit, primary USPC subclass and have an assignee in our data.

3.3 Variables

We focus on several application characteristics that may influence the assignment of applications to individual patent examiners within an art-unit-filing-year.[9] The first of these characteristics is the primary USPC classification of the application, which is defined by a unique combination of primary class and primary subclass codes (for brevity, subclass). If patent examiners specialize in evaluating applications related to particular technologies, we expect to see agglomeration on this variable. While Lemley & Sampat (2012) provide qualitative evidence that this may happen at least in a subset of

[8] Graham et al. (2015) show that about 95% of the regular non-provisional utility patent applications filed between 2001 and 2012 can be found in Public PAIR. Moreover, only 7% of the applications that meet all other criteria for inclusion in our main analysis sample are not published.

[9] We typically compute our test statistics within a art-unit-filing-year cell to account for possible changes in assignment practices over time and turnover in the pool of examiners.

art units, our analysis tests this statistically on the entire population of art units (conditional on sample size) and identifies in what TCs technological specialization is more relevant.

We use technology classification data from published applications rather than granted patents to avoid measuring any agglomeration created by the examination process. In particular, because the USPC classification of an application is based on its claims, which are usually amended during examination, the subclass of many applications changes over time. This could lead to spurious agglomeration if certain examiners are more likely to reject claims in particular classes.[10]

The last two columns in Table 1 show that for patents granted before July 21, 2015, twenty percent of all applications change primary class during the examination process, and almost seventy percent change primary subclass. There is heterogeneity across TCs, with patents in Biotechnology and Chemicals changing classification more often than those in other areas. These changes are not problematic for our tests because, as noted above, we use the initial classification.[11]

The identity of the applicant is a second variable that could influence the allocation of applications — either directly or because of technological specialization. Although the identity of the organization filing the application is not visible to the SPEs at assignment, they can observe the inventors. SPEs may group applications filed around the same time by the same inventors and assign them to the same examiner to increase efficiency. Moreover, applicants usually specialize in certain technologies, and may assigned to examiners who specialize in the same area.

We measure the identity of the applicant with the assignee of an application. Specifically, we retrieve information on the assignment of applications, identify the assignments made by the inventors to their employers before the

[10]The data in Public PAIR provide only the most recent classification of an application, so we utilize the primary classification of applications at publication from PatentsView, which is more likely to reflect the classification contractor's original assignment.

[11]Many papers utilize USPC (sub)classes as a control variable, and future research might consider whether it is better to measure this variation at the time of application publication or grant.

application is docketed to an examiner and keep only applications assigned to a single entity, clean and standardize the assignee names and create clusters of names that are likely to belong to the same organization, to which we assign a unique identifier.[12] After completing this process, we have missing assignee data for 584,313 applications (about 20% of our main analysis sample).[13]

To check the robustness of our assignee measurement, we utilize a second measure of the applicant identity: the customer number assigned by USPTO to each application. This number identifies the correspondent for application-related matters and is usually either the law firm representing the applicant or the legal department of the firm filing the application.[14]

The size of a patent family is often used as a proxy for the economic value of an invention because increased value leads patentees to file in more countries (Harhoff, Scherer & Vopel 2003, Putnam 1996). We utilize the number of applications in the same DOCDB patent family, with filing dates on or before the focal application date, as a proxy for economic value. We treat family-size as a continuous variable, and also construct an indicator variable that equals one if a focal application is above the 95$^{\text{th}}$ percentile in the family size distribution (within an art-unit-filing-year) to test whether certain examiners are assigned a large share of "outlier" applications.[15]

Kuhn & Thompson (2017) show that the length of the first independent

---

[12]We employ an assignee name cleaning and standardization routine that builds upon Thoma, Torrisi, Gambardella, Guellec, Hall & Harhoff (2010) and the name standardization routines developed for the NBER Patent Data Project available at `https://sites.google.com/site/patentdataproject/Home/posts/namestandardizationroutinesuploaded`. Details are available upon request.

[13]Although this is a relatively large percentage of applications with missing values, we note that around 8% of granted patents are not assigned, and would therefore be unassigned at application. Utilizing the normalized differences in average covariates (Imbens 2015), we find that assigned and unassigned applications do not differ substantially in terms of filing year, family size at filing (both DOCDB and INPADOC), length of first independent claim, number of claims and of independent claims. The normalized differences range between -0.03 and 0.06, much below the usual thresholds that reveal serious imbalances. These results are available on request.

[14]Results of the customer number analysis are similar to those for the assignee and are available upon request.

[15]We test the robustness of these results using the INPADOC patent families. The results are similar to those for DOCDB patent families and are available upon request.

claim in a patent is a good measure of patent scope. The idea behind this measure is that shorter claims provide broader scope of patent protection, because every word added to the text of the claims can potentially introduce additional elements or characteristics that must be present to establish infringement. Again, we consider both a simple count of words in the first independent claim, and a dummy variable that equals one if and only if a patent application falls below the 5$^{\text{th}}$ percentile of the word count distribution.[16]

Finally, in our application-pair-level analyses, the outcome variable is an indicator variable equal to one for pairs that share the same examiner. The main explanatory variables are a Jaccard similarity index based on keywords in the abstract and title of each application, and an indicator variable equal to one for pairs filed by the same assignee.

To compute the Jaccard similarity, we pre-process the abstracts and titles of all published patent applications from the PatentsView patent application database following the procedure described in Arts et al. (2018). Then, for every pair where both applications have at least 10 keywords and share the same art unit and filing year, we compute the Jaccard similarity index as 100 times the number of unique keywords in the intersection of the two sets of keywords, divided by the number of unique keywords in the union of the two sets. Intuitively, this variable measures the percentage overlap of unique keywords. Arts et al. (2018) validate the Jaccard similarity index for the population of granted patents using a panel of experts and information on patent families, inventors, assignees and patent citations. In particular, they show that matching on the Jaccard index outperforms matching on the primary USPC subclass, based on the mean scores from an expert evaluation.[17]

[16]Because Kuhn & Thompson (2017) note that their measure of scope is not suitable for the analysis of patent scope in biotechnology, we exclude the Biotechnology TC from the analysis of this variable. We have also conducted tests based on a pair of alternative proxies fro the scope of the application: the total number of claims in the application and the number of independent claims. The results are similar to those reported in Tables 2 and 3, and are available from the authors on request.

[17]Summary statistics for all variables are reported in Table B1.

4 Results

This section presents evidence of patent examiner specialization, and regression results linking specialization to examination outcomes.

4.1 Examiner Specialization

Figure 1 shows that patent examiners handle more applications from a given subclass or assignee than we would expect under random allocation. Specifically, each panel shows a histogram of p-values from a sample of hypothesis tests. For the D-index (top row), we run a separate test for each art-unit-filing-year by subclass or assignee cell containing more than 20 applications. For MTAD (bottom row) we run a separate test for each art-unit-filing-year cell containing more than 50 applications.[18]

Under the null of random assignment, the p-values in Figure 1 should be uniformly distributed between zero and one. However, in each panel a large share of the test-statistics fall below the 1 percent statistical significance threshold, providing strong evidence of specialization. The two histograms in the left column indicate that about 25 percent of the D-index and MTAD tests for random USPC assignment have a p-value below 0.01. The two histograms in the right column show somewhat weaker evidence of specialization by assignee, with about 10 to 20 percent of the p-values falling below the 1-percent threshold. Overall, Figure 1 shows that the allocation of applications within art units is often far from random, and that SPEs take into account the technological classification when assigning applications to an examiner, as described in Lemley & Sampat (2012).

Table 2 examines the degree of examiner specialization in different TCs, and for an additional pair of application characteristics. Specifically, the table reports the share of D-index or MTAD tests that reject the null hypothesis of random allocation at a 1-percent significance level.[19] Panel A

[18] All of our results are robust to varying the within-cell sample size cutoffs, but going much below these thresholds leads to large numbers of uninformative tests. Figure B1 shows the distributions of p-values of D-index and MTAD for subclass and assignee with thresholds equal to, respectively, 10 and 25.

[19] Table B3 in the appendix reports analogous figures with a cutoff at the 5-percent

shows that there is evidence of examiner specialization in every TC, and all of the MTAD results indicate agglomeration rather than over-dispersion. The share of tests that reject random assignment is much lower, however, in the "Computer Architecture" and "Computer Networking" areas than for Biotechnology, Chemistry, Semiconductors and Mechanical Engineering. The results in Panel B, using assignees rather than subclasses, are very similar.

The lower half of Table 2 examines agglomeration for our dichotomous variables that measure whether an application is in the top ventile of the Family Size or Claim Scope distribution. These variables focus on extreme outcomes because we are interested in whether SPEs assign unusual applications to a specific set of examiners. The data suggest that, for the most part, they do not. There is some evidence that applications from very large families are concentrated among a smaller set of examiners for Chemicals, Communications, Semiconductors, Mechanical Engineering and the TC we labeled as "Miscellaneous." There is also evidence of specialization in broader patents (as measured by length of the first claim) in the Chemical and Materials Engineering and Semiconductors TCs. In general, the MTAD tests detect more evidence of specialization than the D-index. But overall, these effects as relatively small, and might easily be caused by the technological specialization observed in Panel A.

While the results in Panels C and D suggest that there is little sorting based on extreme values, this does not imply that the assignment of applications is completely random with respect to value or scope. To test for random assignment with respect to these two characteristics, we run a set of OLS models in which we regress either the count of DOCDB family members at filing date (Panel A) or the count of words in the 1st independent claim (Panel B) against a set of examiner fixed-effects. This analysis is based on the idea that if the assignment of applications is really random, the examiner fixed-effects should not have explanatory power.

We estimate one regression for each art-unit-filing-year subsample and report the results by TC in Table 3. The first column reports the share of F-tests that reject the null hypothesis that the examiner effects are jointly

threshold for statistical significance.

zero at 1% significance level, and the second column reports the number of regressions for each TC. The share of F-tests rejecting the null hypothesis is much larger than the share we would expect under random assignment in each TC. Interestingly, the examiner fixed-effects seem to be less related to value and scope in the Computers & Communications TCs. This is consistent with the idea that the assignment is more random in those TCs.

Because these results may reflect technology specialization, we re-run the models including subclass effects. The results are reported in the third and fourth column of Table 3. The share of F-tests rejecting the null hypothesis drops significantly in all TCs, suggesting that differences across technologies explain a substantial share of the differences across examiners detected by the previous tests. In relative terms, the drop in the share of rejections is less marked in the Computers & Communications TCs.

Overall, the results in this sub-section show that patent examiners specialize in particular technologies, even within fine-grained USPC technology subclasses. There is more specialization in Chemistry, Biotech, Mechanical Engineering and Semiconductors, and less in computer-related technology. While examiner fixed effects are correlated with family size and scope, controlling for differences in technology attenuates these relationships and we find very little evidence that certain examiners specialize in "outlier" patent applications.

4.2 Text-based Specialization Tests

The findings in Tables 2 and 3 raise two questions: (1) What explains heterogeneity in specialization across TCs?, and (2) Is examiner-application matching effectively random within narrower technology subclasses?

With respect to the the first question, there are at least two reasons why we might find less specialization in the computer-related art units. First, examiners in the less agglomerated TCs may be "generalists" who are capable of evaluating most applications within their art unit. This would naturally lead SPEs to adopt a more random allocation process. Alternatively, patent examiners in the Computers and Communications TCs might be just as specialized as their counterparts in other TCs, but this is not apparent to us because the USPC classification system is less representative of actual technological differences in these fields.

If we want to evaluate the USPC classification system, an alternative measure of technological distance is required. Prior research has examined overlap of primary and secondary classification codes to measure technical similarity (Breschi, Lissoni & Malerba 2003). We use the Jaccard similarity index, because it does not depend on technological classifications.

The first column of Table 4 reports the mean Jaccard similarity for application-pairs within art-unit-filing-years by TC, and column (2) reports the mean Jaccard similarity conditional on sharing the same primary subclass. Not surprisingly, pairs sharing the same primary subclass are more similar across all TCs.[20] To compare the magnitude of these differences, column (3) reports the percentage change in mean Jaccard similarity when conditioning on sharing the same primary subclass. The percentage increase in keyword similarity is large for all TCs, indicating that the USPC system does capture a substantial amount of the variation reflected in the Jaccard measure. However, the percentage increase is much smaller in the Computer and Communication TCs than in other areas, suggesting that the USPC system works less well for those technologies.

To compare "examiner generality" across TCs, we repeat the same exercise, conditioning on application pairs with a common examiner instead of a shared primary subclass. Columns (4) and (5) of Table 4 show the results. It is not surprising that having the same examiner is less related to technological similarity of applications than having the same subclass because there are many more subclasses than examiners. Nevertheless, in each TC, application-pairs assigned to the same examiner are more similar than those with different examiners. The percentage changes are much smaller in the

[20] Given the relatively small percentage of application-pairs sharing the same subclass, the unconditional mean similarity is almost identical to the mean similarity conditional on having two different primary subclasses, and we omit the latter statistic for brevity. The same logic applies to the analysis in which we consider the examiner instead of the subclass. The t-tests for the differences in mean Jaccard similarity between application-pairs with the same or a different subclass or examiner are all statistically significant at 1%.

Computers and Communications TCs, suggesting that examiners in that area are assigned more heterogeneous applications. The increase in technological similarity is especially high in Chemistry, where having the same examiner is associated with a 55% increase in the Jaccard similarity.

The results in Table 4 provide evidence supporting *both* of our potential explanations for differences in examiner specialization across TCs. In particular, we find that applications sharing a primary USPC subclass exhibit more similarity in Biotechnology and Chemistry and less in computing. On the other hand, examiners in Computers and Communications examine more diverse applications, which is consistent with a more random allocation of applications.[21]

We turn now to our second question — is examiner-application matching effectively random within subclass? To provide an answer, we construct a data set of all application-pairs sharing the same art unit, filing year and primary subclass, and regress a dummy for having the same examiner on two main explanatory variables: a dummy variable for having the same assignee and the Jaccard similarity index (multiplied by 100 and standardized). We use the Jaccard similarity as a measure of the residual technological specialization that is not captured by the subclasses. Conditional on the subclasses and the Jaccard similarity, the same assignee dummy should capture the agglomeration based on the identity of the assignee that is not explained by observable similarities in technology.[22] Under the null hypothesis of random assignment within primary subclass, the coefficients of the Jaccard similarity index and of the same assignee dummy should be close to zero (and statisti-

[21]Table B4 in the appendix uses the data from Frakes & Wasserman (2017) to show that examiner specialization increases with tenure at the USPTO. Details of this analysis, and additional regression results showing that specialization increases with seniority, are available from the authors on request.

[22]We limit the sample to application pairs where both applications have an assignee and art-unit-filing-year-subclasses with at least two pairs. In this sample, the mean Jaccard similarity is 0.079 (standard deviation 0.1). About 11% of the application pairs in this sample have the same examiner, and about 7% have the same assignee. The application-pairs with the same assignee are technologically very similar: the mean Jaccard similarity for application-pairs assigned to the same organization in this sample is 0.25. Also, this mean is particularly high in Biotech, where it is 0.74, and relatively lower in the computing-related TCs.

cally insignificant).[23]

Table 5 presents the results. For the full sample, the estimates in column (1) imply that having the same assignee is correlated with a 5.3 percentage point increase in the probability of having the same examiner. This is a large increase, given that 11% of the application pairs in this sample have the same examiner. The coefficient on the same assignee dummy drops to 2.5 with the addition of art-unit-filing-year-subclass fixed effects, but remains statistically significant and economically large. The Jaccard similarity index is also positively associated with the probability of having the same examiner. A one standard deviation increase in the Jaccard similarity of an application-pair is correlated with a 1 percentage point increase in the probability of having the same examiner in column (1), and a 2.6 percentage point increase in column (2).

The lower half of Table 5 shows the results by TC.[24] The main message of this set of estimates is that we still observe technological specialization within subclasses. The same assignee dummy is statistically significant at 1% and relatively large in all TCs except Biotechnology. Similarly, the coefficient of the Jaccard similarity is positive, statistically significant at least at 5% and large in all TCs.

Overall, the results in this sub-section suggest that the USPC classification is relatively more effective at capturing actual technological similarities in Biotech and Chemistry, and less in computing. At the same time, examiners in Computers and Communications do examine less similar applications, consistent with the idea that they are more "generalist" than other examiners. We also use the Jaccard similarity measure to show that even within USPC subclasses, there is evidence of examiner specialization.

4.3 Specialization and Examination Outcomes

As a final step in our empirical analysis, we explore the relationship between examiner specialization and patent examination outcomes. We focus on three

[23]We cluster the standard errors at the art-unit-filing-year level in order to avoid the issue of within-dyad correlations, as discussed in Cameron & Miller (2014).

[24]For these estimates, we standardize the Jaccard similarity index within each TC.

outcomes: (i) whether an application is granted, (ii) the change in the number of words in the first independent claim between the published application and the granted patent, and (iii) the number of days required to process the application (i.e., the difference between the date an application is docketed to an examiner for the first time and its disposal date). Our sample for this part of the analysis consists of all applications belonging to an art-unit by examiner by filing-year cell containing more than 10 applications. We also drop a small number of examiner-art-unit-filing-year groups that have only one application in the estimation sample after we dropped pending applications and those filed after year 2009 to account for truncation.

Our unit of analysis is the application, and we adopt a measure of specialization that varies across both examiners and applications. Specifically, our main explanatory variable is the share of an examiner's applications (within an art-unit-filing-year cell) having the same subclass as a focal application. To be more precise, define the set $k_{it}(j)$ of all patents (except for patent i) assigned to examiner j and filed in year t.[25] Let n_{jt} represent the total number of patents reviewed by examiner j filed in year t, and define an indicator 1_{mn} that equals one if and only if two patents (m and n) have the same subclass. Our main explanatory variable can be written as:

$$Share_{ijt} = \frac{\sum_{m \in k_{it}(j)} 1_{mi}}{n_{jt} - 1}. \tag{3}$$

Intuitively, $Share_{ijt}$ equals the probability that a random draw from the pool of applications filed in year t assigned to examiner j has the same subclass as the focal application.

All of our regressions control for application characteristics. To control for economic importance and scope, we include the natural logarithms of the number of DOCDB family members at the filing date and the number of independent claims. We also include an indicator variable for small entity status of the applicant and a dummy for applications whose first inventor has a US address to control for differences in examination related to the identity

[25]For this analysis we consider an examiner affiliated with two (or more) different art units in the same year as two (or more) examiners.

of the applicant (Carley, Hedge & Marco 2015).

Table 6 presents estimates from a series of OLS panel-data regressions that examine the correlation between $Share_{ijt}$ and examination outcomes. To ease interpretation, we standardize $Share_{ijt}$ and the outcome variables except the dummy for granted patents.[26] Standard errors are clustered at art-unit-filing-year level in all models.

Columns (1) through (3) report coefficient estimates from a within-examiner regression with art-unit-examiner-filing-year fixed effects. The coefficients of $Share_{ijt}$ are all positive but very close to zero. Columns (4) through (6) report the results from a between-examiner analysis, where we regress the mean outcome for each art-unit-examiner-filing-year on the mean of $Share_{ijt}$.

The coefficient in Column (4) indicates that a one standard deviation increase in $Share_{ijt}$ leads to a 3 percentage point drop in the grant rate. This suggests that specialized examiners are also more stringent. The coefficient in column (5) also suggests that specialization leads to more stringent examination. However, the economic magnitude of this result is rather small: a one standard deviation change in $Share_{ijt}$ produces a 0.05 standard deviation change in the number of words added to the first claim. Finally, in column (6) we find a small but statistically significant positive association between specialization and the time required to process a patent examination.[27]

The overall message of this part of the analysis is that examiner specialization is related to more stringent examination, although the economic magnitudes are not dramatic. This relationship is driven by differences across examiners, as showed by the "between" estimators in Table 6. We do not find important differences in the relationship between specialization and examination outcomes "within" examiners. One plausible explanation for the finding is that it is easier for examiners that are more specialized to find relevant prior art because they are more familiar with certain fields of tech-

[26] Table B2 displays summary statistics for all variables used in this part of the analysis.

[27] Unreported models that do not include control variables for application characteristics estimate coefficients of $Share_{ijt}$ slightly bigger in magnitude. Under random assignment, the inclusion of control variables should not affect the magnitude of the estimated coefficients.

nology, leading to narrower claims and an increased probability of application abandonment. Under random assignment, these estimates are causal. We prefer a descriptive interpretation. Nevertheless, these results confirm the importance of differences across examiners for examination outcomes.

5 Implications for Examiner Instruments

If patent applications were randomly assigned to patent examiners, then examiner characteristics — including the quality of examination and the propensity to grant — would be uncorrelated with unobserved characteristics of the application. When examiners specialize, however, their attributes and behaviors might be correlated with (or even caused by) the technologies they examine, which could lead to a violation of the exclusion restriction in instrumental variable estimates.

Readers may be tempted to conclude that the preceding paragraph "invalidates" examiner-based instrumental variables. That is not correct. The validity of the IV assumptions need to be evaluated in the context of a specific application, and in any case, are not amenable to direct testing. Rather, our findings imply that random matching cannot be invoked to justify using examiner characteristics as a general-purpose tool for identifying causal effects of the patent examination process.

On a more constructive note, our findings also suggest some ways to "fix" examiner-based instruments in settings where they would be useful. One response to the non-random matching problem is to condition on fine-grained measures of technology. While this is clearly a good idea, the results in Table 3 (correlation between examiner fixed effects and observables) suggest that even within USPC subclasses examiners and observables may be correlated, and those in Table 5 (based on the Jaccard similarity index) show that some unobserved specialization will remain. Furthermore, the results in Table 4 show that the quality of the USPC classification varies across TCs, suggesting that this strategy may be more effective in some art units and less in others.

A related issue regarding (sub)class effects concerns the time at which one observes the technological classification of applications. Table 1 shows that applications often change technological classification throughout examination. Classes and subclasses are thus partially affected by the examination process. Researchers willing to control for (sub)class effects may need to control for technological classifications at the time of filing or publication instead of relying on classifications currently available in Public PAIR or other databases that provide updated data.

The sheer number of subclasses creates an additional difficulty. Adding subclass fixed effects to an IV regression will not purge all specialization-induced technological variation from a leave-one-out grant rate or similar IV constructed at the examiner-year level.[28] But there are typically too few observations within examiner-subclass-year cells to construct the instrument at this more disaggregated level. A possible solution would be to utilize data within 3-digit classes instead of subclasses, but this will not capture part of the technological specialization happening within classes. Another solution would be to focus on large subclasses, or to group together data for subclasses from multiple years.

A second response to potential endogeneity is to check whether a particular instrument is correlated with observed differences in technology. Of course, failing to reject the null hypothesis within a particular sample does not validate the instrument — more data or better measures could lead to a different result.[29]

A third approach would be to focus on areas in which the assignment of

[28]Formally, suppose that the probability of issuance for patent i, in class c, assigned to examiner j is given by $Pr(G_{ijc} = 1) = \theta_j + \eta_c$ (i.e. an examiner-specific leniency and a subclass effect). The examiner-year level IV would then be $Z_{ijc} = E_{k \neq i}[G_{kjc}] = \theta_j + E[\eta|j]$. The last term depends on the classes assigned to examiner j, and will vary across examiners within a subclass-year.

[29]For example, Sampat & Williams (2017) regress patent grant on their measure of examiner leniency with and without subclass fixed effects, and obtain very similar point estimates (not statistically distinguishable) for the two specifications. We performed a similar exercise using our full sample of applications, along with new data on the primary USPC subclass at the time of publication, and find that adding subclass effects produces modest changes in the size of first-stage estimates – between 5 and 10 percent for the full sample. In our models, the subclass effects are always jointly significant, and the change in the first-stage coefficient is generally larger for TCs that exhibit more examiner specialization in our prior results. Table B5 presents the results of this exercise for two different IV-endogenous variable combinations.

applications is more consistent with random assignment. For example, our results suggest it is more plausible to assume that technology is uncorrelated with potential outcomes for information technology art units, while the same assumption is potentially problematic for art units in other TCs. Researchers could also exploit differences in assignment practices even within TCs. For example, Feng & Jaravel (2017) use agglomeration tests to identify art units in which applications are often assigned to examiners based on the last digit of the application, and so assignment practices may be less related to relevant application characteristics.

To sum up, we propose the following implications for those who still wish to use examiner characteristics as instruments, perhaps because (like us) they see the approach as a clear step forward in terms of measuring the causal impacts of intellectual property. First, it is important to carefully control for any observable differences in technology. Class or subclass fixed effects are not a panacea, but they are nevertheless a step in the right direction. Controlling for identity of the applicant, for example with assignee fixed effects, or for different groups of applicants (e.g. small entities or foreign applicants), may also help to capture technological heterogeneity and other characteristics. Second, it may be important to take into account specialization also in the computation of the grant rates or other examiner-based IVs. Third, instead of claiming that applications are randomly matched to examiners, authors should clearly explain the key identification assumption: conditional on observables, examiner characteristics must be uncorrelated with potential outcomes, regardless of any technological sorting. This assumption is not testable, and its reasonableness will vary from one application to the next. Checking that examiner characteristics are not correlated with application characteristics is an important step in assessing the credibility of the identification strategy in a given application. Finally, our results suggest it is more plausible to assume that technology is uncorrelated with potential outcomes for information technology art units, while the same assumption is potentially problematic for art units in other TCs. Researchers may also explore heteroegeneity of assignment practices across art units to understand in what contexts the key identification assumption is more credible.

6 Conclusions

We study a key stage of patent examination: the assignment of applications to examiners. We first focus on characterizing the degree of examiner specialization. Using two statistical tests designed to study industry agglomeration, we find strong evidence that examiners specialize in particular subclasses, even within relatively homogeneous art units. The degree of specialization varies across fields, with examiners in the Computers and Communications area exhibiting relatively little specialization compared to those in other TCs. We also use a measure of textual similarity (and the identity of the assignee) to illustrate residual technological specialization within subclasses. Finally, we show that more specialized examiners are more stringent on average — they have a lower grant rate, and produce a larger reduction in the scope of issued patents' first independent claim.

It may not seem surprising that we can reject the hypothesis of random matching between applications and examiners. After all, one reason for having a patent classification system is to help route applications to appropriate examiners. Also, it is reasonable that SPEs take into account the different skills of their examiners when they assign applications. However, several studies have argued that more-or-less random matching (within art units) provides a justification for using examiner behaviors and characteristics as instruments for examination outcomes. Our findings do not invalidate this identification strategy – patent examiner characteristics might still satisfy the relevant exclusion restrictions – but they do imply that we cannot rely on purely random assignment to justify the approach.

On a more positive note, our results provide some evidence on how the USPTO balances the demands of efficiency and fairness. Technological specialization is likely to be efficient. Fairness can be achieved by enforcing uniform examination standards, which is difficult, or through random assignment, which guarantees all applicants an equal shot at the more "friendly" examiners. Examiner assignment appears relatively more (but not entirely) random in the computer-related TCs. And even without controlling for technology, there is little evidence that certain examiners within a given art unit are assigned to the applications with the largest families or broadest claims. We leave to future researchers the question of whether the current level of procedural fairness to applicants is also the best policy in terms of social welfare.

Finally, our analysis opens up several avenues for research into the causes and consequences of how work is organized at the USPTO. For example, it might be interesting to study how specialization varies over time for individual examiners, or within a given art unit. Future research could also study additional outcomes, such as re-examinations, re-issues or invalidation rates of patents issued by examiners with different levels of specialization to better understand the consequences of specialization for patent quality.

References

Arts, S., Cassiman, B. & Gomez, J. C. (2018), 'Text Matching to Measure Patent Similarity', *Strategic Management Journal* **39**(1), 62–84.

Breschi, S., Lissoni, F. & Malerba, F. (2003), 'Knowledge-relatedness in firm technological diversification', *Research policy* **32**(1), 69–87.

Cameron, C. & Miller, D. (2014), 'Robust inference for dyadic data', *Workign Paper* .

Carley, M., Hedge, D. & Marco, A. (2015), 'What is the probability of receiving a us patent', *Yale JL & Tech.* **17**(1), 203.

Cockburn, I. M., Kortum, S. & Stern, S. (2002), 'Are All Patent Examiners Equal? The Impact of Examiner Characteristics', *NBER Working Paper 8980* .

Farre-Mensa, J., Hegde, D. & Ljungqvist, A. (2017), 'What is a Patent Worth? Evidence from the U.S. Patent "Lottery"', *NBER Working Paper 23268* .

Feng, J. & Jaravel, X. (2017), 'Crafting Intellectual Property Rights: Implications for Patent Assertion Entities, Litigation, and Innovation', *Working Paper* .

Frakes, M. D. & Wasserman, M. F. (2017), 'Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Microlevel Application Data', *Review of Economics and Statistics* **99**(3), 550–563.

Gaulé, P. (2015), 'Patents and the Success of Venture-Capital Backed Startups: Using Examiner Assignment to Estimate Causal Effects', *Working Paper* .

Graham, S. J. H., Marco, A. C. & Miller, R. (2015), 'The USPTO Patent Examination Research Dataset: A Window on the Process of Patent Examination', *Working paper* .

Harhoff, D., Scherer, F. M. & Vopel, K. (2003), 'Citations, Family Size, Opposition and the Value of Patent Rights', *Research policy* **32**(8), 1343–1363.

Imbens, G. W. (2015), 'Matching Methods in Practice: Three Examples', *Journal of Human Resources* **50**(2), 373–419.

Kuhn, J. M. (2016), 'The Right to Exclude: Property Rights and Frictions in the Sale of Patents', *Working Paper* .

Kuhn, J. M. & Thompson, N. C. (2017), 'How to Measure and Draw Causal Inferences with Patent Scope', *Working Paper* .

Kullback, S. & Leibler, R. A. (1951), 'On Information and Sufficiency', *Annals of Mathematical Statistics* **22**(1), 79–86.

Lemley, M. A. & Sampat, B. (2012), 'Examiner Characteristics and Patent Office Outcomes', *Review of Economics and Statistics* **94**(3), 817–827.

Marco, A. C., Myers, A. F., Graham, S., D'Agostino, P. & Kucab, J. (2015), 'The USPTO Patent Assignment Dataset: Descriptions, Lessons, and Insights', *Working paper* .

Marco, A. C., Sarnoff, J. D. & deGrazia, C. (2016), 'Patent Claims and Patent Scope', *Working paper* .

Merges, R. P. (1999), 'As Many as Six Impossible Patents before Breakfast: Property Rights for Business Concepts and Patent System Reform', *Berkeley Technology Law Journal* **14**(2), 577–615.

Mori, T., Nishikimi, K. & Smith, T. E. (2005), 'A Divergence Statistic for Industrial Localization', *Review of Economics and Statistics* **87**(4), 635–651.

Putnam, J. (1996), 'The Value of International Patent Protection', *Unpublished Ph. D. Dissertation, Yale University* .

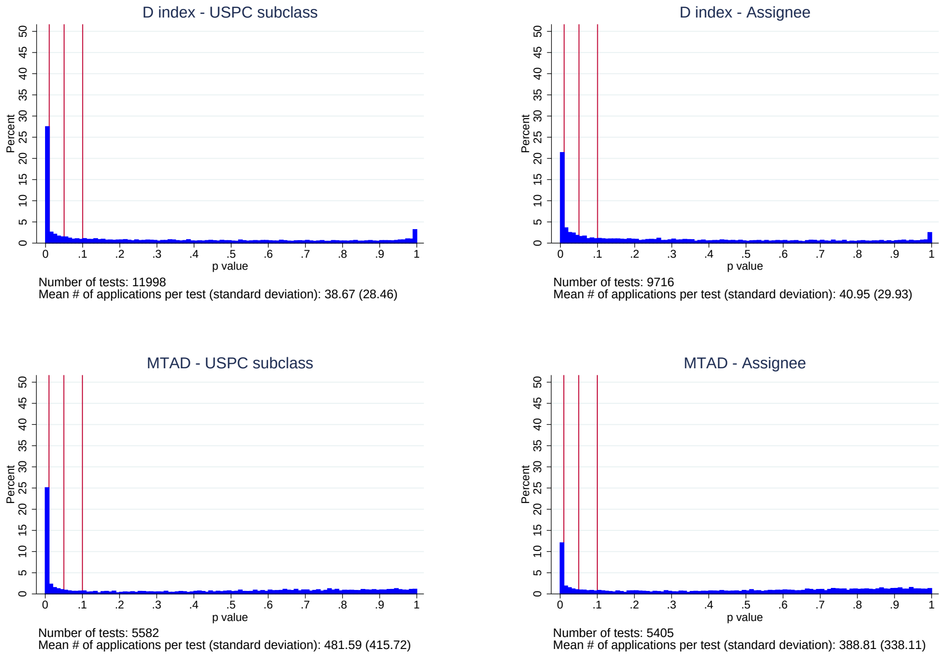
Rysman, M. & Greenstein, S. (2005), 'Testing for Agglomeration and Dispersion', *Economics Letters* **86**(3), 405–411.

Sampat, B. & Williams, H. L. (2017), 'How Do Patents Affect Follow-on Innovation? Evidence from the Human Genome', *NBER Working Paper 21666* .

Thoma, G., Torrisi, S., Gambardella, A., Guellec, D., Hall, B. H. & Harhoff, D. (2010), 'Harmonizing and Combining Large Datasets–An Application to Firm-Level Patent and Accounting Data', *NBER Working Paper 15851* .

Tables and Figures

Figure 1: Distribution of P-values from D-index (Top) and MTAD (Bottom) for USPC Subclass and Assignee



Distribution of p-values of D-index and MTAD analysis for USPC subclass and Assignee (cleaned and standardized). Tests on subsamples with more than 20 applications for D-index and 50 applications for MTAD. Vertical red lines are standard thresholds for statistical significance (0.01, 0.05 and 0.10)

Table 1: Summary Statistics by Technology Center

	Art Units	Examiners	Classes	Sub-Classes	Applications	App pairs*	App pairs, same subclass*	Class Changed†	Subclass Changed†
Biotechnology (1600)	57	1,013	268	11,530	221,586	51.1M	2.5M	32.3	75.3
Chemicals (1700)	75	1,377	411	34,509	367,371	166.6M	1.2M	31.3	77.7
Comp/Comm (2100)	79	1,733	303	7,314	208,102	31.1M	1.2M	23.4	70.7
Comp/Comm (2400)	77	1,289	159	4,492	157,852	23.7M	0.8M	18.5	59.7
Comp/Comm (2600)	82	2,046	310	14,393	338,088	208.0M	3.7	15.4	65.7
Electrical (2800)	80	2,161	382	31,815	637,929	321.0M	4.3M	17.5	65.2
Miscellaneous (3600)	75	1,617	410	38,760	360,691	114.5M	2.0M	16.8	68.8
Mechanical (3700)	65	1,949	411	40,240	425,413	180.5M	1.9M	18.6	68.9
Full sample	590	12,338	452	119,448	2,717,032	1,096.4M	17.6M	20.4	68.5

*Application pairs with same art unit, filing year (and subclass) with at least 10 unique keywords for both applications. †Percent of applications with change in (sub)class conditional on grant before July 21, 2015. Classification data for published applications and granted patents from PatentsView. Full sample counts remove duplicates across TCs. Abbreviated TC names from Graham et al. (2015). Full names of the TCs currently responsible for examination of utility patent applications are:

- 1600 - Biotechnology and Organic Chemistry
- 1700 - Chemical and Materials Engineering
- 2100 - Computer Architecture, Software, and Information Security
- 2400 - Computer Networks, Multiplex communication, Video Distribution, and Security
- 2600 - Communications
- 2800 - Semiconductors, Electrical and Optical Systems and Components
- 3600 - Transportation,Construction, Electronic Commerce, Agriculture, National Security and License & Review
- 3700 - Mechanical Engineering, Manufacturing, Products

Table 2: D-index and MTAD Tests within Art-Unit-Filing-Year (Share Rejecting Random Allocation at 1% Significance Level, by Technology Center)

Panel A: USPC subclass

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Biotechnology and Organic Chemistry	32.9	906	19.6	0.2	551
Chemical and Materials Engineering	58.6	814	55.8	0.0	721
Computer Architecture, Software, and Security	2.2	1,170	0.7	0.0	723
Computer Networking and Video Distribution	6.5	753	0.8	0.0	628
Communications	17.7	2,268	16.7	0.0	694
Semiconductors, Electrical and Optical Systems	37.4	3,389	39.5	0.0	843
Miscellaneous†	15.4	1,162	21.7	0.1	742
Mechanical Engineering, Manufacturing, Products	38.9	1,536	39.4	0.0	680
All tests	27.5	11,998	25.0	0.0	5,582

Panel B: Assignee

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Biotechnology and Organic Chemistry	50.2	225	9.1	0.0	527
Chemical and Materials Engineering	46.1	866	30.9	0.0	699
Computer Architecture, Software, and Security	4.2	970	0.0	0.0	709
Computer Networking and Video Distribution	5.3	509	0.2	0.0	616
Communications	11.0	1,879	6.3	0.0	668
Semiconductors, Electrical and Optical Systems	19.6	3,360	15.0	0.1	824
Miscellaneous†	29.3	818	13.4	0.0	703
Mechanical Engineering, Manufacturing, Products	36.0	1,089	19.1	0.0	659
All tests	21.4	9,716	12.0	0.0	5,405

Panel C: Top Ventile of DOCDB Family Size

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Biotechnology and Organic Chemistry	0.8	772	3.1	0.0	549
Chemical and Materials Engineering	3.7	1,018	5.4	0.0	716
Computer Architecture, Software, and Security	0.5	860	2.1	0.0	723
Computer Networking and Video Distribution	0.3	742	1.1	0.0	627
Communications	2.5	1,011	4.5	0.0	690
Semiconductors, Electrical and Optical Systems	3.8	1,427	6.1	0.0	841
Miscellaneous†	2.4	1,149	4.7	0.0	738
Mechanical Engineering, Manufacturing, Products	3.9	1,089	7.2	0.0	678
All tests	2.5	8,068	4.4	0.0	5,562

Panel D: Bottom Ventile of Words in 1st Claim

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Chemical and Materials Engineering	3.5	1,129	5.1	0.0	721
Computer Architecture, Software, and Security	0.0	895	0.1	0.0	723
Computer Networking and Video Distribution	0.0	755	0.0	0.0	627
Communications	0.2	1,052	0.1	0.0	693
Semiconductors, Electrical and Optical Systems	2.1	1,524	5.1	0.0	843
Miscellaneous†	0.7	1,194	1.8	0.0	741
Mechanical Engineering, Manufacturing, Products	0.8	1,160	0.7	0.0	679
All tests	1.2	7,709	2.0	0.0	5,027

For D-index, columns labelled "Rej." report the share of tests that reject the null hypothesis of equality between the observed and the reference distribution at 1% level. For MTAD, columns labelled "Agg." ("Disp.") report the share of tests that reject the null hypothesis of random allocation at 1% level in favor of agglomeration (dispersion). All tests are conducted within art-unit-filing-year cells with more than 20 applications for the D-index and more than 50 applications for MTAD. † Miscellaneous = "Transportation, Construction, Electronic Commerce, Agriculture, National Security and License & Review."

Table 3: Regressions of DOCDB Family Size and Words in 1st Claim against Examiner Fixed Effects: Share of F-tests Rejecting Null Hypothesis at 1% Significance Level, by Technology Center.

Panel A: DOCDB Family Size

	Examiner FEs		Examiner FEs, controlling for subclass	
Technology Center	Rej.	N	Rej.	N
Biotechnology (1600)	22.4	509	6.5	489
Chemicals (1700)	31.8	651	6.3	443
Comp/Comm (2100)	12.6	666	8.3	640
Comp/Comm (2400)	23.7	566	15.3	561
Comp/Comm (2600)	29.8	568	17.7	491
Electrical (2800)	26.8	776	10.4	657
Miscellaneous (3600)	24.4	676	8.4	585
Mechanical (3700)	34.0	642	12.3	575

Panel B: Words in 1st Claim

	Examiner FEs		Examiner FEs, controlling for subclass	
Technology Center	Rej.	N	Rej.	N
Chemicals (1700)	32.6	651	7.0	442
Comp/Comm (2100)	5.6	665	3.3	639
Comp/Comm (2400)	3.7	566	2.9	561
Comp/Comm (2600)	12.9	568	5.9	491
Electrical (2800)	27.4	776	8.2	658
Miscellaneous (3600)	13.3	676	3.9	586
Mechanical (3700)	22.4	642	4.2	575

We run a separate OLS regression for the applications in each art-unit-filing-year group, using the number of DOCDB family members at filing (Panel A) or the number of words in the 1st independent claim of the application (Panel B) as outcome variable. For estimation we use only applications assigned to examiners associated with at least 10 applications, and art-unit-filing-year groups with at least 50 applications and at least two examiners. The first set of models includes only the examiner fixed effects as explanatory variables ("Examiner FEs"). In the second set of models, we also include USPC subclass effects ("Examiner FEs, cotrolling for subclass"). We exclude models in which some examiner fixed effects are collinear with the subclass effects. Columns labeled "Rej." report the share of F-tests that reject the null hypothesis that all the examiner fixed effects are equal to zero at 1% level. Columns labeled "N" report the number of tests.

Table 4: Mean Jaccard Similarity by Technology Center

	(1)	(2)	(3)	(4)	(5)	(6)
Technology Center	Mean Jaccard Similarity	Mean Jaccard Similarity Same Subclass	% Δ Different - Same Subclass	Mean Jaccard Similarity Same Examiner	% Δ Different - Same Examiner	Mean Jaccard Similarity Same Subclass & Examiner
Biotechnology (1600)	4.9	11.4	149	5.8	19	15.5
Chemicals (1700)	2.9	7.3	152	4.4	55	8.7
Comp/Comm (2100)	4.4	5.5	27	4.6	6	6.5
Comp/Comm (2400)	4.6	6.1	33	4.8	5	7.1
Comp/Comm (2600)	4.5	6.7	52	4.9	10	7.8
Electrical (2800)	4.5	7.8	74	5.2	16	8.2
Miscellaneous (3600)	3.3	5.6	73	3.7	16	6.7
Mechanical (3700)	3.2	6.6	109	3.8	21	8.1

We use a sample of patent application pairs to compute the statistics reported in this table. See main text for the definition of the Jaccard similarity. We compute the Jaccard similarity between applications for all application pairs within art-unit-filing-year whose number of keywords is at least 10 for both applications (1,096,408,105 pairs), and multiply it by 100 for easier interpretation. Column (1) reports the mean Jaccard similarity for pairs in each TC. Columns (2), (4) and (6) report the mean Jaccard similarity for pairs that have the same subclass, the same examiner, and both the same subclass and the same examiner, respectively. Columns (3) and (5) report the percentage changes in Jaccard similarity going from different subclass (examiner) to same subclass (examiner). T-tests for the difference in means between application-pairs with the same subclass (examiner) and those with different subclass (examiner) are statistically significant at 1% level in all TCs.

Table 5: Probability of Having the Same Examiner within USPC Subclasses

Outcome		1[Same Examiner]*100	
Technology Center	Regressors	(1)	(2)
Full sample	Same assignee	5.30	2.50
		(0.37)	(0.11)
	Jaccard Similarity*100	0.96	2.56
		(0.40)	(0.16)
	Art-unit-year-subclass FEs		✓
Biotechnology (1600)	Same assignee	2.27	1.35
		(2.47)	(1.29)
	Jaccard Similarity*100	1.37	4.94
		(0.64)	(0.45)
Chemicals (1700)	Same assignee	12.60	3.88
		(2.46)	(0.47)
	Jaccard Similarity*100	3.17	3.23
		(0.37)	(0.15)
Comp/Comm (2100)	Same assignee	3.01	2.74
		(0.25)	(0.21)
	Jaccard Similarity*100	1.88	1.72
		(0.09)	(0.08)
Comp/Comm (2400)	Same assignee	4.82	4.52
		(0.34)	(0.30)
	Jaccard Similarity*100	1.84	1.73
		(0.10)	(0.09)
Comp/Comm (2600)	Same assignee	3.08	2.39
		(0.29)	(0.21)
	Jaccard Similarity*100	1.00	1.15
		(0.14)	(0.07)
Electrical (2800)	Same assignee	3.73	1.66
		(0.41)	(0.13)
	Jaccard Similarity*100	0.45	1.25
		(0.16)	(0.08)
Miscellaneous (3600)	Same assignee	5.45	2.72
		(0.48)	(0.28)
	Jaccard Similarity*100	1.79	1.67
		(0.55)	(0.30)
Mechanical (3700)	Same assignee	2.86	2.76
		(0.52)	(0.26)
	Jaccard Similarity*100	3.47	2.77
		(0.23)	(0.14)

Unit of observation is a patent application pair. In all models, the outcome is an indicator variable equal to one if applications in the pair have the same examiner, multiplied by 100 for easier interpretation of the coefficients. The two main explanatory variables are the Jaccard similarity between keywords of the applications in the pair (multiplied by 100 and standardized within estimation subsample) and an indicator variable equal to 1 if the applications in the pair have the same assignee. Column (2) also includes art-unit-filing-year-subclass fixed effects. The sample includes application pairs from our agglomeration analysis sample (i) examined by the same art unit, (ii) filed in the same year, (iii) classified in the same primary subclass, (iv) whose number of keywords utilized for the computation of the Jaccard similarity is at least 10 for both applications, and (v) with an assignee for both applications. We exclude from the sample art-unit-filing-year-subclasses with just one pair. The full sample contains about 11.7 million patent application pairs. All models estimated by OLS. Robust standard errors in parenthesis, clustered at art-unit-filing-year level.

Table 6: Examiner Specialization and Examination Outcomes.

Model	Within Examiner			Between Examiner		
Outcome	Granted	Words	Days	Granted	Words	Days
	(1)	(2)	(3)	(4)	(5)	(6)
$Share_{ijt}$	0.00***	0.00	0.01***	-0.03***	0.05***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Application characteristics	✓	✓	✓	✓	✓	✓
Art-unit-year-examiner FEs	✓	✓	✓			
Observations	1,746,664	1,069,219	1,746,444	48,950	43,490	48,950
Art-unit-year-examiners	48,950	43,490	48,950			

All models estimated with OLS. Unit of observation is a patent application for the within regressions and an art-unit-filing-year-examiner for the between regressions. Sample contains all applications in our primary analysis sample belonging to an art-unit by examiner by filing-year cell containing more than 10 applications. We also drop a small number of examiner-art-unit-filing-year groups that have only one application in the estimation sample after excluding pending applications and those filed after year 2009 to account for truncation. Outcome in column (1) is an indicator variable equal to one for granted applications. Outcome in column (2) is the change in the number of words in the first independent claim between the published application and the granted patent (standardized). Outcome in column (3) is the difference in days between the date an application is docketed to an examiner for the first time and its disposal date (standardized). See the main text for the definition of $Share_{iijt}$. Also this variable is standardized. All models also include control variables for application characteristics, including the natural logarithms of DOCDB family size at filing date and number of independent claims in the published application, and indicator variables for applications filed by small entities and applications whose first inventor has a US address. The mean of the outcome of the regression in column (1) is 0.65. Between regressions in columns (4)-(6) estimated on the group means. Standard errors clustered by art-unit-filing-year in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Appendix A: D-index and MTAD Statistics

Divergence Index

Suppose we have a set of applications characterized by category $i \in \mathbf{I} = \{1, ..., I\}$, assigned to a set of examiners denoted by $r \in \mathbf{R} = \{1, ..., R\}$. In our application, the categories i may correspond to USPC subclasses, assignees or any other predetermined observable characteristic of a patent application. Under random allocation, examiner r's share of all applications from category i should equal her share of the overall population.

To formalize that idea, define n_{ir} as the number of applications in category i assigned to examiner r, and $N_i = \sum_{r=1}^{R} n_{ir}$ as the total number of applications in category i. The reference distribution $p_0 = (p_{0r} : r \in \mathbf{R})$, where $p_{0r} = \frac{\sum_{i=1}^{I} n_{ir}}{\sum_{i=1}^{I} N_i}$ measures examiner r's share of all applications, is the share we expect her to be allocated from each category under the null of random assignment.

Let p_{ir} denote the true probability that a randomly sampled application in category i is assigned to examiner r, so the distribution across examiners for the category is $p_i = (p_{ir} : r \in \mathbf{R})$. We can measure the divergence between p_i and p_0 using the relative entropy of p_i with respect to p_0, called the D-index by Mori et al. (2005):

$$D(p_i|p_0) = \sum_{r \in \mathbf{R}} p_{ir} \ln\left(\frac{p_{ir}}{p_{0r}}\right).$$

$D(p_i|p_0)$ is nonnegative, achieves its minimum at $p_i = p_0$ and its local maxima when all applications in category i are assigned to a single examiner.

To estimate the D-index, we use the observed data to estimate the probabilities p_{ir}, with $\hat{p}_{ir} = \frac{n_{ir}}{N_i}$, thus estimating:

$$D(\hat{p}_i|p_0) = \sum_{r \in \mathbf{R}} \hat{p}_{ir} \ln\left(\frac{\hat{p}_{ir}}{p_{0r}}\right).$$

These probability estimates converge to the true value exponentially fast with the increase in sample size for a given category N_i.

As shown by Mori et al. (2005) the D-index can be related to the the log likelihood ratio (λ):

$$-\frac{\ln \lambda}{N_i} = \sum_{r \in \mathbf{R}} \frac{n_{ir}}{N_i} \ln\left(\frac{\hat{p}_{ir}}{p_{0r}}\right) = D(\hat{p}_i|p_0).$$

Given that $-2\ln\lambda$ is distributed asymptotically as a chi-square with $R-1$ degrees of freedom, we can use this relationship to test the null hypothesis that $p_i = p_0$. In practice, we compute $2N_i D(\hat{p}_i|p_0)$ and use it for a chi-square test with $R-1$ degrees of freedom.

Multinomial Test for Agglomeration and Dispersion

To provide a brief formal description of MTAD, we adapt the notation provided in Rysman & Greenstein (2005). Suppose we have R examiners, each receiving n_r applications, with $r = 1, ..., R$. The variable n_r is bounded between $\underline{n} = 0$ and $\overline{n} = \infty$ and distributed according to the discrete distribution $f(n_r)$. Each examiner can be assigned applications of c types. The unconditional probability of being assigned type c is p_c for $c = 1, ..., C$. The observed number of applications of type c assigned to examiner r is x_r^c. Define $\mathbf{x}_r$ as the vector of elements $x_r^1, ..., x_r^C$, $\mathbf{p}$ as the vector of probabilities $p_1, ..., p_C$, $\mathbf{n}$ as the $R \times 1$ vector of applications assigned to each examiner, and $\mathbf{X}$ as the $R \times C$ matrix of allocations. If examiners are assigned applications independently, the likelihood of observing outcome $\mathbf{x}_r$ for examiner r is the multinomial pdf

$$\mathcal{L}(\mathbf{x}_r, n_r, \mathbf{p}) = \binom{n_r}{x_r^1, ..., x_r^C} p_1^{x_r^1} ... \ p_C^{x_r^C}$$

and the average log-likelihood for the data is

$$l(\mathbf{X}, \mathbf{n}, \mathbf{p}) = \frac{1}{R} \sum_{r=1}^{R} \ln \Big(\mathcal{L}(\mathbf{x}_r, n_r, \mathbf{p}) \Big).$$

We want to compare this log-likelihood with the value we would observe under independent random assignment. Let the random variable $l(f, \mathbf{p})$ be distributed according to the distribution $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ if $\mathbf{X}$ was *actually* drawn from a multinomial distribution and n_r was drawn from f. Then the expected log-likelihood under random allocation is given by

$$E[l(f, \mathbf{p})] = \sum_{n_r} \Bigg(\sum_{\mathbf{z} \in \Phi(n_r)} \ln \mathcal{L}(\mathbf{z}, n_r, \mathbf{p}) \times \mathcal{L}(\mathbf{z}, n_r, \mathbf{p}) \Bigg) f(n_r)$$

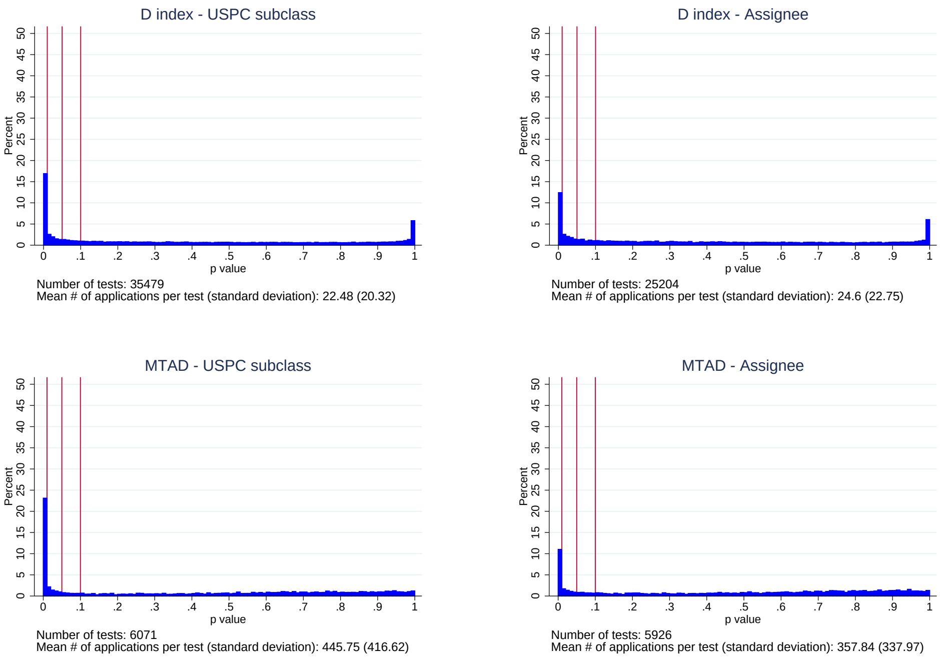
where $\Phi(n_r)$ is the set of all possible allocations of the n_r applications. To compute $E[l(f, \mathbf{p})]$ we treat $\mathbf{p}$ as known and take f to be the empirical distribution of n_r. The MTAD test-statistic is

$$t(\mathbf{X}, \mathbf{n}, \mathbf{p}) = l(\mathbf{X}, \mathbf{n}, \mathbf{p}) - E[l(f, \mathbf{p})]. \tag{4}$$

A negative (positive) value of $t(\mathbf{X}, \mathbf{n}, \mathbf{p})$ signals agglomeration (dispersion) of patent applications compared to the null of random assignment. This statistic is distributed asymptotically normal and we use simulation to generate its confidence intervals.

Appendix B: Additional Tables and Figures (Online Publication Only)

Figure B1: Distribution of P-values from D-index (Top) and MTAD (Bottom) for USPC Subclass and Assignee (Lower Thresholds)



Distribution of p-values of D-index and MTAD analysis for USPC subclass and Assignee (cleaned and standardized). Tests on subsamples with more than 10 applications for D-index and 25 applications for MTAD. Vertical red lines are standard thresholds for statistical significance (0.01, 0.05 and 0.10)

Table B1: Summary Statistics for Sample of Applications

Panel A: categorical variables

Variable	# of categories	Applications per category Mean	Std dev	Min	5th percentile	1st quartile	Median	3rd quartile	95th percentile	Max
Examiners	12,338	220.22	227.54	1	2	48	156	313	714	1,655
Art units	590	4,605.14	3,942.42	3	446	2,007	3,228.50	6,157	13,459	21,905
Subclasses	119,448	22.75	106.65	1	1	2	5	14	85	13,836
Assignees	164,195	12.99	301.16	1	1	1	1	3	19	59,998

Panel B: quantitative variables

Variable	N	Mean	Std dev	Min	5th percentile	1st quartile	Median	3rd quartile	95th percentile	Max
DOCDB family size	2,716,195	2.88	5.66	1	1	1	2	3	8	378
Words in 1st claim	2,712,367	124.95	128.00	1	35	70	103	151	269	46,194

The number of applications characterized by a big DOCDB family and a low number of words in the first independent claim are respectively 106,408 and 116,665.

Table B2: Summary Statistics for Examiners' Specialization and Examination Outcomes.

Variable	N	Mean	Std dev	Min	Median	Max
$Share_{ijt}$	1,750,188	0.04	0.09	0.00	0.00	1.00
Granted	1,750,188	0.65	0.48	0.00	1.00	1.00
Days	1,749,967	918.09	510.19	0.00	826.00	17,835.00
Change in words	1,069,829	49.16	87.56	-10,351.00	30.00	9,248.00
DOCDB family at filing	1,750,165	2.80	3.86	1.00	2.00	323.00
Independent claims	1,746,787	3.11	3.00	1.00	3.00	620.00
Small entity	1,750,188	0.24	0.43	0.00	0.00	1.00
US fist inventor	1,750,067	0.31	0.46	0.00	0.00	1.00

Table B3: D-index and MTAD Tests within Art-Unit-Filing-Year (Share Rejecting Random Allocation at 5% Significance Level, by Technology Center)

Panel A: USPC subclass

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Biotechnology and Organic Chemistry	43.8	906	28.1	0.2	551
Chemical and Materials Engineering	66.3	814	64.6	0.0	721
Computer Architecture, Software, and Security	7.1	1,170	1.4	0.1	723
Computer Networking and Video Distribution	14.9	753	2.1	0.2	628
Communications	22.5	2,268	20.6	0.0	694
Semiconductors, Electrical and Optical Systems	46.5	3,389	47.1	0.1	843
Miscellaneous†	23.2	1,162	28.8	0.3	742
Mechanical Engineering, Manufacturing, Products	49.0	1,536	49.3	0.0	680
All tests	35.3	11,998	31.0	0.1	5,582

Panel B: Assignee

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Biotechnology and Organic Chemistry	65.8	225	13.3	0.0	527
Chemical and Materials Engineering	56.0	866	40.3	0.0	699
Computer Architecture, Software, and Security	11.3	970	0.3	0.0	709
Computer Networking and Video Distribution	15.5	509	0.3	0.0	616
Communications	17.0	1,879	10.0	0.1	668
Semiconductors, Electrical and Optical Systems	31.4	3,360	22.6	0.1	824
Miscellaneous†	40.2	818	21.1	0.0	703
Mechanical Engineering, Manufacturing, Products	51.1	1,089	29.1	0.0	659
All tests	31.7	9,716	17.6	0.0	5,405

Panel C: Top Ventile of DOCDB Family Size

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Biotechnology and Organic Chemistry	2.8	772	6.4	0.0	549
Chemical and Materials Engineering	6.9	1,018	9.4	0.0	716
Computer Architecture, Software, and Security	0.8	860	4.4	0.0	723
Computer Networking and Video Distribution	0.7	742	2.6	0.0	627
Communications	4.0	1,011	6.5	0.0	690
Semiconductors, Electrical and Optical Systems	7.6	1,427	10.5	0.0	841
Miscellaneous†	5.3	1,149	8.5	0.0	738
Mechanical Engineering, Manufacturing, Products	8.2	1,089	11.9	0.0	678
All tests	5.0	8,068	7.7	0.0	5,562

Panel D: Bottom Ventile of Words in 1st Claim

	D-index		MTAD		
Technology Center	Rej.	N	Agg.	Disp.	N
Chemical and Materials Engineering	7.2	1,129	9.7	0.0	721
Computer Architecture, Software, and Security	0.1	895	1.5	0.0	723
Computer Networking and Video Distribution	0.0	755	0.0	0.0	627
Communications	0.5	1,052	0.9	0.0	693
Semiconductors, Electrical and Optical Systems	5.0	1,524	8.8	0.1	843
Miscellaneous†	3.4	1,194	4.2	0.1	741
Mechanical Engineering, Manufacturing, Products	3.2	1,160	4.7	0.1	679
All tests	3.1	7,709	4.5	0.1	5,027

For D-index, columns labelled "Rej." report the share of tests that reject the null hypothesis of equality between the observed and the reference distribution at 5% level. For MTAD, columns labelled "Agg." ("Disp.") report the share of tests that reject the null hypothesis of random allocation at 5% level in favor of agglomeration (dispersion). All tests are conducted within art-unit-filing-year cells with more than 20 applications for the D-index and more than 50 applications for MTAD. † Miscellaneous = "Transportation, Construction, Electronic Commerce, Agriculture, National Security and License & Review."

Table B4: Mean Jaccard Similarity by GS-level.

GS-level	(1) Mean Jaccard	(2) Mean Jaccard, same subclass
GS-5	4.54	7.90
GS-7	4.50	8.15
GS-9	4.54	7.91
GS-11	4.55	8.08
GS-12	4.63	8.44
GS-13	4.60	8.52
GS-14	4.68	10.18
GS-15	4.70	11.20

For this table, we match the data utilized in Table 4 with data on the GS-level of examiners in years between 2000 and 2012 (inclusive) utilized by Frakes & Wasserman (2017). We focus on application-pairs with the same filing year examined by the same examiner. Column (1) reports the mean Jaccard similarity for the pairs examined by examiners at different GS-levels (43,252,367 pairs). Column (2) reports the mean Jaccard similarity for the pairs that have the same subclass examined by examiners at different GS-levels (1,408,830 pairs). Examiners at GS-14 level are "primary" examiners. Examiners at GS-15 level are SPEs.

Table B5: IV "First-Stage" With and Without USPC Subclass Effects

Outcome†	1[Granted]*100			log(Words-in-1^{st}-claim)		
Potential Instrument	Leave-one-out grant rate *100			Leave-one-out scope change		
	(1)	(2)	(1)/(2)	(3)	(4)	(3)/(4)
Full Sample	0.75	0.68	0.90	0.37	0.35	0.94
	(0.00)	(0.00)		(0.01)	(0.01)	
Art-unit-year FEs	✓			✓		
Art-unit-year-subclass FEs		✓			✓	
Biotechnology (1600)	0.69	0.61	0.89			
	(0.01)	(0.01)				
Chemicals (1700)	0.80	0.66	0.83	0.31	0.18	0.57
	(0.01)	(0.01)		(0.02)	(0.02)	
Comp/Comm (2100)	0.66	0.64	0.98	0.32	0.32	1.00
	(0.01)	(0.01)		(0.01)	(0.01)	
Comp/Comm (2400)	0.56	0.53	0.95	0.35	0.34	0.98
	(0.01)	(0.02)		(0.01)	(0.02)	
Comp/Comm (2600)	0.75	0.72	0.96	0.38	0.36	0.95
	(0.01)	(0.01)		(0.01)	(0.01)	
Electrical (2800)	0.79	0.74	0.94	0.44	0.43	0.96
	(0.01)	(0.01)		(0.01)	(0.01)	
Miscellaneous (3600)	0.74	0.63	0.85	0.34	0.30	0.88
	(0.01)	(0.01)		(0.01)	(0.01)	
Mechanical (3700)	0.79	0.69	0.88	0.34	0.30	0.88
	(0.01)	(0.01)		(0.01)	(0.01)	

†Outcome is the endogenous variable in an instrumental variable regression: an indicator variable equal to one if an application is granted multiplied by 100 for easier interpretation of the coefficients in models (1) and (2), and the natural logarithm of the number of words in the first independent claim of a granted patent in models (3) and (4). Coefficients and standard errors for this second set of models are multiplied by 100 for easier interpretation of the results. Sample contains all applications in our primary sample that were either granted or abandoned by the end of the sample period; whose leave-one-out potential instrumental variable is computed with more than 10 applications; and whose art-unit-filing-year-subclass cell in the estimation sample contains at least two applications. The estimates for the patent scope models use only granted applications. Models in columns (1) and (3) include art-unit-filing-year fixed effects, while models in columns (2) and (4) include art-unit-filing-year-subclass fixed effects. Each "first-stage" estimate in this table comes from a separate OLS regression of Outcome on Potential Instrument for applications assigned to a given TC. Robust standard errors, clustered by art-unit-filing-year, in parentheses. All estimates are statistically significant at the 1% level. We exclude Biotechnology patents (TC 1600) from the second set of estimates because Kuhn & Thompson (2017) suggest that counting words in the first claim does not yield a meaningful measure of claim-scope for those applications. To test whether differences between the estimates in columns (1) and (2) and those in columns (3) and (4) are statistically significant, we demean the leave-one-out variables within art-unit-filing-year and within art-unit-filing-year-subclass, re-run the models (without the fixed effects, as demeaning within groups at the level of the fixed effects produces the same coefficients) and test the statistical significance of the differences in the coefficients. All tests but those for the patent scope models in TCs 2100 and 2400 are statistically significant at 1% level. We also run a battery of likelihood ratio tests to compare models analogous to those in columns (1)-(2) and (3)-(4) without clustering the standard errors. All likelihood ratio tests are statistically significant at the 1% level.